*International Statistical Review* (2003), 71, 2, 435–459, Printed in The Netherlands


# Bootstrap Methods for Time Series

**Wolfgang Härdle[1], Joel Horowitz[2] and Jens-Peter Kreiss[3]**

*[1]CASE - Center for Applied Statistics and Economics, Humboldt-Universität zu Berlin, Berlin, Germany. [2]Department of Economics, Northwestern University, Evanston, IL, USA. [3]Institute for Mathematical Stochastics, Technical University of Braunschweig, Braunschweig, Germany*

## Summary

**The bootstrap is a method for estimating the distribution of an estimator or test statistic by resampling one's data or a model estimated from the data. The methods that are available for implementing the bootstrap and the accuracy of bootstrap estimates depend on whether the data are an independent random sample or a time series. This paper is concerned with the application of the bootstrap to time-series data when one does not have a finite-dimensional parametric model that reduces the data generation process to independent random sampling. We review the methods that have been proposed for implementing the bootstrap in this situation and discuss the accuracy of these methods relative to that of first-order asymptotic approximations. We argue that methods for implementing the bootstrap with time-series data are not as well understood as methods for data that are independent random samples. Although promising bootstrap methods for time series are available, there is a considerable need for further research in the application of the bootstrap to time series. We describe some of the important unsolved problems.**

*Key words:* Resampling; Block bootstrap; Asymptotic approximation confidence interval.

## 1 Introduction

The bootstrap is a method for estimating the distribution of an estimator or test statistic by resampling one's data or a model estimated from the data. Under conditions that hold in a wide variety of applications, the bootstrap provides approximations to distributions of statistics, coverage probabilities of confidence intervals, and rejection probabilities of tests that are at least as accurate as the approximations of first-order asymptotic distribution theory. Often, the bootstrap provides approximations that are more accurate than those of first-order asymptotic theory. This is important in applied research, because first-order asymptotic approximations are often inaccurate with the sample sizes found in applications.

The methods that are available for implementing the bootstrap and the improvements in accuracy that it achieves relative to first-order asymptotic approximations depend on whether the data are an independently and identically distributed (*iid*) random sample or a time series. If the data are *iid*, the bootstrap can be implemented by sampling the data randomly with replacement or by sampling a parametric model of the distribution of the data. The distribution of a statistic is estimated by its empirical distribution under sampling from the data or parametric model. Beran & Ducharme (1991), Hall (1992), Efron & Tibshirani (1993), and Davison & Hinkley (1997) provide detailed discussions of bootstrap methods and their properties for *iid* data.

The situation is more complicated when the data are a time series because bootstrap sampling must be carried out in a way that suitably captures the dependence structure of the data generation process (DGP). This is not difficult if one has a finite-dimensional parametric model (e.g., a finite-order ARMA model) that reduces the DGP to independent random sampling. In this case and under suitable regularity conditions, the bootstrap has properties that are essentially the same as they are when the data are *iid*. See Andrews (2001) and Bose (1988, 1990).

This paper is concerned with the situation in which the data are a time series but one does not have a finite-dimensional parametric model that reduces the DGP to independent random sampling. We focus on the question of whether the bootstrap provides approximations to distribution functions, coverage probabilities, and rejection probabilities that are more accurate than first-order asymptotic approximations. We argue that bootstrap methods for time-series data are not as well understood as methods for *iid* data. Moreover, the relative accuracy of bootstrap and first-order asymptotic approximations tends to be poorer with time series data than with *iid* samples. With *iid* data the bootstrap is often much more accurate than are first-order approximations. With time-series data, the bootstrap's ability to provide improved accuracy is more problematic. We conclude that although promising bootstrap methods for time series exist, there is a considerable need for further research on the topic. We describe some of the important unsolved problems.

Section 2 describes the estimation and inference problems that will be discussed in the remainder of the paper. Section 2 also provides background information on the performance of the bootstrap when the data are *iid* and on the theory underlying the bootstrap's ability to improve on first-order approximations. Section 3 reviews the block bootstrap, which is the oldest and best known nonparametric method for implementing the bootstrap with time-series data. The block bootstrap imposes relatively few *a priori* restrictions on the DGP, but this flexibility comes at the price of estimation errors that converge to zero relatively slowly. Section 4 discusses methods that make stronger assumptions about the DGP but offer the possibility of faster converging errors. Section 5 presents the results of Monte Carlo experiments that illustrate the finite-sample performance of some of the methods described in Sections 3 and 4 and the empirical relevance of the theory. Section 6 presents our subjective views on the practical usefulness of existing bootstrap methods for time series and on areas where further research would be beneficial.

The regularity conditions required by bootstrap methods for time-series tend to be highly technical, and they vary among investigators and methods. To enable us to concentrate on important ideas rather than technicalities, we do not give detailed regularity conditions in this paper. They are available in the references that are cited. We assume throughout the paper that the DGP is stationary and weakly dependent in a sense that is made precise in Section 2. Bootstrap methods for DGP's that are nonstationary or have stronger forms of dependence, notably long-memory and unit-root processes, are important topics for research but are at a much more preliminary stage of development.

## 2 Problem Definition and Background Information

This section has three parts. Section 2.1 describes the estimation and inference problems that will be treated in the remainder of the paper. Section 2.2 reviews the performance of the bootstrap when the data are *iid*. This provides a benchmark for judging the bootstrap's performance with time-series data. Section 2.3 reviews the theory underlying the bootstrap's ability to improve on first-order asymptotic approximations.

### *2.1 Statement of the Problem*

Let $\{X_i : i = 1, \dots, n\}$ be observations from the sequence $\{X_i : -\infty < i < \infty\}$, where $X_i \in \mathbb{R}^d$ for each integer $i$ and some integer $d$ satisfying $1 \le d < \infty$. We assume that $\{X_i\}$ is a realization of a strictly stationary, discrete-time stochastic process (the DGP) whose mean, $\mu \equiv \mathbf{E}(X_1)$, exists. We also assume that $\{X_i\}$ satisfies a weak dependence condition called *geometric strong mixing* (GSM).

To define GSM, let $\mathcal{F}_j^k$ be the $\sigma$-algebra generated by $\{X_i : j \leq i \leq k\}$, and define

$$\alpha(k) = \sup_j \sup_{A \in \mathcal{F}_{-\infty}^j, B \in \mathcal{F}_{j-k}^{\infty}} |\mathbf{P}(A \cap B) - \mathbf{P}(A)\mathbf{P}(B)|.$$

Then $\{X_i\}$ is GSM if $\alpha(k) \leq C\rho^k$ for some finite $C > 0$ and $\rho < 1$. In other words, $\{X_i\}$ is GSM if the events $A$ and $B$ become "nearly independent" sufficiently rapidly as their separation in time increases. The GSM assumption is stronger than needed for some of the results described later in this paper, but it enables us to avoid technical distinctions among strengths of dependence that are of little consequence in applied research.

Now define $m_n = n^{-1} \sum_{i=1}^n X_i$. Let $\theta : \mathbb{R}^d \to \mathbb{R}$ be a function. In this paper, we are concerned with making inferences about $\theta = \theta(\mu)$ based on the estimator $\theta_n \equiv \theta(m_n)$. As is discussed by Hall & Horowitz (1996) and Andrews (2002a), a wide variety of estimators that are important in applications either are or can be approximated with (higher-order) asymptotically negligible errors by smooth functions of sample moments of suitably defined random variables. Therefore, the focus on estimators of the form $\theta_n$ is not highly restrictive. In particular, generalized-method-of-moments (GMM) estimators can be approximated this way under mild regularity conditions (Hall & Horowitz, 1996; Andrews, 2002a). GMM is a method for estimating a possibly vector-valued parameter $\psi$ that is identified by the moment condition $\mathbf{E}g(X, \psi) = 0$, where $g$ is a function whose dimension equals or exceeds that of $\psi$. The class of GMM estimators includes linear and nonlinear least squares estimators and maximum likelihood estimators, among many others. Hansen (1982) provides details of the GMM method and gives conditions under which GMM estimators are $n^{1/2}$-consistent and asymptotically normal.

Assume now that $\mathbf{E}\theta(m_n)$ and $\sigma_\theta^2 \equiv \mathrm{Var}[\theta(m_n)]$ exist. Define $\sigma_\infty^2$ to be the variance of the asymptotic distribution of $n^{1/2}(\theta_n - \theta)$, and let $s_n^2$ be a weakly consistent estimator of $\sigma_\infty^2$. In the remainder of this paper, we discuss the use of the bootstrap to estimate the following quantities, all of which are important in applied research:

1. The bias and variance of $\theta_n$, that is $\mathbf{E}\theta_n - \theta$ and $\sigma_\theta^2$.
2. The one-sided distribution functions $\mathbf{P}[n^{1/2}(\theta_n - \theta) \leq \tau]$, $\mathbf{P}[n^{1/2}(\theta_n - \theta)/\sigma_\infty \leq \tau]$, and $\mathbf{P}[n^{1/2}(\theta_n - \theta)/s_n \leq \tau]$ for any real $\tau$.
3. The symmetrical distribution functions $\mathbf{P}[n^{1/2}|(\theta_n - \theta)| \leq \tau]$, $\mathbf{P}[n^{1/2}|(\theta_n - \theta)|/\sigma_\infty \leq \tau]$, and $\mathbf{P}[n^{1/2}|(\theta_n - \theta)|/s_n \leq \tau]$.
4. The coverage probabilities of one-sided and symmetrical confidence intervals for $\theta$. These are $-\infty < \theta \leq \theta_n - \hat{z}_{1-\alpha}s_n$, $\theta_n - \hat{z}_\alpha s_n \leq \theta < \infty$, and $\theta_n - \hat{z}_{\alpha/2}s_n \leq \theta \leq \theta_n + \hat{z}_{\alpha/2}s_n$, where $\hat{z}_\alpha$ is a bootstrap estimator of the $1 - \alpha$ quantile of the distribution of $n^{1/2}(\theta_n - \theta)/s_n$.
5. The probabilities that one-sided and symmetrical tests reject the correct null hypothesis $\mathrm{H}_0 : \theta = \theta_0$ (hereinafter rejection probabilities). For one-sided tests, the test statistic is $n^{1/2}(\theta_n - \theta_0)/s_n$ with bootstrap critical values $\hat{z}_\alpha$ and $\hat{z}_{1-\alpha}$ for upper- and lower-tail tests, respectively. For a symmetrical test, the test statistic is $n^{1/2}|(\theta_n - \theta)|/s_n$, and the bootstrap critical value is $\hat{z}_{\alpha/2}$.

The conclusions regarding coverage probabilities of confidence intervals and rejection probabilities of tests are often identical, so we will not always treat both explicitly here.

### 2.2 *Performance of the Bootstrap when the Data Are IID*

The rates of convergence of the errors made by first-order asymptotic approximations and by bootstrap estimators with data that are *iid* provide benchmarks for judging the bootstrap's performance with time series. As will be discussed in Sections 3 and 4, the errors that the bootstrap makes in estimating distribution functions, coverage probabilities, and rejection probabilities converge to zero more slowly when the data are a time series than when they are *iid*. In some cases, the rates of convergence for time series data are close to those for *iid* data, but in others they are only slightly faster than the rates with first-order approximations.

If the data are *iid*, then under regularity conditions that are given by Hall (1992),

1. The bootstrap estimates $\sigma_\theta^2$ consistently and reduces the bias of $\theta_n$ to $O(n^{-2})$. That is $\mathbf{E}(\theta_n - \theta - \hat{B}_n) = O(n^{-2})$, where $\hat{B}_n$ is the bootstrap estimator of $\mathbf{E}(\theta_n - \theta)$. By contrast, $\mathbf{E}(\theta_n - \theta) = O(n^{-1})$. Horowitz (2001) gives an algorithm for computing $\hat{B}_n$.
   If, in addition, $n^{1/2}(\theta_n - \theta)/s_n \to^d N(0, 1)$, then:
2. The errors in the bootstrap estimates of the one-sided distribution functions $\mathbf{P}[n^{1/2}(\theta_n - \theta)/\sigma_\infty \leq \tau]$ and $\mathbf{P}[n^{1/2}(\theta_n - \theta)/s_n \leq \tau]$ are $O_p(n^{-1})$. The errors made by first order asymptotic approximations are $O(n^{-1/2})$. By "error" we mean the difference between a bootstrap estimator and the population probability that it estimates.
3. The errors in the bootstrap estimates of the symmetrical distribution functions $\mathbf{P}[n^{1/2}|(\theta_n - \theta)|/\sigma_\infty \leq \tau]$ and $\mathbf{P}[n^{1/2}|(\theta_n - \theta)|/s_n \leq \tau]$ are $O_p(n^{-3/2})$, whereas the errors made by first-order approximations are $O(n^{-1})$.
4. When the bootstrap is used to obtain the critical value of a one-sided hypothesis test, the resulting difference between the true and nominal rejection probabilities under the null hypothesis (error in the rejection probability or ERP) is $O(n^{-1})$, whereas it is $O(n^{-1/2})$ when the critical value is obtained from first-order approximations. The same result applies to the error in the coverage probability (ECP) of a one-sided confidence interval. In some cases, the bootstrap can reduce the ERP (ECP) of a one-sided test (confidence interval) to $O(n^{-3/2})$ (Hall, 1992, p. 178; Davidson & MacKinnon, 1999).
5. When the bootstrap is used to obtain the critical value of a symmetrical hypothesis test, the resulting ERP is $O(n^{-2})$, whereas it is $O(n^{-1})$ when the critical value is obtained from first-order approximations. The same result applies to the ECP of a symmetrical confidence interval.

### 2.3 *Why the Bootstrap Provides Asymptotic Refinements*

The term *asymptotic refinements* refers to approximations to distribution functions, coverage probabilities and rejection probabilities that are more accurate than those of first-order asymptotic distribution theory. This section outlines the theory underlying the bootstrap's ability to provide asymptotic refinements. To minimize the length of the discussion, we concentrate on the distribution function of the asymptotically $N(0, 1)$ statistic $T_n \equiv n^{1/2}(\theta_n - \theta)/s_n$. Hall (1992) gives regularity conditions for the results of this section when the data are *iid*. References cited in Sections 3–4 give regularity conditions for time series.

Let $\hat{\mathbf{P}}$ denote the probability measure induced by bootstrap sampling, and let $\hat{T}_n$ denote a bootstrap analog of $T_n$. If the data are *iid*, then it suffices to let $\hat{\mathbf{P}}$ be the empirical distribution of the data. Bootstrap samples are drawn by sampling the data $\{X_i : i = 1, \ldots, n\}$ randomly with replacement. If $\{\hat{X}_i : i = 1, \ldots, n\}$ is such a sample, then

$$\hat{T}_n \equiv n^{1/2}(\hat{\theta}_n - \theta_n)/\hat{s}_n,$$

where $\hat{\theta}_n = \theta(\hat{m}_n)$, $\hat{m}_n = n^{-1}\sum_{i=1}^n \hat{X}_i$, and $\hat{s}_n^2$ is obtained by replacing the $\{X_i\}$ with $\{\hat{X}_i\}$ in the formula for $s_n^2$. For example, if $\Sigma_n$ denotes the sample covariance matrix of $X$ and $s_n^2 = \nabla\theta(m_n)'\Sigma_n\nabla\theta(m_n)$, then $\hat{s}_n^2 = \nabla\theta(\hat{m}_n)'\hat{\Sigma}_n\nabla\theta(\hat{m}_n)$, where $\hat{\Sigma}_n$ is the sample covariance matrix of the bootstrap sample. Bootstrap versions of $\hat{T}_n$ for time-series data are presented in Sections 3–4. The discussion in this section does not depend on the details of $\hat{T}_n$.

The arguments showing that the bootstrap yields asymptotic refinements are based on Edgeworth expansions of $\mathbf{P}(T_n \leq z)$ and $\hat{\mathbf{P}}(\hat{T}_n \leq z)$. The arguments show that the bootstrap provides a low-order

Edgeworth expansion of $\mathbf{P}(T_n \leq z)$. Additional notation is needed to describe the expansions. Let $\Phi$ and $\phi$, respectively, denote the standard normal distribution function and density. The $j$'th cumulant of $T_n$ $(j \leq 4)$ has the form $n^{-1/2}\kappa_j + o(n^{-1/2})$ if $j$ is odd and $I(j = 2) + n^{-1}\kappa_j + o(n^{-1})$ if $j$ is even, where $\kappa_j$ is a constant and $I$ is the indicator function (Hall, 1992, p. 46). Define $\kappa = (\kappa_1, \ldots, \kappa_4)'$. Conditional on the data $\{X_i : i = 1, \ldots, n\}$, the $j$'th cumulant of $\hat{T}_n$ almost surely has the form $n^{-1/2}\hat{\kappa}_j + o(n^{-1/2})$ if $j$ is odd and $I(j = 2) + n^{-1}\hat{\kappa}_j + o(n^{-1})$ if $j$ is even. The quantities $\hat{\kappa}_j$ depend on $\{X_i\}$. They are nonstochastic relative to bootstrap sampling but are random variables relative to the stochastic process that generates $\{X_i\}$. Define $\hat{\kappa} = (\hat{\kappa}_1, \ldots, \hat{\kappa}_4)'$.

$\mathbf{P}(T_n \leq z)$ has the Edgeworth expansion

$$\mathbf{P}(T_n \leq z) = \Phi(z) + \sum_{j=1}^{2} n^{-j/2} g_j(z, \kappa)\phi(z) + O(n^{-3/2}) \tag{2.1}$$

uniformly over $z$, where $g_j(z, \kappa)$ is a polynomial function of $z$ for each $\kappa$, a polynomial function of the components of $\kappa$ for each $z$, an even function of $z$ if $j = 1$, and an odd function of $z$ if $j = 2$. Moreover, $\mathbf{P}(|T_n| \leq z)$ has the expansion

$$\mathbf{P}(|T_n| \leq z) = 2\Phi(z) - 1 + 2n^{-1} g_2(z, \kappa)\phi(z) + O(n^{-2}) \tag{2.2}$$

uniformly over $z$. Conditional on the data, the bootstrap probabilities $\hat{\mathbf{P}}(\hat{T}_n \leq z)$ and $\hat{\mathbf{P}}(|\hat{T}_n| \leq z)$ have the expansions

$$\hat{\mathbf{P}}(\hat{T}_n \leq z) = \Phi(z) + \sum_{j=1}^{2} n^{-j/2} g_j(z, \hat{\kappa})\phi(z) + O(n^{-3/2}) \tag{2.3}$$

and

$$\hat{\mathbf{P}}(|\hat{T}_n| \leq z) = 2\Phi(z) - 1 + 2n^{-1} g_2(z, \hat{\kappa})\phi(z) + O(n^{-2}) \tag{2.4}$$

uniformly over $z$ almost surely. Let $\nabla_\kappa g_j$ denote the gradient of $g_j$ with respect to its second argument. Then a Taylor series expansion yields

$$|\hat{\mathbf{P}}(\hat{T}_n \leq z) - \mathbf{P}(T_n \leq z)| = n^{-1/2}\left[\nabla_\kappa g_1(z, \kappa)(\hat{\kappa} - \kappa)\right]\phi(z) + O\left(n^{-1/2}\left\|\hat{\kappa} - \kappa\right\|^2\right) + O(n^{-1}) \tag{2.5}$$

and

$$|\hat{\mathbf{P}}(|\hat{T}_n| \leq z) - \mathbf{P}(|T_n| \leq z)| = 2n^{-1}\left[\nabla_\kappa g_2(z, \kappa)(\hat{\kappa} - \kappa)\right]\phi(z) + O\left(n^{-1}\left\|\hat{\kappa} - \kappa\right\|^2\right) + O(n^{-2}) \tag{2.6}$$

almost surely uniformly over $z$. Thus, the leading terms of the errors made by bootstrap estimators of one-sided and symmetrical distribution functions are $n^{-1/2}\left[\nabla_\kappa g_1(z, \kappa)(\hat{\kappa} - \kappa)\right]\phi(z)$ and $2n^{-1}\left[\nabla_\kappa g_2(z, \kappa)(\hat{\kappa} - \kappa)\right]\phi(z)$, respectively. If the data are *iid*, then $\left\|\hat{\kappa} - \kappa\right\| = O_p(n^{-1/2})$, so the errors of the bootstrap estimators are $O_p(n^{-1})$ and $O_p(n^{-3/2})$ for one-sided and symmetrical distribution functions, respectively. The root-mean-square estimation errors (RMSE's) also converge at these rates. In contrast, the asymptotic normal approximation keeps only the leading terms on the right-hand sides of (2.1) and (2.2). Therefore, its errors are $O(n^{-1/2})$ and $O(n^{-1})$ for one-sided and symmetrical distributions, respectively. Thus, the bootstrap's estimation errors converge to zero more rapidly than do those of the asymptotic normal approximation.

Now consider the ERP of a symmetrical hypothesis test. Let $z_{\alpha/2}$ and $\hat{z}_{\alpha/2}$, respectively, denote the $1 - \alpha/2$ quantiles of the distributions of $T_n$ and $\hat{T}_n$. Then $\mathbf{P}(|T_n| \leq z_{\alpha/2}) = \hat{\mathbf{P}}(|\hat{T}_n| \leq \hat{z}_{\alpha/2}) = 1 - \alpha$.

The bootstrap-based symmetrical test at the nominal $\alpha$ level accepts $H_0$ if $|T_n| \leq \hat{z}_{\alpha/2}$. Thus, the ERP of the test is $\mathbf{P}(|T_n| \leq \hat{z}_{\alpha/2}) - (1 - \alpha)$. If the data are *iid*, then by carrying out an Edgeworth expansion of $\mathbf{P}(|T_n| \leq \hat{z}_{\alpha/2})$, it can be shown that

$$\mathbf{P}(|T_n| \leq \hat{z}_{\alpha/2}) = 1 - \alpha + O(n^{-2}). \tag{2.7}$$

See Hall (1992, pp. 108–114). Thus, the ERP of a symmetrical test (and the ECP of a symmetrical confidence interval) based on the bootstrap critical value $\hat{z}_{\alpha/2}$ is $O(n^{-2})$ when the data are *iid*. As will be discussed in Sections 3 and 4, this rate of the ERP is not available with current bootstrap methods for time series. Rather, the ERP for time-series data is $O(n^{-1-a})$ for some $a$ satisfying $0 < a < 1/2$.

## 3 The Block Bootstrap

The block bootstrap is the best-known method for implementing the bootstrap with time-series data. It consists of dividing the data into blocks of observations and sampling the blocks randomly with replacement. The blocks may be non-overlapping (Hall, 1985; Carlstein, 1986) or overlapping (Hall, 1985; Künsch, 1989; Politis & Romano, 1993). To describe blocking methods more precisely, let the data consist of observations $\{X_i : i = 1, \ldots, n\}$. In general, the parameter of interest may depend on current and lagged values of $X$ up to order $q \geq 0$. An example is $\theta = \mathbf{E}[G(X_1, \ldots, X_{1+q})]$, where $G$ is a known function. We assume that $q < \infty$. Politis & Romano (1992) discuss the case of $q = \infty$ (e.g, spectral density estimation). See, also, Bühlmann & Künsch (1995). For $q < \infty$, define $Y_i = \{X_i, \ldots, X_{i-q}\}$. With non-overlapping blocks of length $\ell$, block 1 is observations $\{Y_j : j = 1, \ldots, \ell\}$, block 2 is observations $\{Y_{\ell+j} : j = 1, \ldots, \ell\}$, and so forth. With overlapping blocks of length $\ell$, block 1 is observations $\{Y_j : j = 1, \ldots, \ell\}$, block 2 is observations $\{Y_{j+1} : j = 1, \ldots, \ell\}$, and so forth. The bootstrap sample is obtained by sampling blocks randomly with replacement and laying them end-to-end in the order sampled. The procedure of sampling blocks of $Y_i$'s instead of $X_i$'s is called the blocks-of-blocks bootstrap. Monte Carlo experiments have shown that the blocks-of-blocks method provides greater finite-sample accuracy than does sampling blocks of $X_i$'s, but the two methods have the same (higher-order) asymptotic properties.

It is also possible to use overlapping blocks with lengths that are sampled randomly from the geometric distribution (Politis & Romano, 1993). The block bootstrap with random block lengths is also called the *stationary bootstrap* because the resulting bootstrap data series is stationary, whereas it is not with overlapping or non-overlapping blocks of non-stochastic lengths.

As will be explained in Section 3.2.1, the RMSEs of bootstrap estimators of distribution functions are smaller with overlapping blocks than with non-overlapping ones. This suggests that overlapping blocks are preferred for applications, although the differences between the numerical results obtained with the two types of blocking are often very small (Andrews, 2002b) The rates of convergence of the errors made with overlapping and non-overlapping blocks are the same (see Section 3.2.1), and theoretical arguments are often simpler with non-overlapping blocks than with overlapping ones. Therefore, many of the arguments in the remainder of this paper will be based on non-overlapping blocks with the understanding that the results also apply to overlapping blocks.

The errors made by the stationary bootstrap are larger than those of the bootstrap with non-stochastic block lengths and either overlapping or non-overlapping blocks (Lahiri, 1999). Therefore, the stationary bootstrap is unattractive relative to the bootstrap with non-stochastic block lengths. In the remainder of this paper, we assume that block lengths are non-stochastic.

Regardless of whether the blocks are overlapping or non-overlapping, the block length must increase with increasing sample size $n$ to make bootstrap estimators of moments and distribution functions consistent (Carlstein, 1986; Künsch, 1989; Hall *et al.*, 1995). The block length must also increase with increasing $n$ to enable the block bootstrap to achieve asymptotically correct coverage probabilities for confidence intervals and rejection probabilities for tests. When the objective is to estimate a moment or distribution function, the asymptotically optimal block length may be defined as the one that minimizes the asymptotic mean-square error of the block bootstrap estimator. When the objective is to form a confidence interval or test a hypothesis, the asymptotically optimal block length may be defined as the one that minimizes the ECP of the confidence interval or ERP of the test. The asymptotically optimal block length and the corresponding rates of convergence of block bootstrap estimation errors, ECP's and ERP's depend on what is being estimated (e.g., bias, a one-sided distribution function, a symmetrical distribution function, etc.). The optimal block lengths and the rates of convergence of block bootstrap estimation errors with non-stochastic block lengths are discussed in detail in Section 3.2. Before presenting these results, however, it is necessary to deal with certain problems that arise in centering and Studentizing statistics based on the block bootstrap. These issues are discussed in Section 3.1.

### *3.1 Centering and Studentizing with the Block Bootstrap*

Two problems are treated in this section. The first is the construction of a block bootstrap version of the centered statistic $\Delta_n \equiv \theta(m_n) - \theta(\mu)$ that does not have excessive bias. This is discussed in Section 3.1.1. The second problem is Studentization of the resulting block bootstrap version of $n^{1/2}\Delta_n$. This is topic of Section 3.2.2.

#### *3.1.1 Centering*

The problem of centering and its solution can be seen most simply by assuming that $X_i$ is a scalar and $\theta$ is the identity function. Thus, $\Delta_n = m_n - \mu$ and $\mathbf{E}\Delta_n = 0$. An obvious block bootstrap version of $\Delta_n$ is $\hat{\Delta}_n = \hat{m}_n - m_n$, where $\hat{m}_n = n^{-1}\sum_{i=1}^{n}\hat{X}_i$, and $\{\hat{X}_i\}$ is the block bootstrap sample using either non-overlapping or overlapping blocks. Let $\hat{\mathbf{E}}$ denote the expectation operator with respect to the probability measure induced by block bootstrap sampling. If the blocks are non-overlapping, then $\hat{\mathbf{E}}\hat{m}_n = m_n$, so $\hat{\mathbf{E}}\hat{\Delta}_n = 0$. With overlapping blocks, however, observations $X_i$ that are within one block length of either end of the data occur in fewer blocks and, therefore, are sampled with lower probability than observations in the middle of the data. Because of this edge effect, $\hat{\mathbf{E}}\hat{m}_n \neq m_n$ with overlapping blocks. It can be shown that

$$\hat{\mathbf{E}}\hat{m}_n = m_n + [\ell(n-\ell-1)]^{-1}[\ell(\ell-1)m_n - \tau_1 - \tau_2],$$

where $\ell$ is the block length $\tau_1 = \sum_{j=1}^{\ell-1}(\ell-j)X_j$, $\tau_2 = \sum_{j=n-\ell-2}^{n}[j-(n-\ell+1)]X_j$, and it is assumed for simplicity that $n$ is an integer multiple of $\ell$ (Hall *et al.*, 1995). The bias caused by the fact that $\hat{\mathbf{E}}\hat{m}_n \neq m_n$ decreases the rate of convergence of the estimation errors of the overlapping-blocks bootstrap. This problem can be solved by centering the overlapping-blocks bootstrap estimator at $\hat{\mathbf{E}}\hat{m}_n$ instead of at $m_n$. The resulting bootstrap version of $\Delta_n$ is $\hat{\Delta}_n = \hat{m}_n - \hat{\mathbf{E}}\hat{m}_n$. More generally, the block bootstrap version of $\theta(m_n) - \theta(\mu)$ is $\theta(\hat{m}_n) - \theta(\hat{\mathbf{E}}\hat{m}_n)$. This centering can also be used with non-overlapping blocks because $\hat{\mathbf{E}}\hat{m}_n = m_n$ with non-overlapping blocks.

#### *3.1.2 Studentization*

This section addresses the problem of Studentizing $n^{1/2}\hat{\Delta}_n = n^{1/2}[\theta(\hat{m}_n) - \theta(\hat{\mathbf{E}}\hat{m}_n)]$. The source of the problem is that blocking distorts the dependence structure of the DGP. As a result, the most obvious methods for Studentizing the bootstrap version $n^{1/2}\Delta_n$ create excessively large estimation errors. Various forms of this problem have been discussed by Lahiri (1992), Davison & Hall (1993), and Hall & Horowitz (1996). The discussion here is based on Hall & Horowitz (1996).

To illustrate the essential issues with a minimum of complexity, assume that the blocks are non-overlapping, $\theta$ is the identity function, and $\{X_i\}$ is a sequence of uncorrelated (though not necessarily independent) scalar random variables. For example, many economic time series are martingale difference sequences and, therefore, serially uncorrelated. Let $V$ denote the variance operator relative to the process that generates $\{X_i\}$. Then $n^{1/2}\Delta_n = n^{1/2}(m_n - \mu)$, $n^{1/2}\hat{\Delta}_n = n^{1/2}(\hat{m}_n - m_n)$ and $V(n^{1/2}\Delta_n) = \mathrm{E}(X_1 - \mu)^2$. The natural choice for $s_n^2$ is the sample variance, $s_n^2 = n^{-1}\sum_{i=1}^n (X_i - m_n)^2$, in which case $s_n^2 - \mathrm{Var}(n^{1/2}\Delta_n) = O_p(n^{-1/2})$.

Now consider Studentization of $n^{1/2}\hat{\Delta}_n$. Let $\ell$ and $B$, respectively, denote the block length and number of blocks, and assume that $B\ell = n$. Let $\hat{V}$ denote the variance operator relative to the block bootstrap DGP. An obvious bootstrap analog of $s_n^2$ is $\hat{s}_n^2 = n^{-1}\sum_{i=1}^n (\hat{X}_i - \hat{m}_n)^2$, which leads to the Studentized statistic $\tilde{T}_n \equiv n^{1/2}\hat{\Delta}_n/\hat{s}_n$. However, $\hat{V}(n^{1/2}\hat{\Delta}_n) = \tilde{s}_n^2$, where

$$\tilde{s}_n^2 = n^{-1}\sum_{b=0}^{B}\sum_{i=1}^{\ell}\sum_{j=1}^{\ell}(X_{b\ell+i} - m_n)(X_{b\ell+j} - m_n)$$

(Hall & Horowitz, 1996). Moreover, $\hat{s}_n^2 - \tilde{s}_n^2 = O[(\ell/n)^{1/2}]$ almost surely. The consequences of this relatively large error in the estimator of the variance of $n^{1/2}\hat{\Delta}_n$ can be seen by carrying out Edgeworth expansions of $\mathbf{P}(n^{1/2}\Delta_n/s_n \le z)$ and $\hat{\mathbf{P}}(n^{1/2}\hat{\Delta}_n/\hat{s}_n \le z)$. These yield $\hat{\mathbf{P}}(n^{1/2}\hat{\Delta}_n/\hat{s}_n \le z) - \mathbf{P}(n^{1/2}\Delta_n/s_n \le z) = O[(\ell/n)^{1/2}]$. The error made by first-order asymptotic approximations is $O(n^{-1/2})$. Thus, the block bootstrap does not provide asymptotic refinements and, in fact, is less accurate than first-order approximations when $n^{1/2}\hat{\Delta}_n$ is Studentized with $\hat{s}_n$.

This problem can be mitigated by Studentizing $n^{1/2}\hat{\Delta}_n$ with $\tilde{s}_n$ (Lahiri, 1996a) or the estimator $n^{-1}\sum_{b=0}^{B-1}\sum_{i=1}^{\ell}\sum_{j=1}^{\ell}(X_{b\ell+i} - \hat{m}_n)(X_{b\ell+j} - \hat{m}_n)$ (Götze & Künsch, 1996). The error in the block bootstrap estimator of a one-sided distribution function is then $o(n^{-1/2})$ almost surely (Lahiri, 1996a; Götze & Künsch, 1996). However, the distributions of the symmetrical probabilities $\mathbf{P}(|n^{1/2}\Delta_n|/s_n \le z)$ and $\hat{\mathbf{P}}(|n^{1/2}\hat{\Delta}_n|/\tilde{s}_n \le z)$ differ by $O(n^{-1})$, so the block bootstrap does not provide asymptotic refinements for symmetrical distributions, confidence intervals and tests.

Refinements for symmetrical distribution functions, confidence intervals, and tests can be obtained by replacing $\tilde{T}_n$ with the "corrected" statistic $\hat{T}_n = \tau_n\tilde{T}_n$ (Hall & Horowitz, 1996; Andrews, 2002a). In the remainder of this paper, $\hat{T}_n$ will be called a "corrected" bootstrap test statistic and $\tau_n$ will be called a correction factor. The estimation errors resulting from the use of corrected statistics are discussed in Section 3.3.

The use of a correction factor can be generalized to the case in which $\theta$ is not the identity function and $X_i$ is a vector. Suppose that $\mathrm{cov}(X_i, X_j) = 0$ whenever $|i - j| > M$ for some $M < \infty$. This assumption is weaker than $M$-dependence, which requires $X_i$ and $X_j$ to be independent when $|i - j| > M$. It holds in time series generated by many models that are important in economics, including rational-expectations models and models of behavior by optimizing agents. When the assumption holds, $s_n^2 = \nabla\theta(m_n)'\Sigma_n\nabla\theta(m_n)$, where

$$\Sigma_n = n^{-1}\sum_{i=1}^{n}\left[(X_i - m_n)(X_i - m_n)' + \sum_{j=1}^{M} H(X_i, X_{i+j}, m_n)\right], \tag{3.1}$$

and $H(X_i, X_{i+j}, m_n) = (X_i - m_n)(X_{i+j} - m_n)' + (X_i - m_n)'(X_{i+j} - m_n)$. The bootstrap version of $s_n^2$ is $\hat{s}_n^2 = \nabla\theta(\hat{m}_n)'\hat{\Sigma}_n\nabla\theta(\hat{m}_n)$, where

$$\hat{\Sigma}_n = n^{-1}\sum_{i=1}^{n}\left[(\hat{X}_i - \hat{m}_n)(\hat{X}_i - \hat{m}_n)' + \sum_{j=1}^{M} H(\hat{X}_i, \hat{X}_{i+j}, \hat{m}_n)\right]. \tag{3.2}$$

Define

$$\tilde{\Sigma}_n = n^{-1} \sum_{b=0}^{B} \sum_{i=1}^{\ell} \sum_{j=1}^{\ell} (X_{b\ell+i} - m_n)(X_{b\ell+j} - m_n)',$$

$\tilde{s}_n^2 = \nabla\theta(m_n)'\tilde{\Sigma}_n\nabla\theta(m_n)$, and $\tau_n = s_n/\tilde{s}_n$. Then with non-overlapping blocks, the corrected block bootstrap version of $n^{1/2}(\theta_n - \theta)/s_n$ is $\tau_n n^{1/2}[\theta(\hat{m}_n) - \theta(m_n)]/\hat{s}_n$ (Hall & Horowitz, 1996). Andrews (2002a) gives the overlapping-blocks version of the statistic.

The foregoing discussion assumes that $\mathrm{cov}(X_i, X_j) = 0$ when $|i - j| > M$ for some $M < \infty$. When this assumption is not made, $\Sigma_n$ must be replaced by a kernel-type estimator of the covariance matrix of $m_n - \mu$. See, for example, Newey & West, 1987, 1994; Andrews, 1991; Andrews & Monahan, 1992; Götze & Künsch, 1996. In contrast to (3.1), kernel covariance matrix estimators are not functions of sample moments. This complicates the analysis of rates of convergence of estimation errors. As was discussed in Section 2.3, this analysis is based on Edgeworth expansions of the distributions of the relevant statistics. The most general results on the validity of such expansions assume that the statistic of interest is a function of sample moments (Götze & Hipp, 1983, 1994; Lahiri, 1996b). Consequently, as will be discussed in Section 3.2, the properties of the block bootstrap are less well understood when Studentization is with a kernel covariance matrix estimator than when Studentization is with a function of sample moments.

### *3.2 The Accuracy of Block Bootstrap Estimates*

This section summarizes results on the magnitudes of the estimation errors made by the block bootstrap. Estimation errors with the block bootstrap are larger than they are with *iid* data. There are two reasons for this. First, independent sampling of blocks does not correspond to the dependence structure of the DGP. Second, the use of blocks as sampling units reduces the effective size of the sample available for estimating the block bootstrap versions of the quantities $\hat{\kappa}$ in the Edgeworth expansions of block-bootstrap statistics. This increases the sampling variation of block bootstrap estimators relative to their *iid* counterparts.

The earliest result on the accuracy of the block bootstrap appears to be due to Carlstein (1986), who gave conditions under which the block bootstrap with non-overlapping blocks provides a consistent estimator of the variance of $\{X_1\}$ based on observations $\{X_i : i = 1, \ldots, n\}$ from a strictly stationary time series. For the special case of an AR(1) DGP, Carlstein also calculated the block length that minimizes the asymptotic mean-square error (AMSE) of the variance estimator (the asymptotically optimal block length). This length increases at the rate $\ell \propto n^{1/3}$. The corresponding AMSE of the variance estimator is $O(n^{-2/3})$.

Künsch (1989) investigated the use of the block bootstrap to estimate a distribution function. He gave conditions under which the overlapping-blocks bootstrap consistently estimates the CDF of a sample average. Lahiri (1991, 1992) was the first to investigate the ability of the overlapping-blocks bootstrap to provide asymptotic refinements for estimation of the CDF of a normalized function of a sample mean. He also investigated refinements for the CDF of a Studentized function of a sample mean for the special case in which the DGP is $M$-dependent. Lahiri (1991, 1992) gave conditions under which the error in the bootstrap estimator of the one-sided distribution function of a normalized or Studentized function of a sample mean is $o(n^{-1/2})$ almost surely. In contrast, the errors made by first-order asymptotic approximations are $O(n^{-1/2})$, so the overlapping-blocks bootstrap provides an asymptotic refinement.

Lahiri's results were refined and extended by Hall *et al.* (1995) and Zvingelis (2001), who give exact rates of convergence of block bootstrap estimators of certain moments, one-sided and symmetrical distribution functions, and the ERP's of tests. We first summarize the results of Hall *et al.* (1995). The following notation will be used. Define $\Delta_n = \theta(m_n) - \theta(\mu)$, $\beta = \mathbf{E}\Delta_n$, and

$\sigma_\infty^2 = V(n^{1/2}\Delta_n)$, where the moments are assumed to exist. Define the block bootstrap analogs $\hat{\Delta}_n = \theta(\hat{m}_n) - \theta(\hat{\mathbf{E}}\hat{m}_n)$, $\hat{\beta} = \hat{\mathbf{E}}\hat{\Delta}_n$, and $\hat{\sigma}_\infty^2 = \hat{V}(n^{1/2}\hat{\Delta}_n)$. The block bootstrap may use either overlapping or non-overlapping blocks. Define one-sided and symmetrical distribution functions of the normalized statistic $n^{1/2}\Delta_n/\sigma_\infty$ by $F_1(z) = \mathbf{P}(n^{1/2}\Delta_n/\sigma_\infty \leq z)$ and $F_2(z) = \mathbf{P}(n^{1/2}|\Delta_n|/\sigma_\infty \leq z)$. Define bootstrap analogs of $F_1$ and $F_2$ by $\hat{F}_1(z) = \hat{\mathbf{P}}(n^{1/2}\hat{\Delta}_n/\hat{\sigma}_\infty \leq z)$ and $\hat{F}_2(z) = \hat{\mathbf{P}}(n^{1/2}|\hat{\Delta}_n|/\hat{\sigma}_\infty \leq z)$. Finally, let $(\psi, \hat{\psi})$ denote either $(\beta, \hat{\beta})$ or $(\sigma_\infty^2/n, \hat{\sigma}_\infty^2/n)$, and let $\phi$ denote the standard normal density function.

Hall *et al.* (1995) show that there are constants $C_j (j = 1, \ldots, 6)$ such that

$$\mathbf{E}(\psi - \hat{\psi})^2 \sim n^{-2}(C_1\ell^{-2} + C_2 n^{-1}\ell), \tag{3.3}$$

$$\mathbf{E}[F_1(z) - \hat{F}_1(z)]^2 \sim n^{-1}(C_3\ell^{-2} + C_4 n^{-1}\ell^2)\phi(z)^2, \tag{3.4}$$

and

$$\mathbf{E}[F_2(z) - \hat{F}_2(z)])^2 \sim n^{-2}(C_5\ell^{-2} + C_6 n^{-1}\ell^3)\phi(z)^2, \tag{3.5}$$

where the symbol "$\sim$" indicates that the quantity on the right-hand side is the leading term of an asymptotic expansion. The constants $C_j$ do not depend on $n$ or $\ell$. The terms involving $C_1$, $C_3$, and $C_5$ correspond to the bias of the bootstrap estimator, and the terms involving $C_2$, $C_4$, and $C_6$ correspond to the variance. As will be discussed in Section 3.2.1, the variance terms are smaller if the blocks are overlapping than if they are non-overlapping. It follows from (3.3)–(3.5) that the asymptotically optimal block length (in the sense of minimizing the AMSE) is $\ell \propto n^{1/3}$ for bias or variance estimation, $\ell \propto n^{1/4}$ for estimating a one-sided distribution function, and $\ell \propto n^{1/5}$ for estimating a symmetrical distribution function. The corresponding minimum asymptotic RMSE's are proportional to $n^{-4/3}$ for bias or variance estimation, $n^{-3/4}$ for estimating a one-sided distribution function of a normalized statistic, and $n^{-6/5}$ for estimating a symmetrical distribution function of a normalized statistic. These results hold for overlapping and non-overlapping blocks. By contrast, the errors made by first-order asymptotic approximations are $O(n^{-1/2})$ for estimating a bias, variance, or one-sided distribution function; and $O(n^{-1})$ for estimating a symmetrical distribution function. With *iid* data, the bootstrap's asymptotic RMSE's are $O(n^{-3/2})$ for bias or variance estimation, $O(n^{-1})$ for estimating a one-sided distribution function and $O(n^{-3/2})$ for estimating a symmetrical distribution function. Thus, the estimation errors of the block bootstrap converge to zero more rapidly than do the errors made by first-order approximations but less rapidly than do those of the bootstrap for *iid* data. The RMSE of the block bootstrap estimator of a symmetrical distribution function converges only slightly more rapidly than the RMSE of first-order asymptotic theory.

Zvingelis (2001) has extended these results to the ECP's of confidence intervals and the ERP's of hypothesis tests based on Studentized statistics. Zvingelis uses non-overlapping blocks, and he assumes that $\mathrm{cov}(X_i, X_j) = 0$ whenever $|i - j|$ exceeds some finite value. The Studentized statistic that he considers is $T_n = n^{1/2}[\theta(m_n) - \theta(\mu)]/s_n$, where $s_n^2 = \nabla\theta(m_n)'\Sigma_n\nabla\theta(m_n)$ and $\Sigma_n$ is defined in (3.1). The bootstrap analog is $\hat{T}_n = \tau_n n^{1/2}[\theta(\hat{m}_n) - \theta(m_n)]/\hat{s}_n$, where $\hat{s}_n^2 = \nabla\theta(\hat{m}_n)'\hat{\Sigma}_n\nabla\theta(\hat{m}_n)$, $\hat{\Sigma}_n$ is defined in (3.2), and $\tau_n$ is the correction factor defined in Section 3.1.2. Zvingelis defines an asymptotically optimal block length as one that maximizes the rate of convergence of the ECP of a confidence interval or the ERP of a test. To state his results, let $\hat{z}_\alpha$ be the $1 - \alpha$ quantile of the distribution of $\hat{T}_n$ relative to the probability measure induced by block bootstrap sampling. Then

1. The asymptotically optimal block length for a one-sided confidence interval or test is $\ell \propto n^{1/4}$. The corresponding ECP and ERP are $O(n^{-3/4})$. Thus, for example, $\mathbf{P}(T_n \leq \hat{z}_\alpha) - (1 - \alpha) = O(n^{-3/4})$.
2. The asymptotically optimal block length for a symmetrical confidence interval or test is $\ell \propto n^{1/4}$. The corresponding ECP and ERP are $O(n^{-5/4})$. Thus, for example, $\mathbf{P}(|T_n| \leq \hat{z}_{\alpha/2}) - (1 - \alpha) = O(n^{-5/4})$.

The errors made by first-order approximations are $O(n^{-1/2})$ for one-sided confidence intervals and tests, and $O(n^{-1})$ for symmetrical confidence intervals and tests. Thus, the ECP and ERP of block bootstrap confidence intervals and tests converge more rapidly than do the ECP and ERP based on first-order approximations. However, the rates of convergence of the block bootstrap ECP and ERP are slower than the rates obtained with the bootstrap for *iid* data. These rates are $O(n^{-1})$ (sometimes $O(n^{-3/2})$) for one-sided confidence intervals and tests, and $O(n^{-2})$ for symmetrical confidence intervals and tests.

Götze & Künsch (1996), Lahiri (1996a), and Inoue & Shintani (2001) have investigated the application of the block bootstrap to statistics that are Studentized with a kernel covariance matrix estimator and overlapping blocks. In Götze & Künsch (1996) and Lahiri (1996a), the Studentized statistic is $T_n = n^{1/2}\Delta_n/s_n$, where $s_n^2 = \nabla\theta(m_n)'\overline{\Sigma}_n\nabla\theta(m_n)$,

$$\overline{\Sigma}_n = n^{-1}\sum_{i=0}^{\ell}\sum_{j=1}^{n-\ell}[2 - I(i=0)]\omega(i/\ell)(X_i - m_n)(X_{i+j} - m_n)',$$

$I$ is the indicator function, $\omega : [0, 1) \to [0, 1]$ is a kernel function satisfying $\omega(0) = 1$ and $\omega(u) = 0$ if $|u| \geq 1$, and $\ell$ is both the block length and the width of the lag window. The bootstrap version of $T_n$ is $n^{1/2}[\theta(\hat{m}_n) - \theta(\hat{\mathbf{E}}\hat{m}_n)]/\hat{s}_n$, where $\hat{s}_n^2 = \nabla\theta(\hat{m}_n)'\hat{\Sigma}_n\nabla\theta(\hat{m}_n)$,

$$\hat{\Sigma}_n = \sum_{[i/\ell]=[j/\ell]} m_{ij}^{-1}(\hat{X}_i - \hat{m}_n)(\hat{X}_j - \hat{m}_n)',$$

$[u]$ denotes the integer part of $u$, $m_{ij} = n$ in Götze & Künsch (1996), and $m_{ij} = n(1 - |i - j|/\ell)$ in Lahiri (1996a). In Lahiri (1996a), $T_n$ is a Studentized slope coefficient of a linear mean-regression model. Lahiri gives conditions under which

$$D_n \equiv \sup_z |\hat{\mathbf{P}}(\hat{T}_n \leq z) - \mathbf{P}(T_n \leq z)| = o_p(n^{-1/2}).$$

In contrast, the error made by the asymptotic normal approximation is $O(n^{-1/2})$, so the block bootstrap provides an asymptotic refinement. Götze & Künsch (1996) refine this result. They show that the rate of convergence of $D_n$ is maximized by using a rectangular kernel function $\omega$ and setting $\ell \propto n^{1/4}$. This yields

$$\sup_z |\hat{\mathbf{P}}(\hat{T}_n \leq z) - \mathbf{P}(T_n \leq z)| = O_p(n^{-3/4+\varepsilon})$$

for any $\varepsilon > 0$. The rectangular kernel has the disadvantage of not guaranteeing that $\hat{\Sigma}_n$ is positive definite. This problem can be overcome at the price of a slightly slower rate of convergence of $D_n$ by setting $\omega(u) = (1 - u^2)I(|u| \leq 1)$. With this quadratic kernel,

$$\sup_z |\hat{\mathbf{P}}(\hat{T}_n \leq z) - \mathbf{P}(T_n \leq z)| = O_p(n^{-2/3+\varepsilon})$$

for any $\varepsilon > 0$. Götze & Künsch (1996) show that an asymptotic refinement cannot be achieved with the triangular kernel $\omega(u) = (1 - |u|)I(|u| \leq 1)$. Finally, Götze & Künsch (1996) show that

$$\mathbf{P}(T_n \leq \hat{z}_\alpha) = 1 - \alpha + o(n^{-1/2}),$$

where $\hat{z}_\alpha$ is the $1 - \alpha$ quantile of the distribution of $\hat{T}_n$ under block bootstrap sampling. Thus, the block bootstrap provides an asymptotic refinement for the ECP of a one-sided confidence interval and the ERP of a one-sided hypothesis test.

Inoue & Shintani (2001) have extended the results of Götze & Künsch (1996) and Lahiri (1996a). Inoue & Shintani (2001) apply the overlapping-blocks bootstrap to a Studentized estimator of a slope coefficient in a (possibly overidentified) linear model that is estimated by instrumental variables.

They assume that the block length, $\ell$, and width of the lag window are equal, and they require the rate of increase of $\ell$ to be faster than $n^{1/6}$ but slower than $n^{1/4}$. To summarize their results, let $\omega$ be the kernel used in the covariance matrix estimator, and define $q$ to be the largest integer such that

$$\lim_{u \to 0} [1 - \omega(u)]/|u|^q < \infty.$$

Let $T_n$ and $\hat{T}_n$, respectively, denote the Studentized slope estimator and its block-bootstrap analog. Let $\hat{z}_\alpha$ denote the $1 - \alpha$ quantile of the distribution of $\hat{T}_n$ under bootstrap sampling. Inoue & Shintani (2001) give conditions under which

$$\mathbf{P}(T_n \leq \hat{z}_\alpha) = 1 - \alpha + O(\ell/n) + O(\ell^{-q})$$

and

$$\mathbf{P}(|T_n| \leq \hat{z}_{\alpha/2}) = 1 - \alpha + o(\ell/n) + O(\ell^{-q}).$$

Because Inoue & Shintani (2001) require the rate of increase of $\ell$ to exceed $n^{1/6}$, their tightest bound on the ERP of a symmetrical test is $o(a_n n^{-5/6})$ where $a_n \to \infty$ at an arbitrarily slow rate as $n \to \infty$. It is not known whether the bootstrap can achieve an ERP of $o(n^{-1})$ when a kernel covariance matrix estimator is used for Studentization.

### *3.2.1 Relative accuracy of the bootstrap with overlapping and non-overlapping blocks*

Hall *et al.* (1995) and Lahiri (1999) have compared the estimation errors made by the overlapping- and non-overlapping-blocks bootstraps. They find that when the asymptotically optimal block length is used for estimating a bias or variance, the AMSE is with non-overlapping blocks is $1.5^{2/3}$ times the AMSE with overlapping blocks. Thus the AMSE is approximately 31% larger with non-overlapping blocks. The rates of convergence of the AMSE's are equal, however. Hall *et al.* (1995) compare the AMSE's for estimation of a one-sided distribution function of a normalized sample average (that is, for estimating $\mathbf{P}[n^{1/2}(m_n - \mu)/\sigma_\infty \leq z]$). The AMSE is $1.5^{1/2}$ times or 22% larger with non-overlapping blocks than with overlapping ones. The bootstrap is less accurate with non-overlapping blocks because the variance of the bootstrap estimator is larger with non-overlapping blocks than with overlapping ones. The bias of the bootstrap estimator is the same for non-overlapping and overlapping blocks. It should be noted, however, that the differences between the AMSE's with the two types of blocking occurs in higher-order terms of the statistics of interest and, therefore, is often very small in magnitude. Andrews (2002b) provides Monte Carlo evidence on the magnitudes of the errors made with the two types of blocking.

## *3.3 Subsampling*

The block bootstrap's distortions of the dependence structure of a time series can be avoided by using a subsampling method proposed by Politis & Romano (1994) and Politis *et al.* (1999). To describe this method, let $t_n \equiv t_n(X_1, \ldots, X_n)$ be an estimator of the population parameter $\theta$, and set $\Delta_n = \rho(n)(t_n - \theta)$, where the normalizing factor $\rho(n)$ is chosen so that $\mathbf{P}(\Delta_n \leq z)$ converges to a nondegenerate limit at continuity points of the latter. For $i = 1, \ldots, n - \ell + 1$, let $\{X_j : j = i, \ldots, i + \ell - 1\}$ be a subset of $\ell < n$ consecutive observations taken from the sample $\{X_i : i = 1, \ldots, n\}$. Define $N_{n\ell}$ to be the total number of subsets that can be formed. Let $t_{\ell,k}$ denote the estimator $t_\ell$ evaluated at the $k$'th subset. The subsampling method estimates $\mathbf{P}(\Delta_n \leq z)$ by

$$G_{n\ell}(z) \equiv \frac{1}{N_{n\ell}} \sum_{k=1}^{N_{n\ell}} I[\rho(\ell)(t_{\ell,k} - t_n) \leq z].$$

The intuition behind this method is that each subsample is a realization of length $\ell$ of the true DGP. Therefore, $\mathbf{P}(\Delta_\ell \le z)$ is the exact sampling distribution of $\rho(\ell)(t_\ell - \theta)$, and

$$\mathbf{P}(\Delta_\ell \le z) = \mathbf{E}\{I[\rho(\ell)(t_\ell - \theta) \le z]\}. \tag{3.6}$$

$G_{n\ell}(z)$ is the estimator of the right-hand side of (3.6) that is obtained by replacing the population expectation by the average over subsamples and $\theta$ by $t_n$. If $n$ is large but $\ell/n$ is small, then random fluctuations in $t_n$ are small relative to those in $t_\ell$. Accordingly, the sampling distributions of $\rho(\ell)(t_\ell - t_n)$ and $\rho(n)(t_n - \theta)$ are close. Similarly, if $N_{n\ell}$ is large, the average over subsamples is a good approximation to the population average. These ideas were formalized by Politis & Romano (1994), who give conditions under which the subsampling method consistently estimates $\mathbf{P}(\Delta_n \le z)$ and the coverage probability of a confidence interval for $\theta$.

Hall & Jing (1996) investigated the accuracy of the subsampling method for estimating one-sided and symmetrical distribution functions of a Studentized, asymptotically normal statistic. Hall & Jing (1996) find that when $\ell$ is chosen optimally, the rates of convergence of the RMSE's are $n^{-1/4}$ and $n^{-1/3}$, respectively, for one-sided and symmetrical distribution functions. These rates are slower than those of first-order asymptotic approximations and the block bootstrap. Hall & Jing (1996) also describe an extrapolation technique that accelerates the rate of convergence of the RMSE. With extrapolation, the rate of convergence of $\hat{\mathbf{P}}(t_n \le z) - \mathbf{P}(t_n \le z)$ is $O_p(n^{-2/3})$, and the rate of convergence of $\hat{\mathbf{P}}(|t_n| \le z) - \mathbf{P}(|t_n| \le z)$ is $O_p(n^{-8/7})$. These rates are faster than those of first-order approximations but slower than those provided by the block bootstrap. Thus, in terms of rates of convergence of estimation errors, subsampling with or without extrapolation does not improve on the block bootstrap.

### 3.4 *Modifications of the Block Bootstrap*

There have been several attempts to improve the performance of the block bootstrap by reducing the influence of the "discontinuities" in the bootstrap data series that occur at block boundaries. Carlstein *et al.* (1998) proposed sampling blocks according to a data-based Markov chain so as to increase the likelihood that consecutive blocks match at their ends. They gave conditions under which this *matched-block bootstrap* (MBB) reduces the bias of a bootstrap estimator of a variance. However, the MBB increases the rate of convergence of the bias only if the DGP is a Markov process. The MBB does not reduce the variance of the estimator. Carlstein *et al.* (1998) did not investigate the performance of the MBB for estimating a distribution function, the coverage probability of a confidence interval, or the rejection probability of a hypothesis test.

Paparoditis & Politis (2001) proposed downweighting points $\hat{X}_i$ in the block bootstrap data series that are near block endpoints. They give conditions under which this *tapered bootstrap* procedure increases the rate of convergence of the MSE of the bootstrap estimator of $\mathrm{Var}(n^{1/2}m_n)$ to $O(n^{-4/5})$. This may be contrasted with the rate $O(n^{-2/3})$ that is provided by the block bootstrap without tapering. Paparoditis & Politis (2001) did not investigate the ability of the tapered block bootstrap to provide asymptotic refinements for distribution estimation, ECP's of confidence intervals, or ERP's of hypothesis tests. Thus, it remains unknown whether the MBB or tapered bootstrap can increase the rates of convergence of the errors made by the block bootstrap for estimating distribution functions, coverage probabilities, and rejection probabilities.

Andrews (2002b) has proposed modifying statistics computed from the original data so as to mimic the discontinuities of the block bootstrap sample. This is done by deleting small blocks of data from moment calculations, thereby introducing a blocking effect into the statistic based on the original data. Andrews calls this procedure the *block-block bootstrap*. He gives conditions under which it reduces the ECP's and ERP's of block-bootstrap-based confidence intervals and tests without reducing asymptotic local power. Specifically, for any $\xi < 1/2$, the ECPs and ERPs are

$o[n^{-(1+\xi)}]$ for symmetrical confidence intervals and tests based on the block-block bootstrap. They are $o[n^{-(1/2+\xi)}]$ for one-sided tests and confidence intervals. By contrast, the errors made with the standard block bootstrap are $O(n^{-5/4})$ for symmetrical confidence intervals and tests and $O(n^{-3/4})$ for one-sided confidence intervals and tests. Thus, the block-block bootstrap improves on the standard block bootstrap. Andrews (2002b) presents Monte Carlo evidence indicating that the block-block bootstrap also has improved finite-sample performance.

## 4 Methods That Impose Stronger Restrictions on the DGP

This section describes bootstrap methods that make stronger *a priori* assumptions about the DGP than does the block bootstrap. In return for stronger assumptions, some of these methods achieve faster rates of convergence of estimation errors than does the block bootstrap. They do this by using the assumed structure of the DGP to increase the estimation accuracy of the quantities $\hat{\kappa}$ that enter the Edgeworth expansions of bootstrap statistics.

### *4.1 The Sieve Bootstrap for Linear Processes*

A substantial improvement over the performance of the block bootstrap is possible if the DGP is known to be a linear process. That is, the DGP has the form

$$X_i - \mu = \sum_{j=1}^{\infty} \alpha_j (X_{i-j} - \mu) + U_i, \tag{4.1}$$

where $\mu = \mathrm{E}(X_i)$ for all $i$, $\{U_i\}$ is a sequence of *iid* random variables, and $\{X_i\}$ may be a scalar or a vector process. Assume that $\sum_{j=1}^{\infty} \alpha_j^2 < \infty$ and that all of the roots of the power series 1 - $\sum_{j=1}^{\infty} \alpha_j z^j$ are outside of the unit circle. Bühlmann (1997, 1998), Kreiss (1988, 1992), and Paparoditis (1996) proposed approximating (4.1) by an AR($p$) model in which $p = p(n)$ increases with increasing sample size. Let $\{a_{nj} : j = 1, \ldots, p\}$ denote least squares or Yule–Walker estimates of the coefficients of the approximating process, and let $\{U_{nj}\}$ denote the centered residuals. The sieve bootstrap consists of generating bootstrap samples according to the process

$$\hat{X}_i - m = \sum_{j=1}^{p} a_{nj} (\hat{X}_{i-j} - m) + \hat{U}_j, \tag{4.2}$$

where $m = n^{-1} \sum_{i=1}^{n} X_i$ and the $\hat{U}_j$ are sampled randomly with replacement from the $U_{nj}$. Bühlmann (1997), Kreiss (1992, 2000), and Paparoditis (1996) give conditions under which this procedure consistently estimates the distributions of sample averages, sample autocovariances and autocorrelations, and the regression coefficients $a_{nj}$ among other statistics.

Choi & Hall (2000) investigated the ability of the sieve bootstrap to provide asymptotic refinements to the coverage probability of a one-sided confidence interval for the mean of a linear statistic when the $X_i$'s are scalar random variables. A linear statistic has the form

$$\theta_n = (n - q + 1)^{-1} \sum_{i=1}^{n-q+1} G(X_i, \ldots, X_{i+q-1}),$$

where $q \geq 1$ is a fixed integer and $G$ is a known function. Define $\theta = \mathrm{E}[G(X_1, \ldots, X_q)]$. Choi & Hall (2000) considered the problem of finding a one-sided confidence interval for $\theta$, although they note that their conclusions also apply to a one-sided confidence interval for $H(\theta)$, where $H$ is a continuously differentiable function with $H'(\theta) \neq 0$.

To obtain a confidence interval for $\theta$, define $\sigma_\theta^2 = \mathrm{Var}[(n-q+1)^{1/2}(\theta_n - \theta)]$. Then

$$\sigma_\theta^2 = \gamma_G(0) + 2\sum_{j=1}^{n-q}[1 - j(n-q+1)^{-1}]\gamma_G(j),$$

where $\gamma_G(j) = \mathrm{cov}[G(X_1, \ldots, X_q), G(X_{j+1}, \ldots, X_{j+q})]$. The bootstrap analogs of $\theta_n$ and $\theta$ are

$$\hat{\theta}_n = (n-q+1)^{-1}\sum_{i=1}^{n-q+1} G(\hat{X}_i, \ldots, \hat{X}_{i+q-1})$$

and $\hat{\theta} = \hat{\mathbf{E}}[G(\hat{X}_1, \ldots, \hat{X}_{1+q})]$, respectively, where $\hat{\mathbf{E}}$ is the expectation operator relative to the probability distribution induced by sieve bootstrap sampling. The variance of $(n-q+1)^{1/2}(\hat{\theta}_n - \hat{\theta})$ conditional on the data is

$$\hat{\sigma}_\theta^2 = \hat{\gamma}_G(0) + 2\sum_{j=1}^{n-q}[1 - j(n-q+1)^{-1}]\hat{\gamma}_G(j),$$

where $\hat{\gamma}_G(j) = \mathrm{cov}[G(\hat{X}_1, \ldots, \hat{X}_q), G(\hat{X}_{j+1}, \ldots, \hat{X}_{j+q})]$ and the covariance is relative to the bootstrap probability distribution. Define $\hat{T}_n = (n-q+1)^{1/2}(\hat{\theta}_n - \hat{\theta})/\hat{\hat{\sigma}}_\theta$, where $\hat{\hat{\sigma}}_\theta$ is a bootstrap estimator of $\hat{\sigma}_\theta$, possibly obtained through a double bootstrap procedure. Let $\hat{z}_{n\alpha}$ satisfy $\hat{\mathbf{P}}(\hat{T}_n \leq \hat{z}_{n\alpha}) = 1 - \alpha$, where $\hat{\mathbf{P}}$ denotes the probability measure induced by sieve bootstrap sampling. Then a lower $100(1-\alpha)$ percent confidence interval for $\theta$ is $\theta_n - (n-q+1)^{-1/2}\hat{z}_{n\alpha}\hat{\sigma}_\theta \leq \theta$, and an upper $100(1-\alpha)$ percent confidence interval is $\theta_n - (n-q+1)^{-1/2}\hat{z}_{n,1-\alpha}\hat{\sigma}_\theta \geq \theta$. Choi & Hall (2000) gave conditions under which the ECP's of these intervals are $O(n^{-1+\varepsilon})$ for any $\varepsilon > 0$. This is only slightly larger than the ECP of $O(n^{-1})$ that is available with *iid* data. Thus, relative to the block bootstrap, the sieve bootstrap provides a substantial improvement in the rate of convergence of the ECP when the DGP is a linear process.

### 4.2 *The Bootstrap for Markov Processes*

Estimation errors that converge more rapidly than those of the block bootstrap are also possible if the DGP is a (possibly higher-order) Markov process or can be approximated by such a process. The class of Markov and approximate Markov processes contains many nonlinear autoregressive, ARCH, and GARCH processes, among others, that are important in applications.

When the DGP is a Markov process, the bootstrap can be implemented by estimating the Markov transition density nonparametrically. Bootstrap samples are generated by the stochastic process implied by the estimated transition density. Call this procedure the Markov bootstrap (MB). The MB was proposed by Rajarshi (1990), who gave conditions under which it consistently estimates the asymptotic distribution of a statistic. Datta & McCormick (1995) gave conditions under which the error in the MB estimator of the distribution function of a normalized sample average is almost surely $o(n^{-1/2})$. Hansen (1999) proposed using an empirical likelihood estimator of the Markov transition probability but did not prove that the resulting version of the MB is consistent or provides asymptotic refinements. Chan & Tong (1998) used the MB in a test for multimodality in the distribution of dependent data. Paparoditis & Politis (2002) proposed estimating the Markov transition probability by resampling the data in a suitable way.

Horowitz (2003) investigated the ability of the MB to provide asymptotic refinements for confidence intervals and tests based on Studentized statistics. To describe the results, let the DGP be an order $q$ Markov process that is stationary and GSM. For data $\{X_i : i = 1, \ldots, n\}$ and any $j > q$, define $Y_j = (X_{j-1}, \ldots, X_{j-q})$. Define $\mu = \mathbf{E}(X_1)$ and $m = n^{-1}\sum_{i=1}^n X_i$. Horowitz (2003) considered a confidence interval for $H(\mu)$, where $H$ is a sufficiently smooth function. The confidence interval is based on the statistic

$$T_n = n^{1/2}[H(m) - H(\mu)]/s_n$$

where $s_n^2$ is a consistent estimator of the variance of the asymptotic distribution of $n^{1/2}[H(m)-H(\mu)]$. Thus, for example, a symmetrical $1-\alpha$ confidence interval is

$$H(m) - \hat{z}_{n\alpha}s_n \leq H(\mu) \leq H(m) + \hat{z}_{n\alpha}s_n, \tag{4.3}$$

where $\hat{z}_{n\alpha}$ is the MB estimator of the $1-\alpha$ quantile of the distribution of $|T_n|$. Horowitz (2003) used a kernel estimator of the Markov transition density to implement the MB. The estimator is $f_n(x|y) = p_{nz}(x,y)/p_{ny}(y)$, where

$$p_{nz}(x,y) = \frac{1}{(n-q)h_n^{d(q+1)}} \sum_{j=q+1}^{n} K_f\left(\frac{x-X_j}{h_n}, \frac{y-Y_j}{h_n}\right),$$

$$p_{ny}(y) = \frac{1}{(n-q)h_n^{dq}} \sum_{j=q+1}^{n} K_p\left(\frac{y-Y_j}{h_n}\right),$$

and $K_f$ and $K_p$ are kernel functions. The MB procedure samples the process induced by a version of $f_n(x|y)$ that is trimmed to keep the bootstrap process away from regions where the probability density of $Y$ is close to zero. Define $C_n = \{y : p_{ny}(y) \geq \lambda_n\}$, where $\lambda_n \to 0$ at a suitable rate as $n \to \infty$. Then the MB process of Horowitz (2003) consists of the following steps:

MB 1. Draw $\hat{Y}_{q+1} \equiv (\hat{X}_q, \ldots, \hat{X}_1)$ from the distribution whose density is $p_{ny}$. Retain $\hat{Y}_{q+1}$ if $\hat{Y}_{q+1} \in C_n$. Otherwise, discard the current $\hat{Y}_{q+1}$ and draw a new one. Continue this process until a $\hat{Y}_{q+1} \in C_n$ is obtained.

MB 2. Having obtained $\hat{Y}_j \equiv (\hat{X}_{j-1}, \ldots, \hat{X}_{j-q})$ for any $j \geq q+2$, draw $\hat{X}_j$ from the distribution whose density is $f_n(\cdot|\hat{Y}_j)$ . Retain $\hat{X}_j$ and set $\hat{Y}_{j+1} = (\hat{X}_j, \ldots, \hat{X}_{j-q+1})$ if $(\hat{X}_j, \ldots, \hat{X}_{j-q+1}) \in C_n$. Otherwise, discard the current $\hat{X}_j$ and draw a new one. Continue this process until an $\hat{X}_j$ is obtained for which $(\hat{X}_j, \ldots, \hat{X}_{j-q+1}) \in C_n$.

MB 3. Repeat step 2 until a bootstrap data series $\{\hat{X}_j : j = 1, \ldots, n\}$ has been obtained. Compute the bootstrap test statistic $\hat{T}_n \equiv n^{1/2}[H(\hat{m}) - H(\hat{\mu})]/\hat{s}_n$, where $\hat{m} = n^{-1}\sum_{j=1}^{n} \hat{X}_j$, $\hat{\mu}$ is the mean of $X$ relative to the distribution induced by the sampling procedure of steps MB 1 and MB 2 (bootstrap sampling), and $\hat{s}_n^2$ is a consistent estimator of the variance of the asymptotic distribution of $n^{1/2}[H(\hat{m}) - H(\hat{\mu})]$ under bootstrap sampling.

MB 4. Set $\hat{z}_{n\alpha}$ equal to the $1-\alpha$ quantile of the empirical distribution of $|\hat{T}_n|$ that is obtained by repeating steps MB 1–MB 3 many times.

Horowitz (2003) gives conditions under which the ECP of the resulting symmetrical confidence interval (4.3) is $O(n^{-3/2+\varepsilon})$ for any $\varepsilon > 0$. The ECP for a one-sided confidence interval is $O(n^{-1+\varepsilon})$. The corresponding ECP's for confidence intervals based on the block bootstrap with the asymptotically optimal block length are $O(n^{-4/3})$ and $O(n^{-3/4})$ , respectively. Thus, the ECP's of the MB converge to zero more rapidly than do those of the block bootstrap.

These results also apply to approximate Markov processes. An approximate Markov process is a process whose density conditional on the past can be approximated by its density conditional on lags of order up to $q$ with an error of $O(e^{-bq})$ for some $b > 0$. The MB for an approximate Markov process is implemented by generating bootstrap samples from the estimated order $q$ Markov process, where $q \to \infty$ at a logarithmic rate as $n \to \infty$.

One disadvantage of the MB is that is suffers from a curse of dimensionality in that it requires the Markov transition density to have an increasing number of derivatives as $q$ and $d$ (the dimension of $X_i$) increase. When $q$ or $d$ is large, this reduces the attractiveness of the MB compared to the block bootstrap, which does not have a similar problem. The curse of dimensionality does not necessarily preclude application of the MB to approximate Markov processes since these can often be approximated satisfactorily with order 1 or 2 Markov processes. An example is given in Section 5.

### *4.3 The Nonparametric Autoregressive Bootstrap*

In this section it is assumed that the DGP is the autoregressive process

$$X_i = g(X_{i-1}, \ldots, X_{i-p}) + \sigma(X_{i-1}, \ldots, X_{i-q})U_i; i = 0, 1, 2, \ldots, \tag{4.4}$$

where $g$ and $\sigma$ are unknown functions, $\{U_i\}$ is a sequence of *iid* random variables with zero mean and unit variance, and $p, q \geq 1$ are integers. Equation (4.4) is a nonparametric version of an ARCH model. Allowing the possibility of a non-constant conditional variance function is important in applications in finance (Engle, 1982; Bollerslev *et al.*, 1992; Gouri éroux, 1997).

The DGP (4.4) is a Markov process to which the methods of Section 4.2 may be applied. In this section, however, we describe a procedure due to Franke, Kreiss & Mammen (2002) (hereinafter FKM) that takes advantage of the structure (4.4). This procedure generates bootstrap samples by replacing $g$, $\sigma$, and $U_i$ with nonparametric estimates. It is assumed that $\{X_i\}$ is strictly stationary and GSM. See Diebolt & Guegan (1990), Meyn & Tweedie (1993), and FKM for regularity conditions that insure these properties.

FKM use Nadaraya–Watson kernel estimators of $g$ and $\sigma$. Other estimators such as local polynomials could also be used. In the case that $p = q = 1$, the estimators are

$$g_{nh}(x) = \frac{1}{(n-1)p_{nh}(x)} \sum_{i=1}^{n-1} X_{i+1} K\left(\frac{x - X_i}{h_n}\right) \tag{4.5}$$

and

$$\sigma_{nh}^2(x) = \frac{1}{(n-1)p_{nh}(x)} \sum_{i=1}^{n-1} [X_{i+1} - g_{nh}(X_i)]^2 K\left(\frac{x - X_i}{h_n}\right), \tag{4.6}$$

where $K$, the kernel, is a probability density function,

$$p_{nh}(x) = \frac{1}{(n-1)} \sum_{i=1}^{n-1} K\left(\frac{x - X_i}{h_n}\right), \tag{4.7}$$

and $\{h_n\}$ is a sequence of positive constants (bandwidths) that converges to zero as $n \to \infty$. Product kernels may be used when $p > 1$ or $q > 1$.

Let $r = \max(p, q)$. As with the Markov bootstrap, the procedure used to generate bootstrap samples based on (4.4) must avoid regions in which the density of $(X_{i-1}, \ldots, X_{i-r})$ is close to zero. In addition, the estimates of $g$ and $\sigma$ that are used to generate bootstrap samples must be based on a bandwidth $\hat{h}_n$ that satisfies $\hat{h}_n / h_n \to \infty$ as $n \to \infty$. To this end, let $g_{n\hat{h}}$ and $\sigma_{n\hat{h}}$, respectively, denote estimates of $g$ and $\sigma$ based on the bandwidth $\hat{h}_n$. Let $\tilde{g}_{n\hat{h}}$ and $\tilde{\sigma}_{n\hat{h}}$ denote $g_{n\hat{h}}$ and $\sigma_{n\hat{h}}$ restricted to a compact set $C$ on which the density of $(X_{i-1}, \ldots, X_{i-r})$ exceeds zero. Let $\{\nu_i\}$ denote the residuals $\{X_i - g_{n\hat{h}}(X_{i-1}, \ldots, X_{i-q})\}$ corresponding to points $(X_{i-1}, \ldots, X_{i-r}) \in C$. Bootstrap samples are generated by the recursion

$$\hat{X}_i = \tilde{g}_{n\hat{h}}(\hat{X}_{i-1}, \ldots, \hat{X}_{i-p}) + \tilde{\sigma}_{n\hat{h}}(\hat{X}_{i-1}, \ldots, \hat{X}_{i-q})\hat{U}_i,$$

where $\{\hat{U}_i\}$ is obtained by sampling the standardized $\nu_i$'s randomly with replacement. The process is initialized by setting $(\hat{X}_r, \ldots, \hat{X}_1) = (X_r, \ldots, X_1)$. Bootstrap estimators of $g$ and $\sigma$ are obtained by applying (4.5)–(4.7) with the bandwidth $h_n$ to the bootstrap sample. Denote the resulting bootstrap estimators of $g$ and $\sigma$ by $\hat{g}_{nh}$ and $\hat{\sigma}_{nh}$.

Let $\hat{\mathbf{P}}$ denote the probability measure induced by bootstrap sampling. Let $y_p$ and $y_q$, respectively, denote the vectors $(x_p, \ldots, x_1)$ and $(x_q, \ldots, x_1)$. FKM give conditions under which

$$\sup_z \left| \hat{\mathbf{P}}\{(nh_n^p)^{1/2}[\hat{g}_{nh}(y_p) - \tilde{g}_{n\hat{h}}(y_p)] \leq z\} - \mathbf{P}\{(nh_n^p)^{1/2}[g_{nh}(y_p) - g(y_p)] \leq z\} \right| \to^p 0$$

and

$$\sup_z \left| \hat{\mathbf{P}}\{(nh_n^q)^{1/2}[\hat{s}_{nh}(y_q) - \tilde{s}_{n\hat{h}}(y_q)] \leq z\} - \mathbf{P}\{(nh_n^q)^{1/2}[s_{nh}(y_q) - \sigma(y_q)] \leq z\} \right| \to^p 0.$$

Hafner (1996) and Härdle *et al.* (2001) use the nonparametric autoregressive bootstrap to construct pointwise confidence intervals for $g$ and $\sigma$. Franke, Kreiss, Mammen & Neumann (2002) obtain uniform confidence bands for $g$ and carry out inference about the parameters of a misspecified (finite-dimensional) parametric model. Neumann & Kreiss (1998) use a wild-bootstrap version of the nonparametric autoregressive bootstrap to test parametric models of the conditional mean and variance functions. They also give conditions under which the ECP of a uniform confidence band for $g$ converges to zero at the rate $(\log n)^{3/2}/(nh_n)^{1/4}$. In contrast, first order asymptotic approximations yield the rate $(\log n)^{-1}$ (Hall, 1991). Kreiss, Neumann & Yao (1998) also use the nonparametric autoregressive bootstrap for testing.

### *4.4 The Periodogram Bootstrap*

In this section, it is assumed that the data are generated by a stationary, univariate process with mean zero and the possibly infinite-order moving average representation

$$X_i = \sum_{k=-\infty}^{\infty} b_k \xi_{i-k}. \tag{4.8}$$

The innovations $\{\xi_i\}$ are *iid* with $\mathbf{E}(\xi_i) = 0$, $\mathbf{E}(\xi_i^2) = 1$, and $\mathbf{E}|\xi_i|^5 < \infty$. It is also assumed that the coefficients $\{b_k\}$ satisfy $\sum_{k=-\infty}^{\infty} |kb_k| < \infty$.

Franke & Härdle (1992) (hereinafter FH) proposed using a frequency domain version of the bootstrap to estimate the asymptotic distribution of a kernel estimator of the spectral density function. Nordgaard (1992) and Theiler *et al.* (1994) considered frequency domain bootstrap methods for Gaussian processes. FH do not assume that the DGP is Gaussian. They give conditions under which the frequency-domain bootstrap consistently estimates the asymptotic distribution of a centered, scaled kernel estimator of the spectral density at any specified frequency.

Dahlhaus & Janas (1996) (hereinafter DJ) applied a modified version of the FH procedure to a class of *ratio statistics* that includes sample autocorrelation coefficients and a normalized estimate of the spectral density function. DJ give conditions under which the difference between the CDF of a normalized ratio statistic and the CDF of a bootstrap analog is uniformly $o(n^{-1/2})$ almost surely. By contrast, the error made by the asymptotic normal approximation to the CDF of a normalized ratio statistic is $O(n^{-1/2})$, so the frequency domain bootstrap provides an asymptotic refinement under the conditions of DJ (1996). Some of these conditions are quite restrictive. In particular, $X_i$ must have a known mean of zero, and the innovations must satisfy $\mathbf{E}(\xi_i^3) = 0$. DJ point out that the asymptotic refinement is not achieved if these conditions do not hold. In addition, DJ assume that the variance of the asymptotic distribution of the ratio statistic is known.

Kreiss & Paparoditis (2000) describe a procedure called the *autoregressive aided periodogram bootstrap* (AAPB). In this procedure, bootstrap samples $\{\hat{X}_i\}$ are generated using the sieve procedure of Section 4.1. These samples are used to calculate a bootstrap version of the periodogram and bootstrap versions of statistics that are functionals of the periodogram. Kreiss & Paparoditis (2000) give conditions under which the AAPB procedure consistently estimates the distribution of ratio statistics and statistics based on the integrated periodogram. These conditions are sufficiently general to permit application of the AABP to moving average processes that are not invertible and, therefore, do not have an autoregressive representation.

## 5 Monte Carlo Evidence

This section presents the results of Monte Carlo experiments that illustrate the finite-sample performance of several of the bootstrap methods that are discussed in Sections 3 and 4. Each experiment consists of finding a nominal 90%, one-sided upper confidence interval for the slope parameter of the linear regression of $X_i$ on $X_{i-1}$ or, equivalently the correlation coefficient of $X_i$ and $X_{i-1}$. The slope parameter is estimated by ordinary least squares (OLS). The experiments evaluate the empirical coverage probabilities of the nominal confidence intervals. We give results for confidence intervals obtained from first-order asymptotic distribution theory, the sieve bootstrap, the overlapping-blocks bootstrap (with $Y_i = (X_i, X_{i-1})$), the Markov bootstrap, and subsampling with the extrapolation procedure of Hall & Jing (1996).

Simulated data were generated by three widely used models: an AR(1) model, an MA(1) model, and an ARCH(1) model. We used parameter values 0.9, 0.75, and 0.3, respectively, for the AR(1), MA(1), and ARCH(1) models. Thus, the DGP's are:

(i) AR(1) process: $X_i = 0.9X_{i-1} + U_i; i = 1, \ldots, n$
(ii) MA(1) process: $X_i = U_i + 0.75U_{i-1}; i = 1, \ldots, n$
(iii) ARCH(1) process: $X_i = U_i(1 + 0.3X_i^2)^{1/2}; i = 1, \ldots, n$

The $U_i$'s are independently distributed as $N(0, 1)$. DGP's (i) and (iii) are first-order Markov processes. DGP (ii) is an approximate Markov process in the sense defined in Section 4.2. In the experiment with the Markov bootstrap, we approximated the MA(1) process with a first-order Markov process. The sieve bootstrap is invalid for DGP (iii) because this DGP is nonlinear and does not have the representation (4.1).

The confidence interval has the form $\theta \leq \theta_n + \hat{z}_{1-\alpha} s_n$, where $\theta_n$ is the OLS estimate of the slope parameter $\theta$, $s_n$ is a standard error, and $\hat{z}_{1-\alpha}$ is an estimate of the $\alpha$ quantile of the distribution of $(\theta_n - \theta)/s_n$. The estimates $\hat{z}_{1-\alpha}$ were obtained from first-order asymptotic theory (the asymptotic normal approximation), the sieve bootstrap, the block bootstrap, the Markov bootstrap, and subsampling with extrapolation. We used the following formulae for $s_n$:

(i) AR(1) process: $s_n^2 = (1 - \theta_n^2)/n$
(ii) MA(1) process: $s_n^2 = (1 - 3\theta_n^2 + 4\theta_n^4)/n$
(iii) ARCH(1) process: $s_n^2 = B_n/A_n^2$, where $A_n = \sum_{i=1}^n (X_i - \overline{X})^2$, $\overline{X} = \sum_{i=1}^n X_i$, and $B_n = A_n + (1 - n/A_n)\sum_{i=1}^n X_i^2(X_i - \overline{X})^2$.

These formulae are sample analogs of the variances of the asymptotic distributions of $\theta_n$ for the three DGPs. Since they assume knowledge of the DGP, they are not realistic for applied research. However, they permit the experiments to focus on the relative performance of different bootstrap methods, rather than the complexities and often poor finite-sample performance of more realistic methods for obtaining standard errors.

Implementation of the sieve bootstrap requires choosing the order of the sieve (the parameter $p$ in (4.2)). Following a suggestion of Choi & Hall (2000), we set $p$ equal to twice the value that maximizes the AIC criterion. The Markov bootstrap requires choosing the order of the Markov process and the bandwidth for use in estimating the transition density. We used a first-order Markov process. Following a suggestion of Silverman (1986, Section 4.3.2), we pre-whitened the simulated data for $(X_i, X_{i-1})$ by linearly transforming them to have a unit empirical covariance matrix. We then estimated the density of the transformed variables using the bandwidth $h_n = 0.85n^{-1/6}$. Finally, we transformed the random variables with the estimated density so that their covariance matrix was the same as the empirical covariance matrix of $(X_i, X_{i-1})$ in the data. The standard normal density was used for the kernel. Satisfactory data-based methods for selecting block lengths for block-bootstrap confidence intervals are not yet available. Nor are satisfactory methods available for selecting subsample sizes for subsampling with extrapolation. Accordingly, we report results obtained with a range of block lengths and subsample sizes.

We carried out experiments using sample sizes of $n = 100$ and $n = 500$. These sample sizes are common in economics and finance among other applications fields. Bootstrap (or subsampling) estimates were based on 1000 bootstrap samples (subsamples). There were 5000 Monte Carlo replications in each experiment.

The results of the experiments are shown in Tables 1–3, which give the empirical coverage probabilities of the nominal 90% confidence intervals. In most respects, the Monte Carlo results are consistent with the theory reviewed in Sections 3–4. The ECP of the asymptotic normal approximation is large with the AR(1) and MA(1) processes, especially with $n = 100$. As expected, the sieve bootstrap has the smallest ECP with the AR(1) and MA(1) processes. However, the sieve bootstrap has a very large ECP when the data are generated by the ARCH(1) process, which does not have the autoregressive representation (4.2). The Markov bootstrap has a small ECP with the AR(1) model and with the MA(1) model, which is only an approximate Markov process. With the AR(1) and MA(1) models, the ECP of the Markov bootstrap tends to be larger than that of the sieve bootstrap. This is consistent with the theoretical results of Sections 4.1–4.2. The ECPs of the block bootstrap and subsampling are sensitive to the block length and size of the subsample. Depending on the DGP and the block length, the ECP of the block bootstrap can be either comparable to that of the Markov bootstrap or even larger than that of the asymptotic normal approximation. The ECP's of the subsampling procedure tend to be larger than those of both the Markov bootstrap and the block bootstrap with the "best" block length.

## 6 Discussion

First-order asymptotic approximations (the asymptotic normal approximation) often provide inaccurate estimates of distribution functions, coverage probabilities of confidence intervals, and rejection probabilities of hypothesis tests. This paper has explored the ability of the bootstrap to provide approximations that are more accurate when the data are a time series but one does not have a parametric model that reduces the DGP to independent random sampling.

Theoretical calculations and the results of Monte Carlo experiments show that there are significant benefits from taking advantage of any knowledge of the structure of the DGP. If the DGP is linear, then the sieve bootstrap appears to be the best bootstrap method. It has smaller ECPs and ERPs than do either the asymptotic normal approximation or other bootstrap methods. If an assumption of linearity is too strong but the DGP is or can be approximated by a low-order Markov process, then the Markov bootstrap provides good results.

When the DGP is neither linear nor a low-order Markov or approximate Markov process, then the block bootstrap or, possibly, subsampling with extrapolation are the only currently available resampling methods for reducing ECPs and ERPs. However, the accuracy of these methods is sensitive to the choice of block length or subsample size. With a poor choice, the ECPs and ERPs of the block bootstrap and subsampling can be larger than those of the asymptotic normal approximation. At present, there are no satisfactory methods for selecting block lengths for computing confidence intervals or critical values of tests in applications. Consequently, the problem of block-length selection is a severe impediment to use of the block bootstrap in applications, and it is an important topic for further research.

Another potentially important question whether it is possible to develop methods that are more accurate than the block bootstrap but impose less structure on the DGP than do the Markov bootstrap and the sieve bootstrap for linear processes. It would also be useful to know whether dimension- reduction methods can be used to mitigate the curse of dimensionality of the Markov bootstrap without imposing unacceptably strong restrictions on the DGP. In nonparametric estimation of conditional mean functions, for example, dimension reduction is achieved by single-index and additive models, among others. It is an open question whether useful dimension-reduction methods can be found for estimating Markov transition densities.

Finally, the existing theoretical explanation of the bootstrap's ability to provide asymptotic refinements is imperfect and, in time series settings, can be misleading. The existing theory is based on Edgeworth expansions. If the DGP is GSM, then the parameters $\kappa$ that enter the Edgeworth expansions (2.1) and (2.2) can be estimated analytically with errors that are only slightly larger than $O(n^{-1/2})$. By substituting these estimates into an analytic Edgeworth expansion, one can obtain theoretical ECP's and ERP's that are comparable to or smaller than those of the block, sieve and Markov bootstraps. In Monte Carlo experiments, however, analytic Edgeworth expansions are often much less accurate than the bootstrap. Thus, Edgeworth expansions are imperfect guides to the relative accuracy of alternative methods for achieving asymptotic refinements. This is a serious problem for research on bootstrap methods for time series because comparing the relative accuracies of alternative bootstrap approaches is an essential element of this research. Accordingly the development of a more complete theory of the performance of the bootstrap is a potentially important, though undoubtedly very difficult, area for future research.

## Acknowledgement

The research of Wolfgang Härdle and Jens-Peter Kreiss was supported in part by Deutsche Forschungsgemeinschaft Sonderforschungsbereich 373, "Quantifikation und Simulation Ökonomischer Prozesse". The research of Joel L. Horowitz was supported in part by NSF Grant SES-9910925 and by the Alexander von Humboldt Foundation.

**Table 1**

*Empirical coverage probabilities of nominal 90% confidence intervals for the AR(1) DGP.*

| Sample size / Method | $n = 100$ | $n = 500$ |
|---|---|---|
| Asymptotic Normal Approximation | 0.973 | 0.943 |
| Sieve Bootstrap | 0.893 | 0.889 |
| Block Bootstrap ($\ell$ Block Length) | 0.949 ($\ell = 10$) | 0.936 ($\ell = 10$) |
| | 0.905 ($\ell = 15$) | 0.916 ($\ell = 20$) |
| | 0.883 ($\ell = 20$) | 0.890 ($\ell = 30$) |
| | 0.863 ($\ell = 25$) | 0.883 ($\ell = 40$) |
| Markov Bootstrap | 0.882 | 0.891 |
| Subsampling ($m$ = Subsample Size) | 0.972 ($m = 25$) | 0.895 ($m = 125$) |
| | 0.956 ($m = 33$) | 0.869 ($m = 167$) |
| | 0.928 ($m = 50$) | 0.826 ($m = 250$) |

**Table 2**

*Empirical coverage probabilities of nominal 90% confidence intervals for the MA(1) model.*

| Sample size / Method | $n = 100$ | $n = 500$ |
|---|---|---|
| Asymptotic Normal Approximation | 0.927 | 0.911 |
| Sieve Bootstrap | 0.904 | 0.899 |
| Block Bootstrap ($\ell$ = Block Length) | 0.896 ($\ell = 5$) | 0.896 ($\ell = 10$) |
| | 0.863 ($\ell = 10$) | 0.883 ($\ell = 20$) |
| | 0.848 ($\ell = 15$) | 0.877 ($\ell = 30$) |
| | 0.828 ($\ell = 20$) | 0.869 ($\ell = 40$) |
| Markov Bootstrap | 0.903 | 0.932 |
| Subsampling ($m$ = Subsample Size) | 0.887 ($m = 25$) | 0.888 ($m = 125$) |
| | 0.870 ($m = 33$) | 0.854 ($m = 167$) |
| | 0.828 ($m = 50$) | 0.807 ($m = 250$) |

**Table 3**

*Empirical coverage probabilities of nominal 90% confidence intervals for the ARCH(1) DGP.*

| Sample size / Method | $n = 100$ | $n = 500$ |
|---|---|---|
| Asymptotic Normal Approximation | 0.916 | 0.909 |
| Sieve Bootstrap | 0.841 | 0.826 |
| Block Bootstrap ($\ell$ = Block Length) | 0.883 ($\ell = 2$) | 0.890 ($\ell = 5$) |
| | 0.857 ($\ell = 5$) | 0.884 ($\ell = 10$) |
| | 0.839 ($\ell = 10$) | 0.874 ($\ell = 20$) |
| | 0.825 ($\ell = 15$) | 0.866 ($\ell = 30$) |
| Markov Bootstrap | 0.854 | 0.868 |
| Subsampling ($m$ = Subsample Size) | 0.790 ($m = 25$) | 0.793 ($m = 125$) |
| | 0.819 ($m = 33$) | 0.824 ($m = 167$) |
| | 0.860 ($m = 50$) | 0.873 ($m = 250$) |

## References

Andrews, D.W.K. (1991). Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation. *Econometrica*, **59**, 817–858.

Andrews, D.W.K. (2001). Higher-order Improvements of the Parametric Bootstrap for Markov Processes, Cowles Foundation Discussion Paper No. 1334, Yale University, New Haven, CT.

Andrews, D.W.K. (2002a). Higher-Order Improvements of a Computationally Attractive $k$-Step Bootstrap for Extremum Estimators. *Econometrica*, **70**, 119–162.

Andrews, D.W.K. (2002b). The Block-Block Bootstrap: Improved Asymptotic Refinements, Cowles Foundation Discussion Paper No. 1370, Yale University, New Haven, CT.

Andrews, D.W.K. & Monahan, J.C. (1992). An Improved Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*, **59**, 817–858.

Beran, R. & Ducharme, G.R. (1991). *Asymptotic Theory for Bootstrap Methods in Statistics*. Les Publications CRM, Centre de recherches mathematiques, Universitė de Montréal, Montréal, Canada.

Bollerslev, T., Chou, R. & Kroner, K. (1992). ARCH Modelling in Finance: A Review of Theory and Empirical Evidence. *Journal of Econometrics*, **52**, 5–60.

Bose, A. (1988). Edgeworth Correction by Bootstrap in Autoregressions. *Annals of Statistics*, **16**, 1709–1722.

Bose, A. (1990). Bootstrap in Moving Average Models. *Annals of the Institute of Statistical Mathematics*, **42**, 753–768.

Bühlmann, P. (1997). Sieve Bootstrap for Time Series. *Bernoulli*, **3**, 123–148.

Bühlmann, P. (1998). Sieve Bootstrap for Smoothing Nonstationary Time Series. *Annals of Statistics*, **26**, 48–83.

Bühlmann, P. & Künsch, H.R. (1995). The Blockwise Bootstrap for General Parameters of a Stationary Time Series. *Scandinavian Journal of Statistics*, **22**, 35–54.

Carlstein, E. (1986). The Use of Subseries Methods for Estimating the Variance of a General Statistic from a Stationary Time Series. *Annals of Statistics*, **14**, 1171–1179.

Carlstein, E., Do, K.-A., Hall, P., Hesterberg, T. & Künsch, H.R. (1998). Matched-block Bootstrap for Dependent Data. *Bernoulli*, **4**, 305–328.

Chan, K.S. & Tong, H. (1998). A Note on Testing for Multi-Modality with Dependent Data. Working paper, Department of Statistics and Actuarial Science, University of Iowa, Iowa City, IA.

Choi, E. & Hall, P. (2000). Bootstrap Confidence Regions Computed from Autoregressions of Arbitrary Order. *Journal of the Royal Statistical Society, Series B*, **62**, 461–477.

Dahlhaus, R. & Janas, D. (1996). A Frequency Domain Bootstrap for Ratio Statistics in Time Series. *Annals of Statistics*, **24**, 1934–1963.

Datta, S. & McCormick, W.P. (1995). Some Continuous Edgeworth Expansions for Markov Chains with Applications to Bootstrap. *Journal of Multivariate Analysis*, **52**, 83–106.

Davidson, R. & MacKinnon, J.G. (1999). The Size Distortion of Bootstrap Tests. *Econometric Theory*, **15**, 361–376.

Davison, A.C. & Hall. P. (1993). On Studentizing and Blocking Methods for Implementing the Bootstrap with Dependent Data. *Australian Journal of Statistics*, **35**, 215–224.

Davison, A.C. & Hinkley, D.V. (1997). *Bootstrap Methods and Their Application*. Cambridge, UK: Cambridge University Press.

Diebolt, J. & Guegan, D. (1990). Probabilistic Properties of the General Nonlinear Markovian Process of Order One and Applications to Time Series Modelling. Rapport Technique de L.S.T.A. 125, Universitė Paris XIII.

Efron, B. & Tibshirani, R.J. (1993). *An Introduction to the Bootstrap*. New York: Chapman & Hall.

Engle, R.F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of U.K. Inflation. *Econometrica*, **50**, 987–1008.

Franke, J. & Härdle, W. (1992). On Bootstrapping Kernel Spectral Estimates. *Annals of Statistics*, **20**, 121–145.

Franke, J., Kreiss, J.-P. & Mammen, E. (2002). Bootstrap of Kernel Smoothing in Nonlinear Time Series. *Bernoulli*, **8**, 1–37.

Franke, J., Kreiss, J.-P., Mammen, E. &. Neumann, M.H. (2002). Properties of the Nonparametric Autoregressive Bootstrap. *Journal of Time Series Analysis*, **23**, 555–585.

Götze, F. & Hipp, C. (1983). Asymptotic Expansions for Sums of Weakly Dependent Random Vectors. *Zeitschrift fir Warscheinlichkeitstheorie und verwandte Gebiete*, **64**, 211–239.

Götze, F. & Hipp, C. (1994). Asymptotic Distribution of Statistics in Time Series. *Annals of Statistics*, **22**, 2062–2088.

Götze, F. & Künsch, H.R. (1996). Blockwise Bootstrap for Dependent Observations: Higher Order Approximations for Studentized Statistics. *Annals of Statistics*, **24**, 1914–1933.

Gouriéroux, C. (1997). *ARCH Models and Financial Applications*. New York: Springer-Verlag.

Hafner, C. (1996). *Nonlinear Time Series Analysis with Applications to Foreign Exchange Rate Volatility*. Heidelberg: Physica-Verlag.

Hall, P. (1985). Resampling a Coverage Process. *Stochastic Process Applications*, **19**, 259–269.

Hall, P. (1991). On the Distribution of Suprema. *Probability Theory and Related Fields*, **89**, 447–455.

Hall, P. (1992). *The Bootstrap and Edgeworth Expansion*. New York: Springer-Verlag.

Hall, P. & Horowitz, J.L. (1996). Bootstrap Critical Values for Tests Based on Generalized-Method-of-Moments Estimators. *Econometrica*, **64**, 891–916.

Hall, P., Horowitz, J.L.& Jing, B.-Y. (1995). On Blocking Rules for the Bootstrap with Dependent Data. *Biometrika*, **82**, 561–574.

Hall, P. & Jing, B.-Y. (1996). On Sample Reuse Methods for Dependent Data. *Journal of the Royal Statistical Society, Series B*, **58**, 727–737.

Hansen, B. (1999). Non-Parametric Dependent Data Bootstrap for Conditional Moment Models. Working paper, Department of Economics, University of Wisconsin, Madison, WI.

Hansen, L.P. (1982). Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica*, **50**, 1029–1054.

Härdle, W., Kleinow, T., Korostelev, A., Logeay, C. & Platen, E. (2001). Semiparametric Diffusion Estimation and Application to a Stock Market Index. Preprint No. 24, SFB 373, Humboldt-Universität Berlin.

Horowitz, J.L. (2001). The Bootstrap. In *Handbook of Econometrics*, Eds. J.J. Heckman and E.E. Leamer. Vol. **5** Amsterdam: North-Holland Publishing Co.

Horowitz, J.L. (2003). Bootstrap Methods for Markov Processes. *Econometrica*, forthcoming.

Inoue A. & Shintani, M. (2001). Bootstrapping GMM Estimators for Time Series. Working Paper, Department of Agricultural and Resource Economics, North Carolina State University, Raleigh, NC.

Kreiss, J.-P. (1988). Asymptotic Statistical Inference for a Class of Stochastic Processes. Habilitationsschrift, University of Hamburg, Germany.

Kreiss, J.-P. (1992). Bootstrap Procedures for AR($\infty$) Processes. In *Bootstrapping and Related Techniques*, Eds. K.H. Jöckel, G. Rothe and W. Sender, *Lecture Notes in Economics and Mathematical Systems*, **376**, pp. 107–113. Heidelberg: Springer-Verlag.

Kreiss, J.-P. (2000). Residual and Wild Bootstrap for Infinite Order Autoregressions. Working paper, Institute for Mathematical Stochastics, Technical University of Braunschweig, Germany.

Kreiss, J.-P., Neumann, M.H. & Yao, Q. (1998). Bootstrap Tests for Simple Structure in Nonparametric Time Series Regression. Working paper, Institute for Mathematical Stochastics, Technical University of Braunschweig, Germany.

Kreiss, J.-P. & Paparoditis, E. (2000). Autoregressive Aided Periodogram Bootstrap for Time Series. Working paper, Institute for Mathematical Stochastics, Technical University of Braunschweig, Germany.

Künsch, H.R. (1989). The Jackknife and the Bootstrap for General Stationary Observations. *Annals of Statistics*, **17**, 1217–1241.

Lahiri, S.N. (1991). Second Order Optimality of Stationary Bootstrap. *Statistics and Probability Letters*, **11**, 335–341.

Lahiri, S.N. (1992). Edgeworth Correction by 'Moving Block' Bootstrap for Stationary and Nonstationary Data. In *Exploring the Limits of Bootstrap*, Eds. R. LePage and L. Billard. New York: Wiley.

Lahiri, S.N. (1996a). On Edgeworth Expansion and Moving Block Bootstrap for Studentized $M$-estimators in Multiple Linear Regression Models. *Journal of Multivariate Analysis*, **56**, 42–59.

Lahiri, S.N. (1996b). Asymptotic Expansions for Sums of Random Vectors under Polynomial Mixing Rates. *Sankhya, Series A*, **58**, Pt. 2, 206–224.

Lahiri, S.N. (1999). Theoretical Comparisons of Block Bootstrap Methods. *Annals of Statistics*, **27**, 386–404.

Meyn, S.P. & Tweedie, R.L. (1993). *Markov Chains and Stochastic Stability*. New York: Springer-Verlag.

Neumann, M.H. & Kreiss, J.-P. (1998). Regression-type Inference in Nonparametric Autoregression. *Annals of Statistics*, **26**, 1570–1613.

Newey, W.K. & West, K.D. (1987). A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*, **55**, 703–708.

Newey, W.K. & West, K.D. (1994). Automatic Lag Selection in Covariance Matrix Estimation. *Review of Economic Studies*, **61**, 631–653.

Nordgaard, A. (1992). Resampling a Stochastic Process Using a Bootstrap Approach. In *Bootstrapping and Related Techniques*, Eds. K.H. Jöckel, G. Rothe and W. Sendler, *Lecture Notes in Economics and Mathematical Systems*, **376**. Berlin: Springer-Verlag.

Paparoditis, E. (1996). Bootstrapping Autoregresive and Moving Average Parameter Estimates of Infinite Order Vector Autoregressive Processes. *Journal of Multivariate Analysis*, **57**, 277–296.

Paparoditis, E. & Politis, D.N. (2001). Tapered Block Bootstrap. *Biometrika*, **88**, 1105–1119.

Paparoditis, E. & Politis, D.N. (2002). The Local Bootstrap for Markov Processes. *Journal of Statistical Planning and Inference*, **108**, 301–328.

Politis, D.N. & Romano, J.P. (1992). A General Resampling Scheme for Triangular Arrays of Alpha-mixing Random Variables with Application to the Problem of Spectral Density Estimation. *Annals of Statistics*, **20**, 1985–2007.

Politis, D.N. & Romano, J.P. (1993). The Stationary Bootstrap. *Journal of the American Statistical Association*, **89**, 1303–1313.

Politis, D.N. & Romano, J.P. (1994). Large Sample Confidence Regions Based on Subsamples under Minimal Assumptions. *Annals of Statistics*, **22**, 2031–2050.

Politis, D.N., Romano, J.P. & Wolf, M. (1999). *Subsampling*. New York: Springer-Verlag.

Rajarshi, M.B. (1990). Bootstrap in Markov-Sequences Based on Estimates of Transition Density. *Annals of the Institute of Statistical Mathematics*, **42**, 253–268.

Silverman, B.W. (1986). *Density Estimation for Statistics and Data Analysis*. Chapman & Hall.

Theiler, J., Paul, L.S. & Rubin, D.M. (1994). Detecting Nonlinearity in Data with Long Coherence Time. In *Time Series Prediction*, Eds. A. Weigend and N. Gershenfeld. Reading, MA: Addison-Welsey.

Zvingelis, J. (2001). On Bootstrap Coverage Probability with Dependent Data. In *Computer-Aided Econometrics*, Ed. D. Giles. Marcel Dekker.

## Résumé

Le bootstrap est une méthode pour estimer la distribution d'un estimateur en rééchantillonnant ses données ou un modèle estimé à partir des données. Les méthodes disponibles pour mettre en oeuvre le bootstrap et la précision des estimateurs de bootstrap dépendent de ce que les données proviennent d'un échantillon aléatoire indépendant ou d'une série temporelle. Cet article concerne l'application du bootstrap aux données des séries temporelles quand on ne dispose pas de modèle paramétrique de dimension finie qui réduise le processus de génération des données à l'échantillonnage aléatoire indépendant. Nous examinons les méthodes qui ont été proposées pour mettre en oeuvre le bootstrap dans cette situation et discutons la précision de ces méthodes comparativement à celle des approximations asymptotiques de premier ordre. Nous montrons que les méthodes pour mettre en oeuvre le bootstrap avec les données des séries temporelles ne sont pas aussi bien comprises que les méthodes pour les données des échantillons aléatoires indépendants. Bien que des méthodes de bootstrap prometteuses pour les séries temporelles soient disponibles, il y a un besoin considérable de recherche supplémentaire dans leur application. Nous décrivons quelques problèmes importants non résolus.

[*Received September 2001, accepted November 2002*]