

**A THREE-STEP METHOD FOR CHOOSING THE
NUMBER OF BOOTSTRAP REPETITIONS**

BY

DONALD W. K. ANDREWS and MOSHE BUSHINSKY

COWLES FOUNDATION PAPER NO. 1001



**COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY**

**Box 208281
New Haven, Connecticut 06520-8281**

2000

<http://cowles.econ.yale.edu/>

A THREE-STEP METHOD FOR CHOOSING THE NUMBER OF BOOTSTRAP REPETITIONS

BY DONALD W. K. ANDREWS AND MOSHE BUCHINSKY¹

This paper considers the problem of choosing the number of bootstrap repetitions B for bootstrap standard errors, confidence intervals, confidence regions, hypothesis tests, p -values, and bias correction. For each of these problems, the paper provides a three-step method for choosing B to achieve a desired level of accuracy. Accuracy is measured by the percentage deviation of the bootstrap standard error estimate, confidence interval length, test's critical value, test's p -value, or bias-corrected estimate based on B bootstrap simulations from the corresponding ideal bootstrap quantities for which $B = \infty$.

The results apply quite generally to parametric, semiparametric, and nonparametric models with independent and dependent data. The results apply to the standard nonparametric iid bootstrap, moving block bootstraps for time series data, parametric and semiparametric bootstraps, and bootstraps for regression models based on bootstrapping residuals.

Monte Carlo simulations show that the proposed methods work very well.

KEYWORDS: Bias correction, bootstrap, bootstrap repetitions, confidence interval, hypothesis test, p -value, simulation, standard error estimate.

1. INTRODUCTION

BOOTSTRAP METHODS HAVE GAINED a great deal of popularity in empirical research. Although the methods are easy to apply, determining the number of bootstrap repetitions, B , to employ is a common problem in the existing literature. Typically, this number is determined in a somewhat ad hoc manner. This is problematic, because one can obtain a “different answer” from the same data merely by using different simulation draws if B is chosen to be too small. On the other hand, it is expensive to compute the bootstrap statistics of interest, if B is chosen to be extremely large. Thus, it is desirable to be able to determine a value of B that obtains a suitable level of accuracy for a given problem at hand. This paper addresses this issue in the context of the three main branches of statistical inference, viz., point estimation, interval and region estimation, and hypothesis testing.

We provide methods for determining B to attain specified levels of accuracy for bootstrap standard error estimates, confidence intervals, confidence regions, hypothesis tests, and bias correction. The basic strategy is the same in each case.

¹The authors thank Ariel Pakes, three referees, and the co-editor for helpful comments; Glenna Ames for typing the original manuscript; and Rosemarie Lewis for proofreading the manuscript. The first author acknowledges the research support of the National Science Foundation via Grant Numbers SBR-9410975 and SBR-9730277. The second author acknowledges the research support of the National Science Foundation via Grant Number SBR-9320386 and the Alfred P. Sloan Foundation via a Research Fellowship.

We approximate the distribution of the appropriate bootstrap statistic by its asymptotic distribution as $B \rightarrow \infty$. Here we are referring to the distribution of the statistic with respect to the simulation randomness conditional on the sample. We replace unknown parameters in the asymptotic distribution by consistent estimates. Then, we determine a formula for how large B needs to be to attain a desired level of accuracy based on the asymptotic approximation. A three-step method for choosing B is proposed for each case. Three steps are required because one needs to estimate unknown parameters in the initial two steps before one can determine a suitable choice of B in the third step.

The measure of accuracy employed is the percentage deviation of the bootstrap quantity of interest based on B repetitions from the ideal bootstrap quantity, for which $B = \infty$. In particular, in the different applications considered, accuracy is measured by the percentage deviation of a bootstrap standard error estimate, confidence interval “length,” critical value of a test, p -value, or bias-corrected estimate based on B repetitions from its ideal value based on $B = \infty$. For a symmetric two-sided confidence interval, the “length” is just the distance between the lower and upper bounds of the interval. For a one-sided confidence interval, the interval has an infinite length. In this case, the “length” that we consider is the *lower* or *upper* length of the interval depending upon whether the one-sided interval provides a lower bound or an upper bound. By definition, the *lower length* of a confidence interval for a parameter θ based on a parameter estimate $\hat{\theta}$ is the distance between the lower endpoint of the confidence interval and the parameter estimate $\hat{\theta}$. The *upper length* is defined analogously. For two-sided equal-tailed confidence intervals, we consider both the lower and upper lengths of the confidence interval.

The accuracy obtained by a given choice of B is stochastic, because the bootstrap simulations are random. To determine a suitable value of B , we specify a bound on the relevant percentage deviation, denoted pdb , and we require that the actual percentage deviation be less than this bound with a specified probability, $1 - \tau$, close to one. The three-step method takes pdb and τ as given and specifies a data-dependent method of determining a value of B , denoted B^* , such that the desired level of accuracy is obtained. For example, one might take $(pdb, \tau) = (10, .05)$. Then, the three-step method yields a value B^* such that the relevant percentage deviation is less than 10% with approximate probability .95.

The three-step methods are applicable in parametric, semiparametric, and nonparametric models with independent and identically distributed (iid) data, independent and nonidentically distributed (inid) data, and time series data. The methods are applicable when the bootstrap employed is the standard nonparametric iid bootstrap, a moving block bootstrap for time series, a parametric or semiparametric bootstrap, or a bootstrap for regression models that is based on bootstrapping residuals. The methods are applicable to statistics that have normal and non-normal asymptotic distributions. Essentially, the results are applicable whenever the bootstrap samples are simulated to be iid *across*

different bootstrap samples. The simulations need not be iid within each bootstrap sample.

The results for confidence intervals apply to symmetric two-sided, equal-tailed two-sided, and one-sided *percentile t* confidence intervals, as defined in Hall (1992). Efron's (1987) BC_a confidence intervals are not considered. They are considered in Andrews and Buchinsky (1999b). The results for tests apply to a wide variety of tests of parametric restrictions and model specification based on t statistics, Wald statistics, Lagrange multiplier statistics, likelihood ratio statistics, etc.

We note that the results given here for bootstrap standard error estimates are of interest even to those who believe that the bootstrap should only be used to obtain confidence intervals or tests that exhibit higher-order accuracy. The reason is that a bootstrap standard error estimate can be used to "Studentize" a statistic in order to construct an asymptotically pivotal statistic that is the basis of a bootstrap percentile t confidence interval or test. Calculating a bootstrap confidence interval or test that employs a bootstrap standard error estimate to Studentize the statistic requires that one does a nested bootstrap, which is computationally intensive. Nevertheless, there are situations where this is the best method to use.

For bootstrap standard error estimates, the three-step method depends on an estimate of the *coefficient of excess kurtosis*, γ_2 , of the bootstrap distribution of the parameter estimator. We consider the usual estimator of γ_2 as well as a bias-corrected estimator of it. We compare these two methods via simulation in Andrews and Buchinsky (1999a). Because the computational cost of carrying out the bias correction is small and the gains are significant in some cases, we recommend use of the bias-corrected estimator of γ_2 .

The three-step methods are justified by asymptotic results. The small sample accuracy of the asymptotic results is evaluated via simulation. We assess the performance of the three-step methods for symmetric percentile t confidence intervals. More comprehensive simulation results for the standard error estimates, tests for a given significance level, and p -values are given in Andrews and Buchinsky (1999a). In short, the simulations show that the methods work very well in the cases considered.

The closest results in the literature to the standard error results given here are those of Efron and Tibshirani (1986, Sec. 9). Efron and Tibshirani provide a simple formula that relates the coefficient of variation of the bootstrap standard error estimator, as an estimate of the true standard error, to the coefficient of variation of the ideal bootstrap standard error estimator, as an estimate of the true standard error. Their formula depends on some unknown parameters that are not estimable. Hence, Efron and Tibshirani only use their formula to suggest a range of plausible values of B . An advantage of our approach over that of Efron and Tibshirani is that the unknown parameters in our approach can be estimated. This allows us to specify an explicit method of choosing B to obtain a desired degree of accuracy.

The closest results in the literature to the confidence interval results given here are those of Hall (1986). Hall considers *unconditional* coverage probabilities, i.e., coverage probabilities with respect to the randomness in the data and the bootstrap simulations. In contrast, we consider *conditional* coverage probabilities, i.e., coverage probabilities with respect to the randomness in the data conditional on the bootstrap simulations. We do so because we do not want to be able to obtain “different answers” from the same data due to the use of different simulation draws. Bootstrap simulation randomness is *ancillary* and, hence, should be considered only when making inference according to the principle of ancillarity or conditionality; see Kiefer (1982).

The closest results in the literature to the results given here for tests with a given significance level are those of Davidson and MacKinnon (1997). They propose a pretesting method of choosing B that aims to ensure that the probability is small that there is a difference between the conclusions of the ideal bootstrap test and the bootstrap test based on B bootstrap repetitions for a test with a given significance level α . The method that we consider aims to achieve a bootstrap test that has good conditional significance level and power given the simulation randomness by determining an accurate critical value. However, if desired, one can choose the bound pdb so that the method considered here has the same goal as the Davidson and MacKinnon (1997) method.

No results in the literature other than this paper discuss choosing B for p -values or for bias-correction. We prefer the use of bootstrap p -values over tests with a given significance level because they are more informative.

The remainder of this paper is organized as follows. Section 2 presents the general framework that is employed, introduces notation and definitions, and describes the applications to which the results apply. Section 3 presents formulae for the accuracy of the bootstrap estimator for finite B as an approximation to the ideal bootstrap estimator for the applications of interest. This formula is the basis of the three-step method for determining B . Section 4 introduces the three-step method for determining B . Section 5 states its asymptotic justification. Section 6 presents Monte Carlo simulation results for the three-step method for symmetric two-sided confidence intervals. An Appendix discusses the asymptotics used to justify the three-step method and provides proofs of the results given in Sections 3 and 4. It also defines a bootstrap bias-corrected estimator for the coefficient of excess kurtosis of the bootstrap estimator, which is used in the three-step method for standard error estimates.

2. APPLICATIONS OF INTEREST

2.1. *The General Framework*

The general framework is as follows. We are interested in a quantity λ . We would like to estimate λ using an “ideal” bootstrap estimate denoted $\hat{\lambda}_\infty$. In general, analytic calculation of $\hat{\lambda}_\infty$ is intractable, so we approximate it using bootstrap simulations. The bootstrap approximation of $\hat{\lambda}_\infty$ based on B bootstrap

repetitions is denoted $\hat{\lambda}_B$. Below we specify λ , $\hat{\lambda}_\infty$, and $\hat{\lambda}_B$ for each of the applications of interest. Before doing so, we introduce some notation and definitions.

The observed data are a sample of size n : $\mathbf{X} = (X_1, \dots, X_n)'$. Let $\mathbf{X}^* = (X_1^*, \dots, X_n^*)'$ be a bootstrap sample of size n based on the original sample \mathbf{X} . When the original sample \mathbf{X} is comprised of iid or inid random variables, the bootstrap sample \mathbf{X}^* often is an iid sample of size n drawn from some distribution \hat{F} . For example, for the nonparametric bootstrap, \hat{F} is the empirical distribution function based on \mathbf{X} . For parametric and semiparametric bootstraps, \hat{F} depends on estimators of some parameters.

When the original sample \mathbf{X} is comprised of dependent data, the bootstrap sample often is taken to be a moving block bootstrap or some variation of this; see Carlstein (1986), Kunsch (1989), Hall and Horowitz (1996), Li and Maddala (1996), and Andrews (1999). When the model is a regression model with independent or dependent data, the bootstrap sample is sometimes generated by bootstrapping the residuals; see Freedman (1981), Li and Maddala (1996), and the references therein. All of these bootstrap methods are covered by our results.

Let $\hat{\theta} = \hat{\theta}(\mathbf{X})$ be an estimator of a parameter θ_0 based on the sample \mathbf{X} . Let $\hat{\theta}^* = \hat{\theta}(\mathbf{X}^*)$ denote the bootstrap estimator. Let $T = T(\theta_0, \mathbf{X})$ be a test statistic based on the sample \mathbf{X} for testing the null hypothesis $H_0: \theta = \theta_0$. Let $T^* = T(\hat{\theta}, \mathbf{X}^*)$ be the bootstrap test statistic where $\tilde{\theta} = \hat{\theta}$ if \mathbf{X}^* is defined without imposing the null hypothesis, and $\tilde{\theta} = \theta_0$ if \mathbf{X}^* is defined with the null hypothesis imposed.

Let $\{\mathbf{X}_b^*: b = 1, \dots, B\}$ denote B iid bootstrap samples, each with the same distribution as \mathbf{X}^* . We note that our results are applicable in any bootstrap context in which the simulated bootstrap samples $\{\mathbf{X}_b^*: b = 1, \dots, B\}$ are iid over the index b . Let $\hat{\theta}_b^* = \hat{\theta}(\mathbf{X}_b^*)$ and $T_b^* = T(\hat{\theta}, \mathbf{X}_b^*)$, for $b = 1, \dots, B$, denote the corresponding B bootstrap estimators and test statistics. Let $\{T_{B,b}^*: b = 1, \dots, B\}$ denote the ordered sample of the bootstrap T statistics.

Let E denote expectation with respect to the randomness in \mathbf{X} . Let P^* and E^* refer to probability and expectation, respectively, with respect to the randomness in the bootstrap samples \mathbf{X}^* or $\{\mathbf{X}_b^*: b = 1, \dots, B\}$ conditional on the observed data \mathbf{X} .

We now consider the applications of interest.

2.2. Standard Errors

The first application is to bootstrap standard error estimates for a scalar estimator $\hat{\theta}$. The quantities $(\lambda, \hat{\lambda}_\infty, \hat{\lambda}_B)$ in this case are the standard error, se , of $\hat{\theta}$; the “ideal” bootstrap standard error estimator, \widehat{se}_∞ ; and the bootstrap standard error estimator based on B bootstrap repetitions, \widehat{se}_B ; respectively. See Table I for their definitions.

Note that $\lim_{B \rightarrow \infty} \widehat{se}_B = \widehat{se}_\infty$ almost surely by the law of large numbers provided $E^*(\hat{\theta}(\mathbf{X}^*))^2 < \infty$. The latter holds automatically for the nonparametric bootstrap due to its finite support.

TABLE I
VALUES OF λ , $\hat{\lambda}_{\infty}$, AND $\hat{\lambda}_B$ FOR EACH APPLICATION

Application	Quantity of Interest (λ); and Ideal Bootstrap Quantity ($\hat{\lambda}_{\infty}$)	Bootstrap Quantity Based on B Repetitions: $\hat{\lambda}_B$
1. Standard Errors	$se = (E(\hat{\theta}(\mathbf{X}) - E\hat{\theta}(\mathbf{X}))^2)^{1/2}$ $\hat{se}_{\infty} = (E^*(\hat{\theta}(\mathbf{X}^*) - E^*\hat{\theta}(\mathbf{X}^*))^2)^{1/2}$	$\hat{se}_B = \left(\frac{1}{B-1} \sum_{b=1}^B \left(\hat{\theta}_b^* - \frac{1}{B} \sum_{c=1}^B \hat{\theta}_c^* \right)^2 \right)^{1/2}$
2. Confidence Intervals, Confidence Regions, and Tests with a Given Significance Level	$q_{1-\alpha}^0 = \inf\{q : P(T \leq q) \geq 1 - \alpha\}$ $\hat{q}_{1-\alpha, \infty} = \inf\{q : P^*(T^* \leq q) \geq 1 - \alpha\}$	$\hat{q}_{1-\alpha, B} = T_{B, \nu}^*$ $= 1 - \alpha \text{ quantile of } \{T_b^* : b = 1, \dots, B\}$
3. p -values	$p = \text{exact } p\text{-value}$ $\hat{p}_{\infty} = P^*(T^* > T)$	$\hat{p}_B = \frac{1}{B} \sum_{b=1}^B 1(T_b^* > T)$
4. Bias Correction	$\hat{\theta}_{bc}^0 = \hat{\theta} - (E\hat{\theta} - \theta_0)$ $\hat{\theta}_{bc, \infty} = \hat{\theta} - (E^*\hat{\theta}^* - \hat{\theta})$	$\hat{\theta}_{bc, B} = \hat{\theta} - \left(\frac{1}{B} \sum_{b=1}^B \hat{\theta}_b^* - \hat{\theta} \right)$ $= 2\hat{\theta} - \frac{1}{B} \sum_{b=1}^B \hat{\theta}_b^*$

Notes: $q_{1-\alpha}^0$ and $\hat{q}_{1-\alpha, \infty}$ are the $1 - \alpha$ quantiles of T and T^* respectively. $T_{B, \nu}^*$ is the ν th order statistic of $\{T_b^* : b = 1, \dots, B\}$, where ν is an integer that satisfies $\nu = (B + 1)(1 - \alpha)$.

2.3. Confidence Intervals, Confidence Regions, and Tests

The second group of applications includes confidence intervals, confidence regions, and tests for a given significance level α . In each case, the quantity λ of interest is the $1 - \alpha$ quantile, denoted $q_{1-\alpha}^0$, of a test statistic $T = T(\theta_0, \mathbf{X})$ for testing the null hypothesis $H_0: \theta = \theta_0$. The statistic T is normalized such that $T \rightarrow_d G$ as $n \rightarrow \infty$, where G is some distribution function that has a unique $1 - \alpha$ quantile, denoted $q_{1-\alpha}$, and a density with respect to Lebesgue measure, denoted $g(\cdot)$, in a neighborhood of $q_{1-\alpha}$. The distribution G may depend on unknown parameters.

When T is an asymptotically pivotal test statistic, then the confidence level of the bootstrap confidence interval or region based on T , or the significance level of the bootstrap test based on T , typically exhibits higher order improvements over the corresponding procedure based on the delta method; e.g., see Beran (1988) and Hall (1992). When T is not asymptotically pivotal, such improvements are not obtained.

The ideal bootstrap estimate $\hat{\lambda}_{\infty}$ is the $1 - \alpha$ quantile of T^* , denoted $\hat{q}_{1-\alpha, \infty}$. It is defined precisely in Table I. The bootstrap estimate $\hat{\lambda}_B$ in this case is the $1 - \alpha$ sample quantile of $\{T_b^* : b = 1, \dots, B\}$, denoted $\hat{q}_{1-\alpha, B}$. Following Hall (1992, p. 307), for the applications considered in this subsection, we choose B not to be just any positive integer, but one that satisfies $\nu/(B + 1) = 1 - \alpha$ for some positive integer ν . This has advantages in terms of the unconditional coverage probability of the resultant confidence interval or region or the unconditional significance level of the resultant test; see Hall (1992, p. 307).

Note that B can be chosen as such only if α is rational. We assume therefore that

$$(2.1) \quad \alpha = \alpha_1/\alpha_2$$

for some positive integers α_1 and α_2 (with no common integer divisors). Then, $B = \alpha_2 h - 1$ and $\nu = (\alpha_2 - \alpha_1)/h$ for some positive integer h . For example, if $\alpha = .05$, then $\alpha_1 = 1$, $\alpha_2 = 20$, $B = 20h - 1$, and $\nu = 19h$ for some integer $h > 0$. In this case, $B = 19, 39, 59$, etc.

For B as defined above $\hat{q}_{1-\alpha, B}$ is the ν th order statistic of $\{T_b^*: b = 1, \dots, B\}$. That is, $\hat{q}_{1-\alpha, B} = T_{B, \nu}^*$.

Table II provides a detailed specification of five applications in which $(\lambda, \hat{\lambda}_\infty, \hat{\lambda}_B) = (q_{1-\alpha}^0, \hat{q}_{1-\alpha, \infty}, \hat{q}_{1-\alpha, B})$ for different choices of T . The five applications are: (i) symmetric two-sided percentile t confidence intervals of level $100(1 - \alpha)\%$, (ii) one-sided percentile t confidence intervals of level $100(1 - \alpha)\%$, (iii) equal-tailed two-sided percentile t confidence intervals of level $100(1 - 2\alpha)\%$, (iv) confidence regions of level $100(1 - \alpha)\%$, and (v) tests for a given significance level α . For each of these applications, Table II specifies the test statistic T , the bootstrap test statistic T^* , the ‘‘theoretical’’ statistical procedure, the ideal bootstrap statistical procedure, and the bootstrap statistical procedure

TABLE II
CONFIDENCE INTERVALS, CONFIDENCE REGIONS, AND TESTS

Application	Test Statistic (T); and Bootstrap Test Statistic (T_b^*)	Theoretical Procedure; Ideal Bootstrap Procedure; and Procedure Based on B Bootstrap Repetitions
1. Symmetric Two-sided Percentile t 100(1 - α)% Confidence Intervals	$ n^\kappa(\hat{\theta} - \theta_0) /\hat{\sigma}$ $ n^\kappa(\hat{\theta}_b^* - \hat{\theta}) /\hat{\sigma}^*$	$J_{SY} = [\hat{\theta} - n^{-\kappa}\hat{\sigma}q_{1-\alpha}^0, \hat{\theta} + n^{-\kappa}\hat{\sigma}q_{1-\alpha}^0]$ $\hat{J}_{SY, \infty} = [\hat{\theta} - n^{-\kappa}\hat{\sigma}\hat{q}_{1-\alpha, \infty},$ $\hat{\theta} + n^{-\kappa}\hat{\sigma}\hat{q}_{1-\alpha, \infty}]$ $\hat{J}_{SY, B} = [\hat{\theta} - n^{-\kappa}\hat{\sigma}\hat{q}_{1-\alpha, B},$ $\hat{\theta} + n^{-\kappa}\hat{\sigma}\hat{q}_{1-\alpha, B}]$
2. One-sided Percentile t 100(1 - α)% Confidence Intervals	$n^\kappa(\hat{\theta} - \theta_0)/\hat{\sigma}$ $n^\kappa(\hat{\theta}_b^* - \hat{\theta})/\hat{\sigma}^*$	$J_1 = [\hat{\theta} - n^{-\kappa}\hat{\sigma}q_{1-\alpha}^0, \infty)$ $\hat{J}_{1, \infty} = [\hat{\theta} - n^{-\kappa}\hat{\sigma}\hat{q}_{1-\alpha, \infty}, \infty)$ $\hat{J}_{1, B} = [\hat{\theta} - n^{-\kappa}\hat{\sigma}\hat{q}_{1-\alpha, B}, \infty)$
3. Equal-tailed Percentile t 100(1 - 2 α)% Confidence Intervals	$n^\kappa(\hat{\theta} - \theta_0)/\hat{\sigma}$ $n^\kappa(\hat{\theta}_b^* - \hat{\theta})/\hat{\sigma}^*$	$J_2 = [\hat{\theta} - n^{-\kappa}\hat{\sigma}q_{1-\alpha}^0, \hat{\theta} - n^{-\kappa}\hat{\sigma}q_\alpha^0]$ $\hat{J}_{2, \infty} = [\hat{\theta} - n^{-\kappa}\hat{\sigma}\hat{q}_{1-\alpha, \infty}, \hat{\theta} - n^{-\kappa}\hat{\sigma}\hat{q}_{\alpha, \infty}]$ $\hat{J}_{2, B} = [\hat{\theta} - n^{-\kappa}\hat{\sigma}\hat{q}_{1-\alpha, B}, \hat{\theta} - n^{-\kappa}\hat{\sigma}\hat{q}_{\alpha, B}]$
4. 100(1 - α)% Confidence Regions	$T(\theta_0, \mathbf{X})$ $T(\hat{\theta}, \mathbf{X}^*)$	$CR = \{\theta_0 \in \Theta : T(\theta_0, \mathbf{X}) \leq q_{1-\alpha}^0\}$ $\widehat{CR}_\infty = \{\theta_0 \in \Theta : T(\theta_0, \mathbf{X}) \leq \hat{q}_{1-\alpha, \infty}\}$ $\widehat{CR}_B = \{\theta_0 \in \Theta : T(\theta_0, \mathbf{X}) \leq \hat{q}_{1-\alpha, B}\}$
5. Tests for a Given Significance Level α	$T(\theta_0, \mathbf{X})$ $T(\hat{\theta}, \mathbf{X}^*)$	Reject H_0 if $T > q_{1-\alpha}^0$ Reject H_0 if $T > \hat{q}_{1-\alpha, \infty}$ Reject H_0 if $T > \hat{q}_{1-\alpha, B}$

Notes: All quantities in the table are defined in Section 2.

based on B repetitions. Note that for notational convenience we define the equal-tailed two-sided percentile t confidence intervals to have confidence level $100(1 - 2\alpha)\%$, not $100(1 - \alpha)\%$.

The three confidence interval applications are based on an estimator $\hat{\theta}$ of a scalar parameter θ_0 . We assume that the normalized estimator $n^\kappa(\hat{\theta} - \theta_0)$ has an asymptotic distribution as $n \rightarrow \infty$. In the leading case, the asymptotic distribution is normal. In many cases of interest, $\kappa = 1/2$. We allow for $\kappa \neq 1/2$, however, to cover nonparametric estimators, such as nonparametric estimators of a density or regression function at a point. Let $\hat{\sigma} = \hat{\sigma}(\mathbf{X})$ denote a consistent estimator of the asymptotic standard error of $n^\kappa(\hat{\theta} - \theta_0)$. Let $\hat{\sigma}_b^* = \hat{\sigma}(\mathbf{X}_b^*)$ for $b = 1, \dots, B$.

For symmetric two-sided percentile t confidence intervals, the asymptotic distribution of T in the leading case is that of the absolute value of a standard normal random variable. That is, $G(x) = (2\Phi(x) - 1)1(x > 0)$, where $\Phi(x)$ is the standard normal distribution function, and $q_{1-\alpha} = z_{1-\alpha/2}$, where z_α denotes the α quantile of a standard normal distribution.

For one-sided percentile t and equal-tailed two-sided percentile t confidence intervals, the asymptotic distribution of T in the leading case is a standard normal distribution. That is, $G(x) = \Phi(x)$ and $q_{1-\alpha} = z_{1-\alpha}$.

For the equal-tailed percentile t confidence interval, we are actually interested in two population quantities λ . The first is $q_{1-\alpha}^0$. It determines the lower endpoint of the confidence interval. The second is q_α^0 . It determines the upper endpoint. In the latter case, $(\hat{\lambda}_\alpha, \hat{\lambda}_B) = (\hat{q}_{\alpha, \infty}, \hat{q}_{\alpha, B})$, where $\hat{q}_{\alpha, \infty}$ is the α quantile of $T^* = n^\kappa(\hat{\theta}^* - \hat{\theta})/\hat{\sigma}^*$ and $\hat{q}_{\alpha, B}$ is the α sample quantile of $\{T_b^* = n^\kappa(\hat{\theta}_b^* - \hat{\theta})/\hat{\sigma}_b^* : b = 1, \dots, B\}$. Given the choice of B such that $\nu/(B+1) = 1 - \alpha$ for some positive integer ν , $\hat{q}_{\alpha, B} = T_{B, \eta}^*$, where $\eta = B + 1 - \nu$. That is, $\hat{q}_{\alpha, B}$ is the η th order statistic of $\{T_b^* : b = 1, \dots, B\}$. If $B = \alpha_2 h - 1$, $\nu = (\alpha_2 - \alpha_1)h$, and $\alpha = \alpha_1/\alpha_2$ for some positive integer h , then $\eta = \alpha_1 h$. For example, if $\alpha = .05$, then $B = 20h - 1$, $\nu = 19h$, and $\eta = h$ for some integer $h > 0$.

The confidence region application is for a parameter vector θ_0 . We consider confidence regions that are defined to be the set of parameter vectors θ_0 such that a test of $H_0: \theta = \theta_0$ based on a test statistic $T = T(\theta_0, \mathbf{X})$ fails to reject the null hypothesis. For this application, the test statistic T could be a Wald statistic, a likelihood ratio statistic, a Lagrange multiplier statistic, etc. The asymptotic distribution of T is G . In the leading case, G is a chi-squared distribution with d degrees of freedom.

For the test application, the test statistic T could be a t statistic, the absolute value of a t statistic, an overidentifying restrictions test statistic, a nonasymptotically pivotal statistic, such as Andrews' (1997) conditional Kolmogorov test statistic for testing the specification of a parametric model, or any of the test statistics listed in the previous paragraph. In the leading cases, the asymptotic distribution G of T is a normal distribution, the distribution of the absolute value of a normal random variable, or a chi-squared distribution with d degrees of freedom.

We note that it is crucial for testing applications that the distribution of the bootstrap statistic T^* mimics the *null* distribution of T whether or not the null is actually true. Otherwise, the bootstrap test will have poor power properties; see Hall and Wilson (1991), Hall and Horowitz (1996), and Li and Maddala (1996). More specifically, the bootstrap sample \mathbf{X}^* and $\tilde{\theta}$ should be defined such that the asymptotic distribution of T^* conditional on the data is the asymptotic *null* distribution G with probability one (with respect to the randomness in the data).

2.4. *p-values*

Here we consider a testing problem in which one wants to report a p -value. In this case, the quantities $(\lambda, \hat{\lambda}_\infty, \hat{\lambda}_B)$ of interest are the exact p -value, p ; the ideal bootstrap p -value, \hat{p}_∞ ; and the bootstrap p -value based on B repetitions, \hat{p}_B . Table I provides the definitions of \hat{p}_∞ and \hat{p}_B .

We view the reporting of a p -value to be an efficient method of communicating the result of hypothesis tests for all significance levels $\alpha \in (0, 1)$. The use of a bootstrap p -value exploits the higher-order improvements of the bootstrap when T is asymptotically pivotal. This holds because the p -value can be used to construct tests with given significance levels of interest and these tests possess the higher order accuracy of bootstrap tests.

2.5. *Bias Correction*

In this application, the objective is to bias-correct an estimator $\hat{\theta}$ of a scalar parameter θ_0 . The quantity λ of interest is the exact bias-corrected estimate, denoted $\hat{\theta}_{bc}^0$. The quantities $\hat{\lambda}_\infty$ and $\hat{\lambda}_B$ are the ideal bootstrap bias-corrected estimator, denoted $\hat{\theta}_{bc,\infty}$, and the bootstrap bias-corrected estimator based on B repetitions, denoted $\hat{\theta}_{bc,B}$, respectively. The latter are defined in Table I.

We assume in this application that $0 < \widehat{se}_\infty = E^*(\hat{\theta}^* - E^*\hat{\theta}^*)^2 < \infty$.

3. A FORMULA FOR THE ACCURACY OF $\hat{\lambda}_B$ AS AN ESTIMATE OF $\hat{\lambda}_\infty$

In this section, we give a simple formula that provides a probabilistic statement of how close $\hat{\lambda}_B$ is to $\hat{\lambda}_\infty$ as a function of the number of bootstrap repetitions B . We are interested in this, because we want B to be sufficiently large that $\hat{\lambda}_B$ is close to $\hat{\lambda}_\infty$. Otherwise, two researchers using the same data and the same statistical method could reach different conclusions due only to the use of different simulation draws.

We measure the closeness of $\hat{\lambda}_B$ and $\hat{\lambda}_\infty$ by the percentage deviation of $\hat{\lambda}_B$ from $\hat{\lambda}_\infty$:

$$(3.1) \quad 100 \frac{|\hat{\lambda}_B - \hat{\lambda}_\infty|}{\hat{\lambda}_\infty}.$$

Let $1 - \tau$ denote a probability close to one, such as .95. Let pdb be a bound on the percentage deviation of $\hat{\lambda}_B$ from $\hat{\lambda}_\infty$. We want to determine $pdb = pdb(B, \tau)$ such that

$$(3.2) \quad P^* \left(100 \frac{|\hat{\lambda}_B - \hat{\lambda}_\infty|}{\hat{\lambda}_\infty} \leq pdb \right) = 1 - \tau.$$

Alternatively, for given B and pdb , we want to determine $\tau = \tau(B, pdb)$ such that (3.2) holds. The function $B = B(pdb, \tau)$ is considered in Section 4 below.

The approximate formulae we give for $pdb = pdb(B, \tau)$ and $\tau = \tau(B, pdb)$ are based on the following asymptotic result:

$$(3.3) \quad B^{1/2} (\hat{\lambda}_B - \hat{\lambda}_\infty) / \hat{\lambda}_\infty \rightarrow_d N(0, \omega),$$

where ω is defined in Table III for each of the applications of Section 2. (Note that ω_1 , also specified in Table III, is defined in Section 4 below.) This result holds as $B \rightarrow \infty$ for fixed n in the applications in which $\hat{\lambda}_B$ is a smooth function of a sample average, viz., the standard error, p -value, and bias correction applications, and as $B \rightarrow \infty$ and $n \rightarrow \infty$ in the applications in which $\hat{\lambda}_B$ is a sample quantile, viz., the confidence interval, confidence region, and hypothesis test applications. The proof of (3.3) and a discussion of the treatment of n as fixed or as diverging to infinity is given in the Appendix.

Let $\hat{\omega}_B$ denote a consistent estimator of ω based on the bootstrap samples $\{\mathbf{X}_b^*: b = 1, \dots, B\}$. Table IV specifies $\hat{\omega}_B$ for each of the applications of Section 2.

Note that $1/\hat{g}_B$ in the definition of $\hat{\omega}_B$ in Table IV for the confidence interval, confidence region, and tests applications is Siddiqui's (1960) estimator (analyzed by Bloch and Gastwirth (1968) and Hall and Sheather (1988)) of the reciprocal of the density of T^* with a plug-in estimator of the bandwidth parameter, viz., \hat{m}_B , calculated by Hall and Sheather (1988). (See Section 7.2.3 of the Appendix for an explanation of why this estimator is suitable even if T^* is discrete.) To reduce the noise of the plug-in estimator, we take advantage of the fact that we know the asymptotic value of the density and use it to generate our estimators of the unknown coefficients in the plug-in formula.

In Table IV, the estimator $\hat{\gamma}_{2B}$ is used in the definition of $\hat{\omega}_B$ for the standard error application to estimate the coefficient of excess kurtosis γ_2 of $\hat{\theta}^*$. Simulations show that $\hat{\gamma}_{2B}$ is downward biased, especially when B is small. In this case, it is preferable to use the bootstrap biased-corrected estimator $\hat{\gamma}_{2BR}$ defined in the Appendix. The latter is easy to compute no matter how difficult it is to compute $\hat{\theta}$.

For confidence intervals, confidence regions, and tests for a given significance level, the formula for $\hat{\omega}_B$ depends on c_α , which depends on the asymptotic distribution G of the statistic T . Table IV provides a general formula for these applications, as well as specific formulae that are obtained in the leading cases for G .

TABLE III
VALUES OF ω AND ω_1 FOR EACH APPLICATION

Application	ω	ω_1	Quantities Used in the Definition of ω
1. Standard Errors	$\frac{2 + \gamma_2}{4}$	$\frac{1}{2}$	$\gamma_2 = (E^*(\hat{\theta}^* - E^*\hat{\theta}^*)^4 / \widehat{se}_\omega^4) - 3$
2. General Formulae for Confidence Intervals, Confidence Regions, and Tests for a Given Significance Level	$\frac{\alpha(1 - \alpha)}{q_{1-\alpha}^2 g^2(q_{1-\alpha})}$	$\frac{\alpha(1 - \alpha)}{q_{1-\alpha}^2 g^2(q_{1-\alpha})}$	$g(\cdot)$ is the density of $G(\cdot)$ $q_{1-\alpha}$ is the $1 - \alpha$ quantile of $G(\cdot)$
2(a). Symmetric Two-sided Confidence Intervals and Other Applications with $G(x) = (2\Phi(x) - 1)I(x > 0)$	$\frac{\alpha(1 - \alpha)}{z_{1-\alpha/2}^2 (2\phi(z_{1-\alpha/2}))^2}$	$\frac{\alpha(1 - \alpha)}{z_{1-\alpha/2}^2 (2\phi(z_{1-\alpha/2}))^2}$	$\phi(\cdot)$ is the standard normal density $z_{1-\alpha}$ is the standard normal $1 - \alpha$ quantile
2(b). One-sided and Equal-tailed Two-sided Confidence Intervals and Other Applications with $G(x) = \Phi(x)$	$\frac{\alpha(1 - \alpha)}{z_{1-\alpha}^2 \phi^2(z_{1-\alpha})}$	$\frac{\alpha(1 - \alpha)}{z_{1-\alpha}^2 \phi^2(z_{1-\alpha})}$	Same as in 2(a)
2(c). Confidence Regions and Tests for a Given Significance Level with $G(x) = F_d(x)$	$\frac{\alpha(1 - \alpha)}{q_{d,1-\alpha}^2 f_d^2(q_{d,1-\alpha})}$	$\frac{\alpha(1 - \alpha)}{q_{d,1-\alpha}^2 f_d^2(q_{d,1-\alpha})}$	$f_d(\cdot)$ is the χ_d^2 density $q_{d,1-\alpha}$ is the χ_d^2 $1 - \alpha$ quantile
3. p -value	$\frac{1 - \hat{p}_\omega}{\hat{p}_\omega}$	$\frac{G(T)}{1 - G(T)}$	$\hat{p}_\omega = P^*(T^* > T)$
4. Bias Correction	$\frac{\widehat{se}_\omega}{\hat{\theta}_{b,c,\infty}^2}$	$\frac{\hat{\sigma}^2}{\hat{\theta}^2}$	$\widehat{se}_\omega = E^*(\hat{\theta}^* - E^*\hat{\theta}^*)^2$

Notes: ω is introduced in Section 3, ω_1 is introduced in Section 4, γ_2 is the coefficient of excess kurtosis of the distribution of the bootstrap estimator $\hat{\theta}^*$, $G(\cdot)$ is the distribution of the asymptotic distribution of T , $\Phi(\cdot)$ is the standard normal distribution function, $F_d(\cdot)$ is the distribution function of a chi-squared distribution with d degrees of freedom, $\hat{\sigma}^2$ is an estimate of the asymptotic standard error of $n^{1/2}(\hat{\theta} - \theta_0)$.

TABLE IV
VALUES OF $\hat{\omega}_B$ FOR EACH APPLICATION

Application	$\hat{\omega}_B$	Quantities Used in the Definition of $\hat{\omega}_B$
1. Standard Errors	$(2 + \hat{\gamma}_{2B})/4$	$\hat{\gamma}_{2B} = \left(\frac{1}{B-1} \sum_{b=1}^B \left(\hat{\theta}_b^* - \frac{1}{B} \sum_{c=1}^B \hat{\theta}_c^* \right)^2 \right)^{1/2} - 3$
2. General Formulae for Confidence Tests for a Given Level	$\frac{\alpha(1-\alpha)(1/\hat{g}_B)^2}{\hat{q}_{1-\alpha, B}^2}$	$1/\hat{g}_B = \frac{B}{2\hat{m}_B} (T_{B, \nu}^* - \hat{m}_B - T_{B, \nu}^{**})$ $\nu = (B+1)(1-\alpha)$ $\hat{m}_B = \text{int}(c_\alpha B^{2/3}), \hat{q}_{1-\alpha, B} = T_{B, \nu}^*$ $c_\alpha = \left(\frac{1.5z_{1-\alpha/2}^2 g^4(q_{1-\alpha})}{3g'(q_{1-\alpha})^2 - g(q_{1-\alpha})g''(q_{1-\alpha})} \right)^{1/3}$
2(a). Symmetric Two-sided Confidence Intervals and Other Applications with $G(x) = (2\Phi(x) - 1)I(x > 0)$	Same as in general case	Same as in general case with $c_\alpha = \left(\frac{6z_{1-\alpha/2}^2 \phi^2(z_{1-\alpha/2})}{2z_{1-\alpha/2}^2 + 1} \right)^{1/3}$
2(b). One-sided and Equal-tailed Two-sided Confidence Intervals and Other Applications with $G(x) = \Phi(x)$	Same as in general case	Same as in general case with $c_\alpha = \left(\frac{1.5z_{1-\alpha/2}^2 \phi^2(z_{1-\alpha})}{2z_{1-\alpha/2}^2 + 1} \right)^{1/3}$

2(c). Confidence Regions and Tests for a Given Significance Level with $G(x) = F_\alpha(x)$	Same as in general case	Same as in general case with
		$c_\alpha = \left(\frac{1.5z_{1-\alpha/2}^2 J_d^2(z_{d,1-\alpha})}{(d-1-2q_{d,1-\alpha})(d-2)/(2q_{d,1-\alpha}) + 1/2} \right)^{1/3}$
3. p -value	$(1 - \hat{p}_B) / \hat{p}_B$	$\hat{p}_B = \frac{1}{B} \sum_{b=1}^B 1(T_b^* > T)$
4. Bias Correction	$\widehat{se}_B / \hat{\theta}_{b_c, B}^2$	$\widehat{se}_B = \left(\frac{1}{B-1} \sum_{b=1}^B \left(\hat{\theta}_b^* - \frac{1}{B} \sum_{c=1}^B \hat{\theta}_c^* \right)^2 \right)^{1/2}$

Notes: $G(\cdot)$ is the distribution function of the asymptotic distribution of T , $\phi(\cdot)$, $\phi(\cdot)$, and z_α are the standard normal distribution function, density function, and α quantile, respectively, $F_d(\cdot)$, $f_d(\cdot)$, and $q_{d,\alpha}$ are the distribution function, density function, and α quantile, respectively, of a chi-squared distribution with d degrees of freedom. $(T_{b,\beta}^*)$: $b = 1, \dots, B$ is the ordered sample of (T_b^*) : $b = 1, \dots, B$, $\hat{\theta}_{b_c, \beta} = 2\hat{\theta} - \frac{1}{\beta} \sum_{b=1}^B \hat{\theta}_b^*$. The formulae given for one-sided and equal-tailed two-sided confidence intervals are appropriate for a right-hand interval endpoint. For a left-hand endpoint, the formula for c_α remains the same, but the formula for ν becomes $\nu = (B + 1)\alpha$. In consequence, the term $q_{1-\alpha, B}^2$ that appears in the denominator of \widehat{se}_B equals $(T_{\beta, \nu}^*)^2 = (T_{\beta, (B+1)\alpha}^*)^2$.

We note that a much simpler choice of $\hat{\omega}_B$ than the formulae given in Table IV for the confidence interval, confidence region, and test examples is simply $\hat{\omega}_B = \omega$. For large n this should work well, but for small n the formulae given in Table IV are much preferred (based on the simulations we have run). The reason is that ω may differ substantially from the finite sample variance of $B^{1/2}(\hat{\lambda}_B - \hat{\lambda}_\infty)/\hat{\lambda}_\infty$, when n is not large and it is a good estimate of the finite sample variance of $B^{1/2}(\hat{\lambda}_B - \hat{\lambda}_\infty)/\hat{\lambda}_\infty$ that is needed for the formulae of (3.4) below to be accurate.

Using (3.3), the relationship between B , pdb , and τ that is defined by (3.2) satisfies the following approximate formulae:

$$(3.4) \quad \begin{aligned} pdb &\doteq 100z_{1-\tau/2}(\hat{\omega}_B/B)^{1/2} \quad \text{or, equivalently,} \\ \tau &\doteq 2\left(1 - \Phi\left(pdb(B/\hat{\omega}_B)^{1/2}/100\right)\right), \end{aligned}$$

where $z_{1-\tau/2}$ and $\Phi(\cdot)$ are the $1 - \tau/2$ quantile and distribution function, respectively, of the standard normal distribution. The formulae of (3.4) are justified by the following asymptotic result:

$$(3.5) \quad P^* \left(100 \frac{|\hat{\lambda}_B - \hat{\lambda}_\infty|}{\hat{\lambda}_\infty} \leq 100z_{1-\tau/2}(\hat{\omega}_B/B)^{1/2} \right) \rightarrow 1 - \tau.$$

As in (3.3), equation (3.5) holds as $B \rightarrow \infty$ for fixed n in the examples in which $\hat{\lambda}_B$ is a smooth function of a sample average and as $B \rightarrow \infty$ and $n \rightarrow \infty$ in the examples in which $\hat{\lambda}_B$ is a sample quantile. See the Appendix for the proof of (3.5).

We now illustrate how the formulae of (3.4) can be utilized. Consider the case of a standard error, i.e., $\lambda = se$. Suppose B has been specified, perhaps by the author of some research paper of interest. We are interested in whether this choice of B is sufficiently large to yield $\hat{\lambda}_B$ close to $\hat{\lambda}_\infty$. Take $1 - \tau$ close to one, say .95. Then, $z_{1-\tau/2} = 1.96$ and

$$(3.6) \quad pdb \doteq 98((2 + \hat{\gamma}_{2B})/B)^{1/2}.$$

For example, if $B = 200$ and $\hat{\gamma}_{2B} = 0$ (which corresponds to the kurtosis of the normal distribution), then $pdb \doteq 10$. That is, with probability approximately .95, \widehat{se}_B is within $\pm 10\%$ of \widehat{se}_∞ . Or, with probability approximately .95, \widehat{se}_∞ is within $\pm 10\%$ of \widehat{se}_B . (The latter interpretation is valid because (3.5) holds with $\hat{\lambda}_\infty$ in the denominator replaced by $\hat{\lambda}_B$.) Alternatively, if $B = 200$ and $\hat{\gamma}_{2B} = 2$, then $pdb \doteq 14$. Next, suppose $B = 50$. When $\hat{\gamma}_{2B} = 0$, we obtain $pdb \doteq 20$ and when $\hat{\gamma}_{2B} = 2$, we obtain $pdb \doteq 28$.

As a second example, consider a confidence interval. In this case, $\lambda = q_{1-\alpha}^0$. Suppose n is quite large, so we can simply take $\hat{\omega}_B = \omega$. Let $\tau = .05$. Then,

$$(3.7) \quad pdb \doteq 196 \left(\frac{\alpha(1-\alpha)}{g^2(q_{1-\alpha})q_{1-\alpha}^2 B} \right)^{1/2}.$$

When $\alpha = .05$ and $G(x) = (2\Phi(x) - 1)1(x > 0)$, which is the leading case for a symmetric two-sided 95% confidence interval, this gives $pdb \doteq 186/B^{1/2}$. If $B = 339$, then $pdp \doteq 10.1$. When $\alpha = .025$ and $G(x) = \Phi(x)$, which is the leading case for an equal-tailed two-sided 95% confidence interval, this gives $pdb \doteq 267/B^{1/2}$. If $B = 339$, then $pdp \doteq 14.5$.

4. A THREE-STEP METHOD FOR DETERMINING THE NUMBER OF BOOTSTRAP REPETITIONS

We now specify a three-step method for determining B to achieve a desired accuracy of $\hat{\lambda}_B$ for estimating $\hat{\lambda}_\infty$. The desired accuracy is specified by a (pdb, τ) combination, such as $(10, .05)$.

Let ω_1 denote a preliminary estimate of the asymptotic variance ω of $B^{1/2}(\hat{\lambda}_B - \hat{\lambda}_\infty)/\hat{\lambda}_\infty$ in (3.3). Table III specifies ω_1 for each of the applications of Section 2. The formulae for ω_1 for confidence intervals, confidence regions, and tests for a given significance level depends on G , the asymptotic distribution of the test statistic T . Table III provides the general formula for these applications, as well as the formulae for the leading cases for G . If G depends on unknown parameters, the parameters can be replaced by consistent estimates in the formula for ω_1 .

The choice of ω_1 in Table III for standard errors is optimal for the case where $\hat{\theta}^*$ has a normal distribution. This is a suitable initial choice for ω , because $\hat{\theta}^*$ typically is asymptotically normal when $\hat{\theta}$ is asymptotically normal. Nevertheless, the three-step procedure does not rely on this for its asymptotic justification.

The choice of ω_1 in Table III for confidence intervals, confidence regions, and tests for a given significance level is based on the asymptotic distribution of T^* as $n \rightarrow \infty$. The choice of ω_1 for p -values is also based on the asymptotic distribution T^* as $n \rightarrow \infty$. In both cases, the asymptotic distribution does not need to be close to the finite sample distribution for the three-step method to work well. The reason is that the initial value of ω_1 is used only to generate an initial value of B that is used, in turn, to obtain an improved value of ω that reflects the finite sample distribution of $\hat{\theta}^*$ or T^* .

Let $\text{int}(a)$ denote the smallest integer greater than or equal to a .

The three-step method is as follows:

STEP 1: Given ω_1 from Table III, compute

$$(4.1) \quad B_1 = \text{int}(10,000z_{1-\tau/2}^2 \omega_1 / pdb^2)$$

or, if $\hat{\lambda}_B$ is an α or $1 - \alpha$ sample quantile, compute

$$(4.2) \quad B_1 = \alpha_2 h_1 - 1,$$

where $\alpha = \alpha_1/\alpha_2$ and $h_1 = \text{int}(10,000z_{1-\tau/2}^2 \omega_1 / (pdb^2 \alpha_2))$.

STEP 2: Simulate B_1 bootstrap samples $\{\mathbf{X}_b^*: b = 1, \dots, B_1\}$ and compute an improved estimate $\hat{\omega}_{B_1}$ of ω using the formulae of Table IV with B replaced by B_1 .

STEP 3: Compute

$$(4.3) \quad B_2 = \text{int}\left(10,000z_{1-\tau/2}^2 \hat{\omega}_{B_1}/pdb^2\right)$$

or, if $\hat{\lambda}_B$ is an α or $1 - \alpha$ sample quantile, compute

$$(4.4) \quad B_2 = \alpha_2 h_2 - 1, \quad \text{where } h_2 = \text{int}\left(10,000z_{1-\tau/2}^2 \hat{\omega}_{B_1}/(pdb^2\alpha_2)\right).$$

Take the desired number of bootstrap repetitions to be $B^* = \max\{B_2, B_1\}$.

Note that Steps 2 and 3 could be iterated with little additional computational burden by replacing B_1 in Step 2 by B_2 , replacing B_2 in Step 3 by B_3 , and taking $B^* = \max\{B_3, B_2, B_1\}$. In some cases, this may lead to finite sample properties that are closer to the asymptotic properties of the three-step procedure. In the simulation results we have carried out, however, the finite sample properties of the three-step procedure are quite close to its asymptotic properties even without iteration.

Often one is interested in more than one standard error estimate, confidence interval, etc. In such cases, there is more than one quantity λ of interest. Furthermore, in the equal-tailed two-sided confidence interval example, there are always two quantities of interest: $q_{1-\alpha}^0$ and q_α^0 .

Suppose there are $M > 1$ quantities λ of interest, say $\{\lambda_j: j = 1, \dots, M\}$. In such cases, the three step method is carried out as follows. Given $\{\omega_{j1}: j = 1, \dots, M\}$ according to the formulae of Table III, one computes $\{B_{j1}: j = 1, \dots, M\}$ in Step 1. In Step 2, one takes $B_1 = \max\{B_{j1}: j = 1, \dots, M\}$ and computes $\{\hat{\omega}_{jB_1}: j = 1, \dots, M\}$ according to the formulae of Table IV. In Step 3, one computes $\{B_{j2}: j = 1, \dots, M\}$ using $\{\hat{\omega}_{jB_1}: j = 1, \dots, M\}$ and takes $B^* = \max\{B_1, B_{12}, \dots, B_{M2}\}$. Note that different values of (pdb, τ) can be used for different value of j .

For tests with a given significance level, application of the three-step method for given (pdb, τ) delivers a choice of B that ensures that the bootstrap critical value based on B repetitions is close to the ideal bootstrap critical value. As suggested by Davidson and MacKinnon (1997), in some cases, one may not want a critical value that is accurate to a prespecified level, but rather, one may want to choose B such that the outcome of the testing procedure is the same whether one uses the critical value based on B repetitions or the ideal bootstrap critical value. Of course, no finite choice of B can guarantee this. Instead, suppose we want to choose B such that the probability of this occurring, $1 - \tau$, is quite high, such as .99. That is, we want

$$(4.5) \quad P^*(1(T > \hat{q}_{1-\alpha, B}) = 1(T > \hat{q}_{1-\alpha, \infty})) \doteq 1 - \tau.$$

The left-hand side equals $P^*(100|\hat{q}_{1-\alpha, B} - \hat{q}_{1-\alpha, \infty}|/\hat{q}_{1-\alpha, \infty} \leq 100|T - \hat{q}_{1-\alpha, \infty}|/\hat{q}_{1-\alpha, \infty})$. In consequence, if $pdb = 100|T - \hat{q}_{1-\alpha, \infty}|/\hat{q}_{1-\alpha, \infty}$, then the three-step method delivers the desired choice of B . This choice of pdb is not feasible because $\hat{q}_{1-\alpha, \infty}$ is unknown.

Instead, in the first and third steps of the three-step method, we suggest taking

$$(4.6) \quad \begin{aligned} pdb &= \min\{\max\{100|T - q_{1-\alpha}|/q_{1-\alpha}, \delta\}, \xi\} \quad \text{and} \\ pdb_1 &= \max\{100|T - \hat{q}_{1-\alpha, B_1}|/\hat{q}_{1-\alpha, B_1}, \delta\}, \end{aligned}$$

respectively, in place of pdb , where $\delta > 0$ and $\xi > \delta$ are truncation bounds, such as 1/2 and 25, that ensure that pdb lies in a reasonable interval, $q_{1-\alpha}$ is the $1 - \alpha$ quantile of G , the asymptotic distribution of T , and $\hat{q}_{1-\alpha, B_1} = T_{B_1, \nu_1}^*$ is the ν_1 th sample quantile of $\{T_b^*: b = 1, \dots, B_1\}$.

5. ASYMPTOTIC JUSTIFICATION OF THE THREE-STEP METHOD

The justification of the three-step method is that as $pdb \rightarrow 0$ (and $n \rightarrow \infty$ when $\hat{\lambda}_B$ is a sample quantile), we have

$$(5.1) \quad P^* \left(100 \frac{|\hat{\lambda}_{B_2} - \hat{\lambda}_\infty|}{\hat{\lambda}_\infty} \leq pdb \right) \rightarrow 1 - \tau.$$

Note that B_2 depends on pdb in (5.1) via (4.3) or (4.4). The proof of (5.1) and some innocuous additional assumptions under which it holds are given in the Appendix.

Equation (5.1) implies that the three-step method attains precisely the specified accuracy asymptotically using “small pdb ” asymptotics when $\omega \geq \omega_1$. If $\omega < \omega_1$, then $B^* = B_1 > B_2$ with probability that goes to one as $pdb \rightarrow 0$ (and $n \rightarrow \infty$ when $\hat{\lambda}_B$ is a sample quantile) and the accuracy of $\hat{\lambda}_{B^*}$ for approximating $\hat{\lambda}_\infty$ exceeds that of (pdb, τ) . This is a consequence of the fact that it would be silly to throw away the extra $B_1 - B_2$ bootstrap estimates that have already been calculated in Step 2.

When $\hat{\lambda}_B$ is a sample quantile, (5.1) holds as is, as well as with B_2 replaced by B^* . The reason is that $\omega = \omega_1$ in this case.

Because one normally specifies a small value of pdb , the asymptotic result (5.1) should be indicative of the relevant nonzero pdb behavior of the three-step method. The simulation results of Section 6 are designed to examine this. We note that the asymptotics used here are completely analogous to large sample size asymptotics with pdb driving B_2 to infinity as $pdb \rightarrow 0$ and B_2 playing the role of the sample size.

When there are $M > 1$ quantities of interest, the three-step method ensures that the percentage deviation of $\hat{\lambda}_{jB}$ from $\hat{\lambda}_{j\infty}$ is less than pdb with probability that is greater than or equal to $1 - \tau$ for all $j = 1, \dots, M$ and equal to $1 - \tau$ for some j , asymptotically.

6. MONTE CARLO SIMULATIONS FOR SYMMETRIC
TWO-SIDED CONFIDENCE INTERVALS

In this section, we evaluate the performance of the three-step method introduced in Section 4 for the case of symmetric two-sided percentile t confidence intervals. More extensive simulation results for standard errors, symmetric two-sided confidence intervals, tests for a given significance level α , and p -values are reported in Andrews and Buchinsky (1999a).

The proposed three-step method is justified by the limit result of (5.1). We wish to see whether this limit result is indicative of finite sample behavior for a range of values of pdb and τ in a standard econometric model. More specifically, given several (pdb, τ) combinations, we want to see how close $P^*(100|\hat{\lambda}_{B_2} - \hat{\lambda}_\infty|/\hat{\lambda}_\infty \leq pdb)$ is to $1 - \tau$. Since the limit result also holds with B_2 replaced by B^* in this application, we also want to see how close $P^*(100|\hat{\lambda}_{B^*} - \hat{\lambda}_\infty|/\hat{\lambda}_\infty \leq pdb)$ is to $1 - \tau$.

6.1. Monte Carlo Experimental Design

The model we consider is the linear regression model

$$(6.1) \quad y_i = x_i' \beta + u_i \quad \text{for } i = 1, \dots, n,$$

where $n = 25$, $X_i = (y_i, x_i)'$ are iid over $i = 1, \dots, n$, $x_i = (1, x_{1i}, \dots, x_{5i})' \in R^6$, (x_{1i}, \dots, x_{5i}) are mutually independent normal random variables, x_i is independent of u_i , and $Eu_i = 0$. The simulation results are invariant with respect to the means and variances of (x_{1i}, \dots, x_{5i}) , the variance of u_i , and the value of the regression parameter β , so we need not be specific as to their values. We consider three error distributions: standard normal (denoted $N(0, 1)$), t with five degrees of freedom (denoted t_5), and chi-squared with five degrees of freedom shifted to have mean zero (denoted χ_5^2). These distributions were chosen in order to assess the effect of heavy tails (t_5 and χ_5^2) and asymmetry (χ_5^2) on the performance of the three-step method.

We estimate β by least squares (LS). We focus attention on the first slope coefficient. Thus, the parameter θ in this case is β_2 (the second element of β). The standard error estimator $\hat{\sigma}$ is defined using the standard formula. That is, $\hat{\sigma}^2$ is the (2, 2) term of the matrix $\hat{\sigma}_u^2 (\sum_{i=1, \dots, 25} x_i x_i' / 25)^{-1}$, where $\hat{\sigma}_u^2 = e'e / (n - 6)$ and e is the vector of the LS residuals.

We simulate 250 different samples from each of the three error distributions. For each of the 250 samples, we compute the LS estimate $\hat{\theta}$ and the standard error estimate $\hat{\sigma}$. Then, we simulate $\hat{q}_{1-\alpha, \infty}$ using 250,000 bootstrap repetitions (each of size 25). We explicitly assume that 250,000 is close enough to infinity to accurately obtain $\hat{q}_{1-\alpha, \infty}$. Given $\hat{\theta}$, $\hat{\sigma}$, and $\hat{q}_{1-\alpha, \infty}$, we calculate the ideal bootstrap symmetric confidence interval $\hat{J}_{SY, \infty}$ defined in Table II for each of the 250 samples and for each error distribution.

Next, we run 2,000 Monte Carlo repetitions for each of the 250 samples for a total of 500,000 repetitions. In each Monte Carlo repetition, we compute \hat{J}_{SY, B_2} , \hat{J}_{SY, B^*} , $\hat{q}_{1-\alpha, B_2}$, and $\hat{q}_{1-\alpha, B^*}$, using the three-step method of Section 4. We make this calculation for several combinations of α (viz., .10 and .05), pdb (viz., 15%, 10%, and 5%), and $1 - \tau$ (viz., .10 and .05). For each repetition and each (α, pdb, τ) combination, we check whether $\hat{q}_{1-\alpha, B^*}$ satisfies

$$(6.2) \quad 100 \frac{|\hat{q}_{1-\alpha, B^*} - \hat{q}_{1-\alpha, \infty}|}{\hat{q}_{1-\alpha, \infty}} \leq pdb,$$

or equivalently, whether $L(\hat{J}_{SY, B^*})$ satisfies

$$(6.3) \quad 100 \frac{|L(\hat{J}_{SY, B^*}) - L(\hat{J}_{SY, \infty})|}{L(\hat{J}_{SY, \infty})} \leq pdb,$$

where $L(\hat{J}_{SY, B^*})$ denotes the length of \hat{J}_{SY, B^*} . We call the fraction of times this condition is satisfied, out of the 2,000 repetitions, the *empirical level* based on B^* . The empirical level based on B_2 bootstrap repetitions is computed analogously. In addition, we compute the fraction of times that θ falls within the constructed confidence interval \hat{J}_{SY, B^*} . We call this fraction the *empirical unconditional coverage probability*. The empirical unconditional coverage probability based on B_2 bootstrap repetitions is defined analogously.

The three-step method of Section 4 is considered to perform well if the empirical levels based on B_2 and B^* bootstrap repetitions are close to $1 - \tau$.

6.2. Monte Carlo Simulation Results

The results from this set of experiments are reported in Table V for the $N(0,1)$ and t_5 error distributions. The numbers reported in this table are averages over the 250 samples. The results for the χ_5^2 error distribution are very similar to those given in Table V(B) for the t_5 error distribution in terms of both the empirical levels obtained and the number of bootstrap repetitions B^* needed. These results show that the high skewness of the χ_5^2 error distribution does not have any effect on the performance of the three-step method. For brevity, we do not report these results.

Table V(A) shows that the empirical levels are somewhat higher than the corresponding $1 - \tau$ values for the experiments with the $N(0,1)$ error distribution. Nevertheless, with low pdb (5), the empirical levels are quite close to their asymptotic counterparts.

Table V(A) indicates that the performance of the three-step method is determined by the number of bootstrap repetitions, B_2 or B^* , employed. The (α, pdb, τ) combinations that yield the best results are those that induce a relatively large number of bootstrap repetitions. Thus, the smaller the bound pdb , the closer are the empirical levels to their asymptotic counterparts, and the

TABLE V
MONTE CARLO SIMULATION RESULTS FOR SYMMETRIC TWO-SIDED CONFIDENCE INTERVALS
A. ERROR DISTRIBUTION $N(0, 1)$

$1 - \alpha$	pd_b	$1 - \tau$	Empirical Level		B_1	B_2			
			B^*	B_2		Mean	Med	Min	Max
.90	15	.90	.946	.943	99	258	216	28	1,837
.90	10	.90	.924	.920	219	394	364	74	1,482
.90	5	.90	.907	.905	849	1,317	1,280	481	3,003
.90	15	.95	.970	.968	139	309	273	47	1,532
.90	10	.95	.960	.957	309	524	493	124	1,652
.90	5	.95	.952	.951	1,209	1,825	1,785	756	3,829
.95	15	.90	.952	.949	119	564	360	31	16,709
.95	10	.90	.947	.946	259	754	654	104	4,346
.95	5	.90	.915	.915	979	1,920	1,843	591	5,104
.95	15	.95	.989	.989	159	1,228	804	69	35,579
.95	10	.95	.969	.968	359	884	801	159	3,809
.95	5	.95	.955	.955	1,399	2,611	2,531	947	6,046

B. ERROR DISTRIBUTION t_5

$1 - \alpha$	pd_b	$1 - \tau$	Empirical Level		B_1	B_2			
			B^*	B_2		Mean	Med	Min	Max
.90	15	.90	.945	.942	99	275	230	29	1,927
.90	10	.90	.924	.920	219	418	385	79	1,560
.90	5	.90	.908	.907	849	1,388	1,348	505	3,196
.90	15	.95	.969	.967	139	329	291	48	1,686
.90	10	.95	.959	.957	309	555	521	130	1,792
.90	5	.95	.953	.952	1,209	1,922	1,878	792	4,048
.95	15	.90	.950	.948	119	587	377	32	18,426
.95	10	.90	.947	.946	259	800	696	107	4,439
.95	5	.90	.917	.916	979	2,055	1,972	635	5,320
.95	15	.95	.989	.989	159	1,274	839	75	39,633
.95	10	.95	.969	.968	359	941	854	163	3,958
.95	5	.95	.957	.956	1,399	2,799	2,714	1,007	6,530

Note: The reported numbers are the averages over the simulations performed for 250 samples, each of which consists of 25 observations. For each sample we carry out 2,000 Monte Carlo repetitions.

more so, the higher the $1 - \tau$ value. For example, for the (.10, 5, .10) combination, the median B^* value is 1,348, while for the combination (.10, 15, .10), it is only 230. As a result, the empirical level for the former case is .907, which is quite close to .900, while for the latter it is .942.

Table V(B) reports the results from the Monte Carlo simulations with the t_5 error distribution. The general picture revealed by Table V(B) is very similar to that of Table V(A). The empirical levels are comparable to those reported in

Table V(A). They are somewhat higher than their asymptotic counterparts. The most pronounced difference between the two sets of experiments is that for all (α, pdb, τ) combinations, the number of bootstrap repetitions B^* is larger for the experiment with the t_5 error distribution, but not by much. This indicates that even with a relatively small sample size (25 observations) the bootstrap distribution of T^* with a fat-tailed t_5 error distribution is not much different than with a $N(0, 1)$ error distribution. Certainly, the bootstrap distribution of T^* based on t_5 errors is far from being a t_5 distribution.

We conclude that the three-step method does pretty well in attaining the desired accuracy of the confidence interval length in relation to its ideal bootstrap counterpart. The three-step method is slightly conservative, because the accuracy obtained is slightly greater than the nominal accuracy.

Lastly, we consider the empirical unconditional coverage probabilities. In all cases, they are the same whether based on B_2 or B^* bootstrap repetitions. In Table V(A), they equal .908 or .909 for all cases where $\alpha = .90$ and .957 for all cases where $\alpha = .95$. In Table V(B), they are in the range .900–.902 for all cases where $\alpha = .90$ and in the range .951–.953 for all cases where $\alpha = .95$. Thus, the empirical unconditional coverage probabilities are extremely close to their asymptotic counterparts. This is consistent with Hall's (1986) result that one need not employ a large number of bootstrap repetitions in order to obtain good unconditional coverage probabilities. Nevertheless, our results show that to construct confidence intervals whose length and conditional coverage probability are close to that of the ideal bootstrap confidence interval, one does need to employ a relatively large number of bootstrap repetitions.

Cowles Foundation for Research in Economics, Yale University, 30 Hillhouse Ave., Box 208281, New Haven, CT 06520-8281, U.S.A.

and

Dept. of Economics, Brown University, Box B, Providence, RI 02912, U.S.A.

Manuscript received November, 1996; final revision received November, 1998.

APPENDIX

A BOOTSTRAP BIAS-CORRECTED ESTIMATOR OF γ_2

Here, we specify a bootstrap bias-corrected estimator of γ_2 , which can be used in the three-step procedure for choosing B for bootstrap standard errors. The iid sample of B bootstrap estimates of θ_0 is $\Theta_B^* = (\hat{\theta}_1^*, \dots, \hat{\theta}_B^*)$. By definition, γ_2 is the coefficient of excess kurtosis of the distribution of $\hat{\theta}_b^*$ for any $b = 1, \dots, B$. For present purposes, we think of $(\hat{\theta}_1^*, \dots, \hat{\theta}_B^*)$ as being the original sample and $\hat{\gamma}_{2B}$ as being an estimator based on this sample that we want to bootstrap bias correct.

Let \hat{G} denote the empirical distribution of $(\hat{\theta}_1^*, \dots, \hat{\theta}_B^*)$. Consider R independent bootstrap samples $\{\Theta_{Br}^{**} : r = 1, \dots, R\}$, where each bootstrap sample $\Theta_{Br}^{**} = (\hat{\theta}_{1r}^{**}, \dots, \hat{\theta}_{Br}^{**})$ is a random sample

of size B drawn from \hat{G} . The bootstrap bias-corrected estimator $\hat{\gamma}_{2BR}$ of γ_2 for R bootstrap repetitions is

$$(7.1) \quad \hat{\gamma}_{2BR} = 2\hat{\gamma}_{2B} - \frac{1}{R} \sum_{r=1}^R \hat{\gamma}_2(\Theta_{Br}^{**}), \quad \text{where}$$

$$\hat{\gamma}_2(\Theta_{Br}^{**}) = \frac{\frac{1}{B-1} \sum_{b=1}^B \left(\hat{\theta}_{br}^{**} - \frac{1}{B} \sum_{c=1}^B \hat{\theta}_{cr}^{**} \right)^4}{\left(\frac{1}{B-1} \sum_{b=1}^B \left(\hat{\theta}_{br}^{**} - \frac{1}{B} \sum_{c=1}^B \hat{\theta}_{cr}^{**} \right)^2 \right)^2} - 3.$$

The computational requirements of $\hat{\gamma}_{2BR}$ are quite modest. The estimator $\hat{\gamma}_{2BR}$ requires that one simulate R bootstrap samples and calculate the simple closed form expressions for $\hat{\gamma}_2(\Theta_{Br}^{**})$ for $r = 1, \dots, R$. For example, when B is 192 (which corresponds to $(pdb, \tau) = (10, .05)$) and $R = 400$, the computational time is only about four seconds using a Sun Sparc-20 computer. Note that the computational requirements of $\hat{\gamma}_{2BR}$ are the same no matter how difficult and time consuming the computation of $\hat{\theta}$ is.

ASSUMPTIONS AND PROOFS

General Proofs

Let “& $n \rightarrow \infty \dots$ ” abbreviate “and $n \rightarrow \infty$ when $\hat{\lambda}_B$ is a sample quantile.” All of the probabilistic statements below refer to the bootstrap simulation randomness conditional on the sample \mathbf{X} , unless stated otherwise.

For each of the three applications, we show below that (3.3) holds and

$$(7.2) \quad \hat{\omega}_B \rightarrow_p \omega \quad \text{as} \quad B \rightarrow \infty \quad (\& n \rightarrow \infty \dots).$$

In consequence, $B^{1/2}(\hat{\lambda}_B - \hat{\lambda}_\infty)/(\hat{\lambda}_\infty \hat{\omega}_B^{1/2}) \rightarrow_d N(0, 1)$ and (3.5) holds.

Next, (7.2) implies that the random number of bootstrap repetitions B_2 satisfies

$$(7.3) \quad B_2/B_\omega \rightarrow_p 1 \quad \text{as} \quad pdb \rightarrow 0 \quad (\& n \rightarrow \infty \dots), \quad \text{where}$$

$$B_\omega = 10,000 z_{1-\tau/2}^2 \omega / pdb^2.$$

Note that B_ω is nonrandom. For each of the applications, we use the proof of (3.3) plus (7.3) to establish that

$$(7.4) \quad B_2^{1/2}(\hat{\lambda}_{B_2} - \hat{\lambda}_\infty)/\hat{\lambda}_\infty \rightarrow_d N(0, \omega) \quad \text{as} \quad pdb \rightarrow 0 \quad (\& n \rightarrow \infty \dots).$$

Equation (7.2), the fact that B_1 is nonrandom, and $B_1 \rightarrow \infty$ as $pdb \rightarrow 0$ ($\& n \rightarrow \infty \dots$) imply that $\hat{\omega}_{B_1} \rightarrow_p \omega$ as $pdb \rightarrow 0$ ($\& n \rightarrow \infty \dots$). This result, (7.4), and the substitution of $pdb = 100 z_{1-\tau/2}(\hat{\omega}_{B_1}/B_2)^{1/2}$ into (5.1) establishes (5.1). The latter expression for pdb follows from the definition of B_2 in (4.3) (ignoring the asymptotically negligible effect of the $\text{int}(\cdot)$ function).

It remains to show that (3.3), (7.2), and (7.4) hold in each of the three applications.

Proofs for the Standard Error Application

Suppose $E^*(\hat{\theta}^*)^4 < \infty$; then $\hat{\gamma}_{2B} \rightarrow_p \gamma_2$ as $B \rightarrow \infty$ by the weak law of large numbers and (7.2) holds. Note that $E^*(\hat{\theta}^*)^4 < \infty$ always holds for the nonparametric bootstrap, because $\hat{\theta}^*$ has a discrete distribution in this case.

Next, we prove (3.3). We rewrite \widehat{se}_B , defined in Table I, as

$$(7.5) \quad \widehat{se}_B = \left(\frac{1}{B} \sum_{b=1}^B (\hat{\theta}_b^* - \mu)^2 - \left(\frac{1}{B} \sum_{b=1}^B \hat{\theta}_b^* - \mu \right)^2 \right)^{1/2} = m(A_B), \quad \text{where}$$

$$A_B = \begin{pmatrix} \frac{1}{B} \sum_{b=1}^B (\hat{\theta}_b^* - \mu)^2 \\ \frac{1}{B} \sum_{b=1}^B \hat{\theta}_b^* - \mu \end{pmatrix} \quad \text{and} \quad m(a) = (a_1 - a_2^2)^{1/2}$$

for $a = (a_1, a_2)'$.

For convenience, we have replaced $B - 1$ in the denominator of \widehat{se}_B by B . By the central limit theorem,

$$(7.6) \quad B^{1/2}(A_B - A) \xrightarrow{d} N(0, \Omega) \quad \text{as} \quad B \rightarrow \infty, \quad \text{where}$$

$$A = \begin{pmatrix} \widehat{se}_\infty^2 \\ 0 \end{pmatrix} \quad \text{and} \quad \Omega = \begin{pmatrix} E^*((\hat{\theta}_b^* - \mu)^2 - \widehat{se}_\infty^2)^2 & E^*(\hat{\theta}_b^* - \mu)^3 \\ E^*(\hat{\theta}_b^* - \mu)^3 & \widehat{se}_\infty^2 \end{pmatrix}.$$

We have $(\partial/\partial a)m(a) = \frac{1}{2}(a_1 - a_2^2)^{-1/2}(1, -2a_2)'$ and $(\partial/\partial a)m(A) = (1/(2\widehat{se}_\infty), 0)'$. The delta method now gives

$$(7.7) \quad B^{1/2}(\widehat{se}_B - \widehat{se}_\infty) = B^{1/2}(m(A_B) - m(A)) \xrightarrow{d} N(0, V), \quad \text{where}$$

$$V = \frac{1}{4\widehat{se}_\infty^2} E^*((\hat{\theta}_b^* - \mu)^2 - \widehat{se}_\infty^2)^2 = \frac{\widehat{se}_\infty^2}{4}(2 + \gamma_2).$$

This establishes (3.3).

Next, we prove (7.4). Equations (7.6) and (7.7) hold with B replaced by B_ω throughout and with the limit as $B \rightarrow \infty$ replaced by the limit as $pdb \rightarrow 0$ (because the latter forces $B_\omega \rightarrow \infty$). Now, by the central limit theorem of Doeblin–Anscombe for a sum of independent random variables with a random number of terms in the sum (e.g., see Chow and Teicher (1978, Thm. 9.4.1, p. 317)), because $B_2/B_\omega \rightarrow_p 1$ as $pdb \rightarrow 0$, the result of (7.6) holds with B replaced by B_2 and with the limit as $B \rightarrow \infty$ replaced by the limit as $pdb \rightarrow 0$. In turn, this implies that (7.7) holds with the same changes, which establishes (7.4).

The Asymptotic Framework and Additional Assumptions for the Confidence Interval, Confidence Region, and Test Applications

We start by discussing the reason for letting $n \rightarrow \infty$ as $B \rightarrow \infty$ (or as $pdb \rightarrow 0$) in the asymptotic justification for these applications. In these applications, $\hat{\lambda}_B = \hat{q}_{1-\alpha, B}$ is a sample quantile based on an iid sample of random variables, each with distribution given by the bootstrap distribution of T^* . If the bootstrap distribution of T^* was absolutely continuous at $\hat{q}_{1-\alpha, \infty}$, then $B^{1/2}(\hat{q}_{1-\alpha, B} - \hat{q}_{1-\alpha, \infty})$ would be asymptotically normally distributed as $B \rightarrow \infty$ for fixed n with asymptotic variance given by $\alpha(1-\alpha)/(f^*(\hat{q}_{q-\alpha, \infty}))^2$, where $f^*(\cdot)$ denotes the density of T^* . But, the bootstrap distribution of T^* is a discrete distribution (at least for the nonparametric bootstrap, which is based on the empirical distribution). In consequence, the asymptotic distribution of $B^{1/2}(\hat{q}_{1-\alpha, B} - \hat{q}_{1-\alpha, \infty})$ as $B \rightarrow \infty$ for fixed n is a *pointmass at zero* for all α values except for those in a set of Lebesgue measure zero. (The latter set is the set of values that the distribution function of T^* takes on at its points of support.)

Although T^* has a discrete distribution in the case of the nonparametric iid bootstrap, its distribution is very nearly continuous even for small values of n . The largest probability π_n of any of its atoms is very small: $\pi_n = n!/n^n \approx (2\pi n)^{1/2}e^{-n}$ provided the original sample \mathbf{X} consists of distinct vectors and distinct bootstrap samples \mathbf{X}^* give rise to distinct values of T^* (as is typically the case); see Hall (1992, Appendix I). This suggests that we should consider asymptotics as $n \rightarrow \infty$, as well as $B \rightarrow \infty$, in order to account for the essentially continuous nature of the distribution of T^* . If we do so, then $B^{1/2}(\hat{q}_{1-\alpha, B} - \hat{q}_{1-\alpha, \infty})$ has a nondegenerate asymptotic distribution with asymptotic variance that depends on the value of a density at a point, just as in the case where the distribution of T^* is continuous. This is what we do. It is in accord with Hall's (1992, p. 285) view that: "for many practical purposes the bootstrap distribution of a statistic may be regarded as continuous."

We now introduce a requisite strengthening of the assumption that $T \rightarrow_d G$ as $n \rightarrow \infty$. We assume: For some $\xi > 0$ and all sequences of constants $\{x_n; n \geq 1\}$ for which $x_n \rightarrow q_{1-\alpha}$, we have

$$(7.8) \quad \begin{aligned} P(T \leq x_n) &= G(x_n) + O(n^{-\xi}) \quad \text{as } n \rightarrow \infty \quad \text{and} \\ P^*(T^* \leq x_n) &= G(x_n) + O(n^{-\xi}) \quad \text{as } n \rightarrow \infty. \end{aligned}$$

(The assumption on T^* is assumed to hold with probability one with respect to the randomness in the data, i.e., with respect to $P(\cdot)$.)

Assumption (7.8) holds whenever the statistic T and the bootstrap statistic T^* have one-term Edgeworth expansions. The latter occurs in any context in which the bootstrap delivers higher order improvements. The literature on the bootstrap is full of results that establish (7.8) for different statistics T and T^* . For example, see Hall (1992, Sec. 3.3 and Ch. 5), Hall and Horowitz (1996), Andrews (1999), and references therein. For example, suppose $T = |n^\kappa(\hat{\theta} - \theta_0)|/\hat{\sigma}$, as is typical for symmetric two-sided confidence intervals. When $\kappa = 1/2$ and $\hat{\sigma}$ is an $n^{1/2}$ -consistent estimator of the asymptotic standard error of $\hat{\theta}$, then (7.8) typically holds with $\xi = 1$. (The $n^{-1/2}$ terms in the Edgeworth expansions of T and T^* typically are even functions of x_n and hence cancel out in the Edgeworth expansions of T and T^* , leaving the order of the first terms of the latter equal to n^{-1} .) One example where (7.8) holds with $\kappa = 1/2$ and $\xi < 1$ is when $\hat{\theta}$ is a sample quantile and $\hat{\sigma}$ is an estimator of its asymptotic standard error (which is not $n^{1/2}$ -consistent because it involves the nonparametric estimation of a density at a point); see Hall and Sheather (1988) and Hall and Martin (1991). When $\kappa < 1/2$, as occurs with nonparametric estimators $\hat{\theta}$, then (7.8) typically holds with $\xi < 1$; see Hall (1992, Ch. 4) and references therein.

As a second example, suppose $T = n^\kappa(\hat{\theta} - \theta_0)/\hat{\sigma}$, as is typical for one-sided and equal-tailed two-sided confidence intervals. When $\kappa = 1/2$ and $\hat{\sigma}$ is an $n^{1/2}$ -consistent estimator of the asymptotic standard error of $\hat{\theta}$, then (7.8) typically holds with $\xi = 1/2$.

To obtain the desired asymptotics in which $\hat{q}_{1-\alpha, B}$ behaves like the sample quantile from a sample of continuous random variables, we cannot allow $B \rightarrow \infty$ or $pdB \rightarrow 0$ too quickly relative to the speed at which $n \rightarrow \infty$. Thus, for (3.3) and (7.2) to hold, we require that

$$(7.9) \quad n^\xi/B^{1/2} \rightarrow \infty \quad \text{as } B \rightarrow \infty \quad \text{and } n \rightarrow \infty,$$

where ξ is as in (7.8). Alternatively, for (7.4) to hold, we require that

$$(7.10) \quad n^\xi \times pdB \rightarrow \infty \quad \text{as } pdB \rightarrow 0 \quad \text{and } n \rightarrow \infty.$$

These are both purely technical assumptions whose justification is that they yield good asymptotic approximations.

Proofs for the Confidence Interval, Confidence Region, and Test Applications

First, we prove (3.3). We use an argument developed for proving the asymptotic distribution of the sample median, e.g., see Lehmann (1983, Thm. 5.3.2, p. 354).

We have: For any $x \in R$,

$$(7.11) \quad P^*(B^{1/2}(\hat{q}_{1-\alpha, B} - \hat{q}_{1-\alpha, \infty}) \leq x) = P^*(T_{B, \nu}^* \leq \hat{q}_{1-\alpha, \infty} + x/B^{1/2}).$$

Let S_B be the number of T_b^* 's for $b = 1, \dots, B$ that exceed $\hat{q}_{1-\alpha, \infty} + x/B^{1/2}$. We have

$$(7.12) \quad T_{B, \nu}^* \leq \hat{q}_{1-\alpha, \infty} + x/B^{1/2} \quad \text{if and only if} \quad S_B \leq B - \nu = B\alpha - (1 - \alpha).$$

The random variable S_B has a binomial distribution with parameters $(B, p_{B, n})$, where

$$(7.13) \quad p_{B, n} = 1 - P^*(T^* \leq \hat{q}_{1-\alpha, \infty} + x/B^{1/2}).$$

The probability in (7.11) equals

$$(7.14) \quad \begin{aligned} &P^*(S_B \leq B\alpha - (1 - \alpha)) \\ &= P^*\left(\frac{S_B - Bp_{B, n}}{(Bp_{B, n}(1 - p_{B, n}))^{1/2}} \leq \frac{B\alpha - (1 - \alpha) - Bp_{B, n}}{Bp_{B, n}(1 - p_{B, n})^{1/2}}\right). \end{aligned}$$

Note that the random variable in the right-hand side probability has mean zero and variance one and satisfies the conditions of the Lindeberg central limit theorem (applied with $B \rightarrow \infty$ and $n \rightarrow \infty$).

Using the assumption of (7.8), we obtain

$$(7.15) \quad \begin{aligned} \hat{q}_{1-\alpha, \infty} &= \inf\{q : P^*(T^* \leq q) \geq 1 - \alpha\} = \inf\{q : G(q) \geq 1 - \alpha\} + o(1) \\ &= q_{1-\alpha} + o(1) \quad \text{as} \quad n \rightarrow \infty \quad \text{and} \\ p_{B, n} &= 1 - P^*(T^* \leq \hat{q}_{1-\alpha, \infty} + x/B^{1/2}) \rightarrow \alpha \quad \text{as} \quad B \rightarrow \infty \quad \text{and} \quad n \rightarrow \infty. \end{aligned}$$

The upper bound in the right-hand side probability of (7.14) can be written as

$$(7.16) \quad \begin{aligned} w_{B, n} &= \frac{B^{1/2}(\alpha - p_{B, n}) - (1 - \alpha)/B^{1/2}}{(p_{B, n}(1 - p_{B, n}))^{1/2}} \\ &= ((\alpha(1 - \alpha))^{1/2} + o(1))B^{1/2}(\alpha - p_{B, n}) + o(1) \end{aligned}$$

as $B \rightarrow \infty$ and $n \rightarrow \infty$. In addition, we have

$$(7.17) \quad \begin{aligned} B^{1/2}(\alpha - p_{B, n}) &= B^{1/2}(P^*(T^* \leq \hat{q}_{1-\alpha, \infty} + x/B^{1/2}) - (1 - \alpha)) \\ &= B^{1/2}(P^*(T^* \leq \hat{q}_{1-\alpha, \infty} + x/B^{1/2}) - P(T \leq \hat{q}_{1-\alpha, \infty})) + o(1) \\ &= B^{1/2}(G(\hat{q}_{1-\alpha, \infty} + x/B^{1/2}) - G(\hat{q}_{1-\alpha, \infty})) + o(1) \\ &= B^{1/2}g(\zeta_{B, n})x/B^{1/2} + o(1) \\ &\rightarrow g(q_{1-\alpha})x \quad \text{as} \quad B \rightarrow \infty \quad \text{and} \quad n \rightarrow \infty. \end{aligned}$$

The first equality of (7.17) holds by the definition of $p_{B, n}$. The second and third equalities hold by (7.8) and (7.9) (using the fact that the latter implies that $B^{1/2} = o(n^\xi)$). The fourth equality holds for some $\zeta_{B, n}$ that lies between $\hat{q}_{1-\alpha, \infty} + x/B^{1/2}$ and $\hat{q}_{1-\alpha, \infty}$ by a mean value expansion, using the assumption that $G(\cdot)$ has a density $g(\cdot)$ in a neighborhood of $q_{1-\alpha}$. The convergence result of (7.17) holds because $\zeta_{B, n} \rightarrow q_{1-\alpha}$ as $B \rightarrow \infty$ and $n \rightarrow \infty$.

Equations (7.16) and (7.17) give

$$(7.18) \quad w_{B, n} \rightarrow g(q_{1-\alpha})x/(\alpha(1 - \alpha))^{1/2} \quad \text{as} \quad B \rightarrow \infty \quad \text{and} \quad n \rightarrow \infty.$$

Equations (7.11), (7.14), and (7.18) plus the Lindeberg central limit theorem applied to (7.14) yield

$$(7.19) \quad \begin{aligned} &P^*(B^{1/2}(\hat{q}_{1-\alpha, B} - \hat{q}_{1-\alpha, \infty}) \leq x) \rightarrow \Phi(xg(q_{1-\alpha})/(\alpha(1 - \alpha))^{1/2}) \quad \text{and} \\ &B^{1/2}(\hat{q}_{1-\alpha, B} - \hat{q}_{1-\alpha, \infty}) \xrightarrow{d} N\left(0, \frac{\alpha(1 - \alpha)}{g^2(q_{1-\alpha})}\right) \end{aligned}$$

as $B \rightarrow \infty$ and $n \rightarrow \infty$, which implies that (3.3) holds, as desired.

Second, we establish (7.2). It suffices to show that

$$(7.20) \quad \hat{q}_{1-\alpha, B} \xrightarrow{P} q_{1-\alpha} \quad \text{and} \quad \left(\frac{B}{2\hat{m}} \right) (T_{B, \nu+\hat{m}}^* - T_{B, \nu-\hat{m}}^*) \xrightarrow{P} \frac{1}{g(q_{1-\alpha})}$$

as $B \rightarrow \infty$ and $n \rightarrow \infty$. The former holds by (7.15) and (7.19). The latter is established as follows.

For a distribution function F , define $F^{-1}(t) = \inf\{x : F(x) \geq t\}$. Let $F_{T^*}(\cdot)$ denote the distribution function of T^* (conditional on the sample \mathbf{X}). Let $\{U_b : b = 1, \dots, B\}$ denote iid uniform $[0, 1]$ random variables. Let $\{U_{B,b} : b = 1, \dots, B\}$ denote the ordered sample of $\{U_b : b = 1, \dots, B\}$. Then, $F_{T^*}^{-1}(U_b)$ has the same distribution as T_b^* and $F_{T^*}^{-1}(U_{B,b})$ has the same distribution as $T_{B,b}^*$. It suffices to show that

$$(7.21) \quad \left(\frac{B}{2\hat{m}} \right) (F_{T^*}^{-1}(U_{B, \nu+\hat{m}}) - F_{T^*}^{-1}(U_{B, \nu-\hat{m}})) \xrightarrow{P} \frac{1}{g(z_{1-\alpha})}$$

as $B \rightarrow \infty$ and $n \rightarrow \infty$.

The left-hand side of (7.21) equals

$$(7.22) \quad \left(\frac{F_{T^*}^{-1}(U_+) - F_{T^*}^{-1}(U_-)}{U_+ - U_-} \right) \left(\frac{B}{2\hat{m}} \right) (U_+ - U_-) \\ = \left(\frac{F_{T^*}^{-1}(U_+) - F_{T^*}^{-1}(U_-)}{U_+ - U_-} \right) (1 + o_p(1))$$

where U_+ and U_- abbreviate $U_{B, \nu+\hat{m}}$ and $U_{B, \nu-\hat{m}}$ respectively. Equation (7.22) holds by the argument of Bloch and Gastwirth (1968, Pf. of Thm. 1) (which relies on the fact that the spacings of the order statistics of uniform random variables have beta distributions). The first term in parentheses on the right-hand side of (7.22) equals

$$(7.23) \quad \frac{G^{-1}(U_+) - G^{-1}(U_-)}{U_+ - U_-} \\ + \frac{B^{1/3}(F_{T^*}^{-1}(U_+) - G^{-1}(U_+))}{B^{1/3}(U_+ - U_-)} - \frac{B^{1/3}(F_{T^*}^{-1}(U_-) - G^{-1}(U_-))}{B^{1/3}(U_+ - U_-)}.$$

The first summand of (7.23) satisfies

$$(7.24) \quad U_+ \xrightarrow{P} 1 - \alpha, \quad U_- \xrightarrow{P} 1 - \alpha, \quad \text{and} \\ \frac{G^{-1}(U_+) - G^{-1}(U_-)}{U_+ - U_-} \xrightarrow{P} \frac{\partial}{\partial x} G^{-1}(1 - \alpha) = \frac{1}{g(G^{-1}(1 - \alpha))} = \frac{1}{g(q_{1-\alpha})}$$

as $B \rightarrow \infty$ and $n \rightarrow \infty$. The first two results of (7.24) hold by standard results for the sample quantiles of iid uniform random variables. The third result follows from the first two results using the definition of differentiability of $G^{-1}(\cdot)$ and an almost sure representation argument.

Next, we show that the second and third summands of (7.23) are $o_p(1)$. By the argument of Bloch and Gastwirth, referred to above, $B^{1/3}(U_+ - U_-) \rightarrow_p 2c_\alpha > 0$. Thus, it suffices to show that

$$(7.25) \quad B^{1/3}(F_{T^*}^{-1}(U_+) - G^{-1}(U_+)) \xrightarrow{P} 0 \quad \text{as} \quad B \rightarrow \infty \quad \text{and} \quad n \rightarrow \infty$$

and likewise with “ U_+ ” replaced by “ U_- ”. The proofs of these two results are the same, so we just prove the former.

It suffices to prove (7.25) with $B^{1/3}$ replaced by $n^{2\xi/3}$ because $B^{1/3} = o(n^{2\xi/3})$ by the assumption of (7.9). For any distribution function F , $x_1 < F^{-1}(t) \leq x_2$ if and only if (iff) $F(x_1) < t \leq F(x_2)$;

see Shorack and Wellner (1986, p. 5). Thus, for any $\varepsilon > 0$,

$$(7.26) \quad \begin{aligned} n^{2\xi/3} |F_{T^*}^{-1}(U_+) - G^{-1}(U_+)| &\leq \varepsilon \quad \text{iff} \\ G^{-1}(U_+) - n^{2\xi/3}\varepsilon &< F_{T^*}^{-1}(U_+) \leq G^{-1}(U_+) + n^{2\xi/3}\varepsilon \quad \text{iff} \\ F_{T^*}(G^{-1}(U_+) - n^{2\xi/3}\varepsilon) &< U_+ \leq F_{T^*}(G^{-1}(U_+) + n^{2\xi/3}\varepsilon). \end{aligned}$$

We have

$$(7.27) \quad \begin{aligned} F_{T^*}(G^{-1}(U_+) - n^{2\xi/3}\varepsilon) &= (F_{T^*}(G^{-1}(U_+) - n^{2\xi/3}\varepsilon) - G(G^{-1}(U_+) - n^{2\xi/3}\varepsilon)) \\ &\quad + G(G^{-1}(U_+) - n^{2\xi/3}\varepsilon) \\ &= O_p(n^{-\xi}) + (U_+ + g(\gamma_{B,n}))(-n^{2\xi/3}\varepsilon) \\ &< U_+ \end{aligned}$$

with probability that goes to one as $B \rightarrow \infty$ and $n \rightarrow \infty$. The second equality of (7.27) holds by (i) the assumption of (7.8), the fact that $G^{-1}(U_+) - n^{2\xi/3}\varepsilon \rightarrow_p q_{1-\alpha}$, and the use of an almost sure representation argument and (ii) a mean value expansion, where $\gamma_{B,n}$ lies between $G^{-1}(U_+) - n^{2\xi/3}\varepsilon$ and $G^{-1}(U_+)$ and, hence, $\gamma_{B,n} \rightarrow_p q_{1-\alpha}$. An analogous result (with the inequality reversed) holds for $F_{T^*}(G^{-1}(U_+) + n^{2\xi/3}\varepsilon)$. Hence, the right-hand side of (7.26) holds with probability that goes to one, which establishes (7.25), and the proof of the second result of (7.20) is complete. Thus, (7.2) holds, as desired.

Third, we use (7.3) and the above proof of (3.3) to establish (7.4). We have: For any $x \in \mathcal{R}$,

$$(7.28) \quad P^*(B_\omega^{1/2}(\hat{q}_{1-\alpha, B_2} - \hat{q}_{1-\alpha, \infty}) \leq x) = P^*(T_{B_2, \nu_2}^* \leq \hat{q}_{1-\alpha, \infty} + x/B_\omega^{1/2}).$$

(Note that we take the normalization factor to be B_ω not B_2 .) Let S_{B_2} be the number of T_b^* 's for $b = 1, \dots, B_2$ that exceed $\hat{q}_{1-\alpha, \infty} + x/B_\omega^{1/2}$. We have

$$(7.29) \quad T_{B_2, \nu_2}^* \leq \hat{q}_{1-\alpha, \infty} + x/B_\omega^{1/2} \quad \text{iff} \quad S_{B_2} \leq B_2 - \nu_2 = B_2\alpha - (1 - \alpha).$$

Thus, the probability in (7.28) equals

$$(7.30) \quad \begin{aligned} P^*(S_{B_2} \leq B_2\alpha - (1 - \alpha)) \\ = P^* \left(\frac{S_{B_2} - B_2 p_{B_\omega, n}}{(B_2 p_{B_\omega, n} (1 - p_{B_\omega, n}))^{1/2}} \leq \frac{B_2\alpha - (1 - \alpha) - B_2 p_{B_\omega, n}}{(B_2 p_{B_\omega, n} (1 - p_{B_\omega, n}))^{1/2}} \right). \end{aligned}$$

The random variable depending on S_{B_2} in the right-hand side probability is a normalized sum of independent random variables with a random number, B_2 , of terms in the sum. By the central limit theorem of Doeblin-Anscombe, it has a standard normal asymptotic distribution as $pdb \rightarrow 0$ and $n \rightarrow \infty$, because (i) it has a standard normal asymptotic distribution when B_2 is replaced by the nonrandom quantity B_ω and (ii) $B_2/B_\omega \rightarrow_p 1$ as $pdb \rightarrow 0$ and $n \rightarrow \infty$ by (7.3).

Now, in the present context, equations (7.15)–(7.18) hold with the following changes: $p_{B,n}$, $B^{1/2}$, $B \rightarrow \infty$, and “ \rightarrow ” are replaced by $p_{B_\omega, n}$, $B_2^{1/2}$, $pdb \rightarrow 0$, and “ \rightarrow_p ,” respectively, and the second and third equalities of (7.17) hold by (7.8) and (7.10) using the fact that the latter and the definition of B_ω imply that $B_\omega^{1/2} = O(1/pdb) = n^\xi O(1/(pdb \times n^\xi)) = o(n^\xi)$ and, hence, $B_2^{1/2} = o_p(n^\xi)$. The revised (7.18) and (7.3) imply that the upper bound in the right-hand side of (7.30) converges in probability to $g(q_{1-\alpha})x/(\alpha(1-\alpha))^{1/2}$ as $pdb \rightarrow 0$ and $n \rightarrow \infty$. This result and the result of the previous paragraph combine to verify (7.19) with $B^{1/2}(\hat{q}_{1-\alpha, B} - \hat{q}_{1-\alpha, \infty})$ replaced by $B_\omega^{1/2}(\hat{q}_{1-\alpha, B_2} - \hat{q}_{q-\alpha, \infty})$. In light of (7.3), this establishes (7.4), as desired.

Lastly, we note that because $\omega = \omega_1$ in this application, we have $B_\omega = B_1$, $B_2/B_1 \rightarrow_p 1$ as $pdb \rightarrow 0$ and $n \rightarrow \infty$ using (7.3), $B^*/B_\omega \rightarrow_p 1$, the proof of (7.4) above goes through with B_2 replaced by B^* throughout and (7.4) and (5.1) hold with B_2 replaced by B^* , as is claimed in Section 4.

Proofs for the p-value and Bias Correction Applications

First, by the central limit theorem for iid random variables,

$$(7.31) \quad \frac{B^{1/2}(\hat{p}_B - \hat{p}_\infty)}{(\hat{p}_\infty(1 - \hat{p}_\infty))^{1/2}} \xrightarrow{d} N(0, 1) \quad \text{and}$$

$$\frac{B^{1/2}(\hat{\theta}_{bc,B} - \hat{\theta}_{bc,\infty})}{\widehat{se}_\infty} \xrightarrow{d} N(0, 1) \quad \text{as } B \rightarrow \infty,$$

because $\hat{p}_\infty = P^*(T^* > T)$ does not equal zero or one by assumption and $\widehat{se}_\infty = E^*(\hat{\theta}^* - E^*\hat{\theta}^*)^2$ is positive and finite by assumption. Hence, (3.3) holds. Second, by the law of large numbers for iid integrable random variables, $\hat{p}_B \rightarrow_p \hat{p}_\infty$ and $\hat{\theta}_{bc,B} \rightarrow_p \hat{\theta}_{bc,\infty}$ as $B \rightarrow \infty$. So, (7.2) holds. Third, (7.31) holds with B replaced by B_ω and with the limit as $B \rightarrow \infty$ replaced by the limit as $pdB \rightarrow 0$ (because the latter forces $B_\omega \rightarrow \infty$). By the central limit theorem of Doebelin-Anscombe, which holds using (7.3), the result of (7.31) holds with B replaced by B_2 and with the limit as $B \rightarrow \infty$ replaced by the limit as $pdB \rightarrow 0$. Hence, (7.4) holds.

REFERENCES

- ANDREWS, D. W. K. (1997): "A Conditional Kolmogorov Test," *Econometrica*, 65, 1097–1128.
- (1999): "Higher-order Improvements of a Computationally Attractive k-step Bootstrap for Extremum Estimators," Cowles Foundation Discussion Paper No. 1230, Yale University; <http://cowles.econ.yale.edu>.
- ANDREWS, D. W. K., AND M. BUCHINSKY (1999a): "Evaluation of a Three-step Method for Choosing the Number of Bootstrap Repetitions," Cowles Foundation Discussion Paper, Yale University.
- (1999b): "On the Number of Bootstrap Repetitions for BC_a Confidence Intervals," Cowles Foundation Discussion Paper, Yale University.
- BERAN, R. (1988): "Pivoting Test Statistics: A Bootstrap View of Asymptotic Refinements," *Journal of the American Statistical Association*, 83, 687–697.
- BLOCH, D. A., AND J. L. GASTWIRTH (1968): "On a Simple Estimate of the Reciprocal of the Density Function," *Annals of Mathematical Statistics*, 39, 1083–1085.
- CARLSTEIN, E. (1986): "The Use of Subseries Values for Estimating the Variance of a General Statistic from a Stationary Sequence," *Annals of Statistics*, 14, 1171–1179.
- CHOW, Y. S., AND H. TEICHER (1978): *Probability Theory: Independence Interchangeability Martingales*. New York: Springer-Verlag.
- DAVIDSON, R., AND J. G. MACKINNON (1997): "Bootstrap Tests: How Many Bootstraps?" unpublished working paper, Department of Economics, Queen's University, Kingston, Ontario.
- EFRON, B. (1987): "Better Bootstrap Confidence Intervals" (with discussion), *Journal of the American Statistical Association*, 82, 171–200.
- EFRON, B., AND R. TIBSHIRANI (1986): "Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy," *Statistical Science*, 1, 54–77.
- (1993): *An Introduction to the Bootstrap*. New York: Chapman and Hall.
- FREEDMAN, D. (1981): "Bootstrapping Regression Models," *Annals of Statistics*, 9, 1218–1228.
- HALL, P. (1986): "On the Number of Bootstrap Simulations Required to Construct a Confidence Interval," *Annals of Statistics*, 14, 1453–1462.
- (1992): *The Bootstrap and Edgeworth Expansion*. New York: Springer-Verlag.
- HALL, P., AND J. L. HOROWITZ (1996): "Bootstrap Critical Values for Tests Based on Generalized-Method-of-Moment Estimators," *Econometrica*, 64, 891–916.
- HALL, P., AND M. A. MARTIN (1991): "On the Error Incurred Using the Bootstrap Variance Estimator when Constructing Confidence Intervals for Quantiles," *Journal of Multivariate Analysis*, 38, 70–81.
- HALL, P., AND S. J. SHEATHER (1988): "On the Distribution of a Studentized Quantile," *Journal of the Royal Statistical Society, Series B*, 50, 381–391.
- HALL, P., AND S. R. WILSON (1991): "Two Guidelines for Bootstrap Hypothesis Testing," *Biometrics*, 47, 757–762.

- KIEFER, J. (1982): "Conditional Inference," in *Encyclopedia of Statistical Sciences*, Vol. 2. New York: J. Wiley and Sons, pp. 103–109.
- KUNSCH, H. R. (1989): "The Jackknife and the Bootstrap for General Stationary Observations," *Annals of Statistics*, 17, 1217–1241.
- LEHMANN, E. L. (1983): *Theory of Point Estimation*. New York: Wiley.
- LI, H., AND G. S. MADDALA (1996): "Bootstrapping Time Series Models," *Econometric Reviews*, 15, 115–158.
- SHORACK, G. R., AND J. A. WELLNER (1986): *Empirical Processes with Applications to Statistics*. New York: Wiley.
- SIDDIQUI, M. M. (1960): "Distribution of Quantiles in Samples from a Bivariate Population," *Journal of Research of the National Bureau of Standards*, B, 64, 145–150.