



ELSEVIER

Journal of Econometrics 80 (1997) 297–318

---

---

**JOURNAL OF  
Econometrics**

---

---

## Bootstrapping cointegrating regressions<sup>1</sup>

Hongyi Li<sup>a</sup>, G.S. Maddala<sup>\*.b</sup>

<sup>a</sup>*Department of Decision Sciences and Managerial Economics, Faculty of Business Administration,  
Chinese University of Hong Kong, Hong Kong*

<sup>b</sup>*Department of Economics, The Ohio State University, Columbus, OH 43210, USA*

No receipt date

---

### Abstract

The paper investigates the usefulness of bootstrap methods for small sample inference in cointegrating regression models. It discusses the standard bootstrap, the recursive bootstrap, the moving block bootstrap and the stationary bootstrap methods. Some guidelines for bootstrap data generation and test statistics to consider are provided and some simulation evidence presented suggests that the bootstrap methods, when properly implemented, can provide significant improvement over asymptotic inference. © 1997 Elsevier Science S.A.

*Keywords:* Nonstationary time series; Small sample bias; Size distortion; Moving block bootstrap; Stationary bootstrap

*JEL classification:* C12; C13; C15.

---

### 1. Introduction

The first paper by Granger (1981) on cointegration was presented at a conference organized by one of the authors in Gainesville, Florida in 1980. For the next couple of years work on cointegration was mostly done by Granger and his

---

\* Corresponding author. Address: Room 410, 1945 N. High Street, Columbus, OH 43210-1172, USA.

<sup>1</sup> This is a substantially revised version of a paper presented at the EC<sup>2</sup> conference, Oxford, England, Dec. 1993. We would like to thank David Hinkley for helpful comments and the suggestion to use the stationary bootstrap. Helpful comments were also received from Noud P.A. van Giersbergen, David Hendry, Brett Inder, and three anonymous referees. In particular, the comments of the referees led to a complete rewriting of the paper. The usual disclaimer applies.

associates. During the last decade cointegration has become an extremely active area of research with participants all over the world.

The first procedure for the estimation of cointegrating regressions was the one proposed by Engle and Granger (1987) which depends on ordinary least squares (OLS). Since then several other procedures have been developed including the maximum likelihood method of Johansen (1988, 1991) and the fully modified OLS (FMOLS) of Phillips and Hansen (1990). The paper by Hargreaves (1994) lists eleven categories of procedures and now there are several more. All these estimation methods have been justified on asymptotic grounds. There are now many time series, particularly in finance (stock and bond markets) as well as foreign exchange rates, where data are available by the day (or indeed hour and minute) and thus, the sample sizes are indeed very large. In these cases one might be comfortable using asymptotic theory although such finely tuned data present further problems of very low signal/noise ratios and structural changes over the period analyzed. Many other studies, however, are based on small sample sizes (typically less than 100). In this paper we are concerned with small sample inference and interested in investigating the usefulness of bootstrap methods for this purpose.

The bootstrap method initiated by Efron (1979) is a resampling technique which takes advantage of today's powerful computers. The bootstrap resampling method was originally designed for independent and identically distributed (IID) errors. As pointed out by Hinkley (1988, p. 335), Jeong and Maddala (1993) among others, in most of the econometric applications, the error structures are generally more complicated than simple IID distributions and hence more sophisticated bootstrap techniques are needed.

Correlated errors are not exchangeable, and lagged dependent variables create problems in pseudo data generation. Unit root and cointegration regression models create further complications. In subsequent sections we will elaborate on these issues.

Given the very large number of estimation methods available, researchers trying to estimate cointegrating regressions are faced with two problems related to

- (i) choice of an estimation method and
- (ii) once this choice has been made, the appropriate procedure for conducting inference.

On problem (i) there are now several Monte Carlo studies that compare the different methods of estimation for cointegrated systems. See for instance Banerjee et al. (1986) for an earlier study and Phillips and Loretan (1991), Inder (1993), Gonzalo (1994), Hargreaves (1994) and Li and Maddala (1995) for some recent studies. Based on these studies no one particular method has been found to be superior to others. However, the studies are not all comparable to each other because they are based on different data generation processes (DGP's) and wide difference in signal/noise ratios. The focus of this paper is not on the choice

of an estimation method but on problem (ii), that is, the appropriate inferential procedures, once an estimation method has been chosen.

Turning to question (ii), the standard procedure of inference is to rely on asymptotic theory for tests of significance. It is now widely recognized that inference based on asymptotic distributions has two major drawbacks:

(a) The estimators, though consistent (and often super consistent) have substantial small sample biases and

(b) The tests of significance based on the asymptotic distributions have substantial size distortions.

There is, however, little work on small sample corrections to these two problems. The purpose of this paper is to demonstrate the usefulness of bootstrap methods in providing some corrections to small sample biases in the estimators and the size distortions in the associated tests of significance.

## 2. The different bootstrap methods considered

We will first describe the standard bootstrap, the recursive bootstrap, the moving block bootstrap (MBB), and the stationary bootstrap (SB) methods.

Let  $(y_1, y_2, \dots, y_n)$  be a random sample from a distribution characterized by a parameter  $\theta$ . Inference about  $\theta$  will be based on a statistic  $T$ . The basic bootstrap approach consists of drawing repeated samples (with replacement of size  $n'$ , which may or may not be equal to  $n$ , although it usually is). Call this sample  $(y_1^*, y_2^*, \dots, y_{n'}^*)$ . This is the bootstrap sample. We do this NB times and for each bootstrap sample we compute the statistic  $T$ . Call this  $T^*$ . The distribution of  $T^*$  based on the NB bootstrap samples is known as the bootstrap distribution of  $T$ . We use this to make inferences about  $\theta$ . This procedure has been extended to classical regressions by Freedman (1981a, b). In the case of the classical regression models, it is the residuals that are resampled. Needless to say when the errors are not IID, one needs to modify this procedure. There are some applications of this standard procedure to cointegrating regressions. See, for instance, Shea (1989a, b). However, these papers are based on the assumption of IID errors which is too restrictive and many of the estimation methods suggested for the cointegrating regressions have been designed to solve the problem of endogeneity of the regressors and serial correlation in the errors. In subsequent sections, we discuss the bootstrap methods for these estimators.

### 2.1. The recursive bootstrap

To deal with lagged dependent variables and serially correlated errors with a well specified structure (say stationary ARMA( $p, q$ ) models with known  $p$  and  $q$ ), one can use the *recursive bootstrap* method, first introduced by Freedman and

Peters (1984). This method was also used by Efron and Tibshirani (1986) for bootstrapping the AR(1) and AR(2) models. In the recursive bootstrap method one estimates the model by OLS, or some other consistent methods, obtains the residuals and (after rescaling and centering) resamples them. With the resampled residuals, one next generates the bootstrap samples recursively. In the case of a regression model with say AR(1) errors, such as

$$y_t = \beta x_t + u_t, \quad (1)$$

$$u_t = \rho u_{t-1} + e_t, \quad (2)$$

where  $e_t \sim \text{IID}(0, \sigma^2)$ , one estimates Eq. (1) by OLS, then using the estimated residuals  $\hat{u}_t$ , one estimates  $\hat{\rho}$  using the Cochrane–Orcutt or Prais–Winsten procedures and obtains  $\hat{e}_t$ . Then one resamples  $\hat{e}_t$  and using a recursive procedure generates  $\hat{u}_t$ , and the bootstrap sample on  $y_t$ . Rayner (1991) uses this procedure with some promising results in correcting the test size. We will examine its performance in a cointegrating regression with exogenous regressors in a subsequent section.

## 2.2. The moving block bootstrap

Application of the recursive bootstrap methods is straightforward if the error distribution is specified to be a stationary ARMA( $p, q$ ) process with known  $p$  and  $q$ . However, if the structure of serial correlation is not tractable or is misspecified, the residual based methods will give inconsistent estimates (if lagged dependent variables are present in the system). Other approaches which do not require fitting the data into a parametric form have been developed to deal with general dependent time series data. Carlstein (1986) first discussed the idea of bootstrapping blocks of observations rather than the individual observations. The blocks he considers are non-overlapping. Later, Künsch (1989) and Liu and Singh (1992) (the paper was available as a discussion paper in 1988) independently introduced a more general bootstrap procedure, the moving block bootstrap which is applicable to stationary time series data. In this method the blocks of observations are overlapping.

The methods of Carlstein (non-overlapping blocks) and Künsch (overlapping blocks) both divide the data of  $n$  observations into blocks of length  $l$  and select  $b$  of these blocks (with repeats allowed) by resampling with replacement all the possible blocks. Let us for simplicity assume  $n = bl$ . In the Carlstein procedure, there are just  $b$  blocks. In the Künsch procedure there are  $n - l + 1$  blocks. The blocks are  $L_k = \{x_k, x_{k+1}, \dots, x_{k+l-1}\}$  for  $k = 1, 2, \dots, (n - l + 1)$ . For example with  $n = 6$  and  $l = 3$  suppose the data are:  $x_t = \{3, 6, 7, 2, 1, 5\}$ . The blocks according to Carlstein are  $\{(3, 6, 7), (2, 1, 5)\}$ . The blocks according to Künsch are  $\{(3, 6, 7), (6, 7, 2), (7, 2, 1), (2, 1, 5)\}$ . Now draw a sample of two blocks with replacement in each case. Suppose, the first draw gave (3, 6, 7). The probability

of missing all of (2, 1, 5) is  $\frac{1}{2}$  in Carlstein's scheme and  $\frac{1}{4}$  in the moving block scheme. Thus there is a higher probability of missing entire blocks in the Carlstein scheme. For this reason, it is not popular, and is not often used. Our own experience with Carlstein's non-overlapping block method is that it gave very erratic results as the block length was varied. The MBB did better.

The literature on blocking methods is mostly on the estimation of the sample mean and its variance, although Liu and Singh (1992) talk about the applicability of the results to more general statistics, and Künsch (1989, p. 1235) discusses the AR(1) and MA(1) models. In all these studies the bootstrapping is done by sampling blocks of the data.

### 2.3. The stationary bootstrap

The pseudo time series generated by the moving block method is not stationary, even if the original series  $\{x_t\}$  is stationary. For this reason, Politis and Romano (1994) suggest the *stationary bootstrap* method. The basic steps for the stationary bootstrap are the same as those of the moving block bootstrap. However, there is one major difference between the sampling schemes of the moving block bootstrap and the stationary bootstrap. The stationary bootstrap resamples the data blocks of random length, where the length of each block has a geometric distribution, while the moving block bootstrap resamples blocks of data of the same length. The resampling method for the stationary bootstrap runs as follows.

Let  $\{u_1^*\}$  be the first observation randomly resampled from  $\{\hat{u}_t\}$  ( $t = 1, \dots, n$ ) so that  $\{u_1^* = \hat{u}_j\}$  for  $1 \leq j \leq n$ . The second one  $\{u_2^*\}$  has the following distribution

$$\Pr(\{u_2^* = \hat{u}_{j+1}\}) = 1 - p \quad \text{and} \quad \Pr(\{u_2^* = \hat{u}_i\}) = p,$$

where  $1 \leq i \leq n$  and  $p$  ( $0 \leq p \leq 1$ ) is the probability of the geometric distribution of the random block length. The average block length in the stationary bootstrap is  $1/p$ . Thus, the block length  $k$  in the moving block bootstrap and the probability  $p$  in the stationary bootstrap play the same role when  $k = 1/p$ .

There is some discussion of optimal choice of  $k$  and  $p$  in the papers by Carlstein (1986), Künsch (1989), Hall and Horowitz (1993) and Politis and Romano (1995). The rules are suggestive in small sample cases.

### 3. Issues in the generation of bootstrap samples and the test statistics

There are three important issues that need to be resolved in bootstrapping cointegrating regression models. These are

- (1) Whether to bootstrap the residuals or the data.
- (2) If it is the residuals, how should the residuals be generated?
- (3) How should the appropriate test statistics be defined?

Regarding question (1), although bootstrapping the residuals is a common procedure, there have been some examples in the literature where bootstrapping the data has been suggested. This alternative, however, is not a valid one in the case of cointegrating regressions.<sup>2</sup>

For the case of random regressors (which he calls the ‘correlation model’ as opposed to the ‘regression model’), Freedman (1981a) suggests resampling the pair  $(y, x)$  which have a joint distribution with  $E(y|x) = x\beta$ . Efron (1981) uses the direct method of resampling the data in a problem involving censored data. He bootstraps the data  $(x_i, d_i)$  where  $d_i = 1$  if  $x_i$  is not censored and  $d_i = 0$  if  $x_i$  is censored. The direct method of bootstrapping the data has also been advocated in Efron and Gong (1983).

The direct method of sampling the data has not gone unnoticed in econometrics. An unpublished paper by Kiviet (1984) is the earliest use. Veall (1987, p. 205) considers the direct bootstrap approach but rejects it on grounds that it does not embody all the information used in the residual based approach.

Our basic argument is that information about the structure of the model should be used in the generation of bootstrap samples because this information is used in the estimation of the model from the bootstrap data. This is done in the residual based bootstrap generation but not in the method of bootstrapping the data. In the case of cointegrating regressions, to see what is involved in the direct vs. residuals based bootstrap methods, consider the case where  $y_t$  and  $x_t$  are both  $I(1)$  and we have the regression equation

$$y_t = \beta x_t + u_t, \quad (3)$$

$$x_t = x_{t-1} + v_t. \quad (4)$$

Suppose we bootstrap the data  $(y_t, x_t)$  and estimate equation (3) by OLS. If  $u_t$  is also  $I(1)$  then it is well known that equation (3) is a spurious regression. But there is no way of knowing this if we use the direct bootstrap method without first testing whether (3) is indeed a cointegration relationship.

Suppose we initially apply cointegration tests to equation (3) and make sure that it is not a spurious regression. This implies that  $y_t$  and  $x_t$  are  $I(1)$  and  $u_t$  is  $I(0)$ . But the direct bootstrap method does not use the information that  $u_t$  is  $I(0)$ , and that  $x_t$  is  $I(1)$ .

There is also another important issue that the bootstrap samples of  $y_t$  and  $x_t$  may not be  $I(1)$  by the direct method because the time series structure of an  $I(1)$  process is destroyed. If the direct method for IID distribution is used, then

---

<sup>2</sup> In the earlier version of this paper we used both the procedures of bootstrapping the residuals and bootstrapping the data and found that the latter performed slightly worse. The comments by Hinkley and one another referee convinced us to delete all those results, and add additional reasons why bootstrapping the data is not a valid procedure in the case of cointegrating regressions.

bootstrap samples of  $y_t$  and  $x_t$  are not  $I(1)$  at all. If the block methods for correlated data are used, the bootstrap samples are not  $I(1)$  either. In fact, the MBB assumes that  $\sup_{t>1} E\|x_t\|^4 < \infty$  which is not certainly satisfied if  $x_t$  is  $I(1)$ . Also, the assumptions of Künsch (1989) are violated if  $x_t$  is  $I(1)$ . Thus, the direct method using MBB is not valid. The direct method can not give bootstrap samples that are also  $I(1)$  as the original data. This issue is even more important than whether the information in  $u_t$  is used or not. If the bootstrap samples are not  $I(1)$ , then the limiting distribution of the parameter estimates are entirely different. Hence, in this case there is no validity whatsoever for the direct method of bootstrapping the data. If the residual method of bootstrapping is used, the bootstrap data  $y_t^*$  and  $x_t^*$  are generated according to equations (3) and (4), which guarantees that  $y_t^*$  and  $x_t^*$  are also  $I(1)$ .

There are essentially two pieces of information in this model, namely, (a)  $x_t$  is  $I(1)$  and (b) equation (3) is a cointegrating regression. This suggests the following residual based bootstrap method which is used in this paper.

(i) Estimate equation (3) by OLS to get  $\hat{\beta}$  and  $\hat{u}_t$ . Note that  $\hat{\beta}$  is super consistent and hence it is valid to bootstrap  $\hat{u}_t$ .

(ii) Next define  $\hat{v}_t = \Delta x_t$  and bootstrap the pairs  $(\hat{u}_t, \hat{v}_t)$ . If the errors are serially correlated, the recursive or block bootstrap methods can be used. This uses the information in (a) and (b) above and also preserves the correlation between  $u_t$  and  $v_t$ , that is, the endogeneity of  $x_t$  is taken into account. Note also that bootstrapping just  $\hat{u}_t$  is not a valid procedure. By contrast the direct method of bootstrapping the data does not use information (a) and (b). The important thing to note is that  $x_t$  is not only stochastic as in the case of Freedman (1981a), but it is also  $I(1)$ .

Turning to the next issue (2), how should the bootstrap residuals be generated, we answered one of the problems. The next thing is, suppose we are testing the hypothesis  $\beta = \beta_0$ , then instead of using  $\hat{u}_t = y_t - \hat{\beta}x_t$  we can use  $\tilde{u}_t = y_t - \beta_0 x_t$ , and bootstrap the pairs  $(\tilde{u}_t, \hat{v}_t)$ .

The issue of which of these two residuals to use for bootstrapping hinges on the purpose of the bootstrap method: whether it is hypothesis testing or estimation (and bias correction). When a null hypothesis is tested, the resampling is performed under the null and hence it is  $\tilde{u}_t$  that we consider for bootstrapping. For estimation and bias correction it makes more sense to bootstrap  $\hat{u}_t$  and not  $\tilde{u}_t$ . We discuss the problems related to bootstrapping for hypothesis testing when we come to issue (3). There it will be noted that many test procedures suggested have been based on bootstrapping  $\hat{u}_t$  and not  $\tilde{u}_t$ . In our simulations discussed later, we bootstrapped  $\tilde{u}_t$  in the case of hypothesis testing. Since simulation with bootstrap procedure is very time consuming we used the same resampled data for bias correction as well. As explained later, the bias correction formulae will be different whether  $\hat{u}_t$  or  $\tilde{u}_t$  are used for bootstrapping.

Coming to the estimation problem, we said earlier that estimators like FMOLS, MLVECM and so on have been shown to have substantial biases in small samples. Our purpose is to see how the bootstrap methods help in bias correction. If we denote  $\beta_i^*$  as the estimator of  $\beta$  from the  $i$ -th bootstrap sample and define  $\bar{\beta}^* = (NB)^{-1} \sum \beta_i^*$ , an estimate of the bias will be given by  $\bar{\beta}^* - \hat{\beta}$  and hence the bias corrected estimator of  $\beta$  is

$$\tilde{\beta}_{bc} = \hat{\beta} + (\beta_0 - \bar{\beta}^*)$$

if we use  $\hat{u}_t$  for bootstrapping. If we use the residuals  $\tilde{u}_t$  for bootstrapping, the bias corrected estimator of  $\beta$  is

$$\tilde{\beta}_{bc} = \hat{\beta} + (\beta_0 - \bar{\beta}^*).$$

In the case of cointegrating regressions, if there are no endogeneity and serial correlation problems, the OLS estimator of the cointegrating vector is asymptotically normally distributed, and we can use the above formulae for bias correction. In the case where either or both of these problems are present the OLS estimator is not normally distributed (even asymptotically if endogeneity exists) and it involves nuisance parameters arising from endogeneity and serial correlation. But we can use the FMOLS or MLVECM estimators for bias correction. To illustrate the use of bootstrap methods for bias correction in the case of the OLS estimator, we have to use a model with no endogeneity and no serial correlation. But this restrictive model is not very interesting. Hence we illustrate bias correction with reference to the FMOLS and MLVECM methods.

Finally issue (3) is on how to define the test statistics. We will first discuss this issue in relation to regression models and next come to special problems with cointegrating regressions.

First there is the issue of whether to bootstrap  $\hat{\beta}$  or the pivotal statistic  $\hat{\beta}/SE(\hat{\beta})$ . The importance of bootstrapping the pivotal statistics has been stressed by Hartigan (1986), Hall (1988, 1992) and Beran (1987, 1988). In the case of confidence intervals, it has been established in these papers that the coverage probabilities are  $O(n^{-1})$  for the percentile- $t$  (or bootstrap- $t$ ) methods but only  $O(n^{-1/2})$  for the percentile methods. Given the duality between confidence intervals and tests of hypotheses, these results also carry over to hypothesis testing. This is also confirmed, for instance by the conflicting results in Veall (1986) who found that the bootstrap method did not provide an improvement over asymptotic theory and Rayner (1991) who comes to the opposite conclusion. The former bootstraps the coefficient  $\hat{\beta}$  and the latter bootstraps  $\hat{\beta}/SE(\hat{\beta})$  which is pivotal. However, in our simulations we did not get any better results by using the pivotal statistic for the Johansen procedure. This might be due to the fact that the variance of the variance estimator for the Johansen ML estimator is high as compared to that of the other asymptotically efficient estimation methods.



In addition to this issue of the importance of bootstrapping pivotal statistics, there is another issue that has been brought up by Hall and Wilson (1991). Suppose that we want to test the hypothesis  $H_0: \theta = \theta_0$  vs.  $H_1: \theta \neq \theta_0$ . Given an estimator  $\hat{\theta}$  of  $\theta$ , the usual test procedure would be based on  $T = \hat{\theta} - \theta_0$  and the significance level and  $p$ -values are obtained from the distribution of  $T$  under  $H_0$ . A direct application of the bootstrap procedure would suggest using the bootstrap distribution of  $T^* = \hat{\theta}^* - \theta_0$  instead of the distribution of  $T$ . ( $\hat{\theta}^*$  is the value of  $\theta$  from the bootstrap sample.) However, Hall (1992, Section 3.12) discusses the bad behavior of the power of this test arguing that  $T^*$  does not approximate the null hypothesis when the sample comes from a distribution with parameter  $\theta$  far away from  $\theta_0$ . Hall and Wilson therefore consider another bootstrap procedure based on the empirical distribution of  $(\hat{\theta}^* - \hat{\theta})$ . They compare this with the test procedure based on  $T$  and  $T^*$ .

Hall and Wilson propose two guidelines for hypothesis testing. The first suggests using the bootstrap distribution of  $(\hat{\theta}^* - \hat{\theta})$  but not  $(\hat{\theta}^* - \theta_0)$ . The second guideline suggests using a properly studentized statistic, that is  $(\hat{\theta}^* - \hat{\theta})/\hat{\sigma}^*$  and not  $(\hat{\theta}^* - \hat{\theta})/\hat{\sigma}$  or  $(\hat{\theta}^* - \hat{\theta})/\hat{\sigma}$ , where  $\hat{\sigma}^*$  is the estimate of  $\hat{\sigma}$  from the bootstrap sample.

van Giersbergen and Kiviet (1993) discuss these two rules in the context of hypothesis testing in regression models. To simplify the exposition, we will discuss the case of a simple regression

$$y = \beta x + \varepsilon, \quad \varepsilon \sim \text{iid}(0, \sigma^2)$$

Let  $\hat{\beta}$  and  $\hat{\sigma}$  be the OLS estimators of  $\beta$  and  $\sigma$  respectively and  $\hat{\varepsilon}$  the OLS residuals. Let  $\varepsilon^*$  be the bootstrap residuals obtained by resampling  $\hat{\varepsilon}$ . The null hypothesis to be tested is  $H_0: \beta \sim \beta_0$  vs.  $H_1: \beta \neq \beta_0$ .

Consider two sampling schemes for the generation of the bootstrap samples:

$$S_1: y^* = \hat{\beta}x + \varepsilon^*, \quad S_2: y^* = \beta_0x + \varepsilon^*.$$

Both use  $\varepsilon^*$  but they differ in the way  $y^*$  is generated. In  $S_1$ ,  $\hat{\beta}$  is used, while in  $S_2$ ,  $\beta_0$  (the value specified in the null) is used. For each sampling scheme consider two  $t$ -statistics:

$$T_1: T(\hat{\beta}) = (\hat{\beta}^* - \hat{\beta})/\hat{\sigma}^*, \quad T_2: T(\beta_0) = (\hat{\beta}^* - \beta_0)/\hat{\sigma}^*.$$

Thus four different combinations of the sampling schemes and  $t$ -statistics can be defined. Hall and Wilson consider sampling scheme  $S_1$  only and suggest using  $T_1$  only. They do not consider sampling scheme  $S_2$  which is quite natural for the statistic  $T_2$  in the case of a regression model. van Giersbergen and Kiviet suggest, on the basis of a Monte Carlo study of an AR(1) model, the use of  $T_2$  under sampling scheme  $S_2$  in preference to  $T_1$  under  $S_1$ . The main

conclusions of the paper are:

(i) Inference based on  $T_2$  under  $S_1$  does not just have low power but in fact has size close to zero. Similarly  $T_1$  under  $S_2$  is not recommended. The sampling scheme should mimic the null distribution of the test statistic to be bootstrapped.

(ii) Using  $T_1$  under  $S_1$  and  $T_2$  under  $S_2$  are the same. The two  $t$ -statistics are independent of either  $\hat{\beta}$  or  $\beta_0$ . This equivalence extends to the multiparameter case if one bootstraps the appropriate  $F$ -statistic. However, in dynamic models this equivalence breaks down in finite samples. The Monte Carlo results suggest that it is better to use  $T_2$  under  $S_2$ .

(iii) The limiting distributions of  $T_1$  under  $S_1$  and  $T_2$  under  $S_2$  are identical even with dynamic models. The conclusions that  $T_2$  under  $S_2$  is better is based on small sample performance.

There is yet another sampling scheme that is more reasonable than  $S_1$  or  $S_2$

$$S_3: y^* = \beta_0 x + \varepsilon_0^*,$$

where  $\varepsilon_0^*$  is the residual resampled from  $\tilde{\varepsilon}_0 = y - \beta_0 x$ . Note that in both  $S_1$  and  $S_2$  we use sampling based on the OLS residuals  $\hat{\varepsilon}$ . If the null  $H_0: \beta = \beta_0$  is true but the OLS estimator  $\hat{\beta}$  gives a value of  $\beta$  far from  $\beta_0$ , the empirical distribution of the residuals will suffer from a poor approximation of the distribution of the errors under the null.

The intuition behind  $S_3$  is as follows. If the null hypothesis is true, the distribution of  $\tilde{\varepsilon}_0$  is exactly the true distribution of the regression errors. Hypothesis testing based on this will give (approximately) the correct test size. If the null is not true, then the distribution of  $\tilde{\varepsilon}_0$  is different from the true distribution of the errors. Hypothesis testing will give proper test size and power. Thus, using  $T_2$  under  $S_3$  is better than using  $T_1$  under  $S_1$  or  $T_2$  under  $S_2$  which are equivalent as shown in Giersbergen and Kiviet. In this paper we use these restricted regression errors for sampling. This idea has also been used in Nankervis and Savin (1994).

In the case of bootstrapping unit root models, some of the above sampling schemes have been shown to be invalid. For instance, Basawa et al. (1991a) prove that sampling scheme  $S_1$  is not appropriate in the unit root case. Basawa et al. (1991b) use test statistic  $n(\hat{\beta}^* - 1)$  with sampling scheme  $S_3$ . Ferretti and Romo (1994) establish the result that test statistic  $n(\hat{\beta}^* - 1)$  with sampling scheme  $S_2$  can also be used in bootstrap tests of unit roots. Note that this is similar to the procedure used by Rayner (1990) for the stationary AR(1) model, except that it is not in the pivotal form. We have compared both the Basawa et al. (1991b) and Ferretti and Romo schemes of getting bootstrap samples and did not notice any difference, although more detailed investigation of this is under way.

One other issue relates to the type of statistics to use for bootstrapping when procedures like the moving block bootstrap are used. Davison and Hall (1993) argue that this creates problems in using the percentile-*t* method with the moving block bootstrap. They suggest that the usual estimator  $\hat{\sigma}^2 = n^{-1} \sum_{i=1}^n (x_i - \bar{x}_n)^2$  be modified to  $\tilde{\sigma}^2 = n^{-1} \sum_{i=1}^n \{ (x_i - \bar{x}_n)^2 + \sum_{k=1}^{i-1} \sum_{i=k+1}^n (x_i - \bar{x}_n)(x_{i+k} - \bar{x}_n) \}$ . With this modification the bootstrap-*t* can improve substantially on the normal approximation. The reason for this bias in the estimator of the variance is that the block bootstrap method damages the dependence structure of the data. Unfortunately this formula is valid only for the variance of  $\sqrt{n}\bar{x}_n$ . For more complicated problems there is no such simple correction available. Hence, in this paper no such corrections were applied and we bootstrapped the usual *t*-statistics in the pivotal form. However, the Monte Carlo studies showed that the bootstrap-*t* provided considerable improvement over asymptotic results.

In a subsequent paper, Hall and Horowitz (1994) investigate this problem in the context of tests based on GMM estimators. They argue that because the blocking methods do not replicate the dependence structure of the original data, it is necessary to develop special bootstrap versions of the test statistics and these must have the same distribution as the sample version of the test statistics through  $O_p(n^{-1})$ . They derive the bootstrap versions of the test statistics with Carlstein's blocking scheme (non-overlapping blocks) but argue that Künsch's blocking scheme is more difficult to analyze owing to its use of overlapping blocks.

In the case of hypothesis tests in cointegrating regressions based on the moving block scheme considered in this paper, the derivation of the appropriate bootstrap versions of the test statistics is still more complicated. Although the use of the bootstrap version of the usual test statistics cannot be theoretically justified, the Monte Carlo results unequivocally indicate considerable improvement over the asymptotic results. Thus, in spite of no explicit theoretical justification, we bootstrapped the usual test statistics and conducted the simulations. What our results suggest is that this procedure, nevertheless, is useful for empirical practitioners in this area.

#### 4. Estimation methods used and the DGP investigated

The cointegrating regression model under investigation is

$$y_t = \beta' x_t + u_t, \tag{5}$$

$$x_t = x_{t-1} + v_t \tag{6}$$

for  $t = 1, \dots, n$  where  $y_t \sim I(1)$ , a one-dimensional vector;  $x_t \sim I(1)$ , an  $m$ -dimensional vector. This is a typical triangular cointegrating system with  $m$  regressors.

Define the innovation vector as  $\eta_t = (u_t v_t)'$  and assume  $\eta_t \sim I(0)$ . Thus  $y_t$  and  $x_t$  are cointegrated in the sense of Engle and Granger (1987) and  $\beta$  is the  $m \times 1$  cointegrating vector. Various estimators have been developed for this model.

The Engle–Granger procedure is to estimate this model by OLS. It is now well known that even if  $\hat{\beta}_{OLS}$  is super consistent, its asymptotic distribution involves nuisance parameters arising from:

- (i) endogeneity of the regressors (correlation between  $u_t$  and  $v_t$ ) and
- (ii) serial correlation in the errors.

The several methods that have been suggested for the estimation of cointegrated systems are designed to solve these two problems.

For the sake of brevity, we will just consider only two of these estimation methods: the fully modified OLS (FMOLS) of Phillips and Hansen (1990) and vector error correction model (VECM) estimated by maximum likelihood (ML) by Johansen (1988, 1991). The former procedure is designed to remove the nuisance parameters asymptotically and the latter procedure solves the endogeneity and serial correlation problems by considering a MLVECM. We have also done the simulations with the canonical cointegrating regression (CCR) model of Park (1992) but since this method is very similar to FMOLS we will omit the results for the sake of brevity. The simulations for several other estimators are in a preliminary stage but for illustrative purposes the results on FMOLS and MLVECM are sufficient. Also, since both of these methods have been discussed in the literature, to conserve space we will omit the detailed presentation of these two methods.

*The Data generating process (DGP):* The data generating process (DGP) for the cointegration system (3) and (4) with  $m = 1$  (for the ease of computation, only the case with one regressor is considered) is that  $\eta_t = (u_t v_t)'$  follows a stationary VAR(1) process

$$\begin{pmatrix} u_t \\ v_t \end{pmatrix} = \Phi \begin{pmatrix} u_{t-1} \\ v_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_t \\ \omega_t \end{pmatrix},$$

where

$$\Phi = \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix} \text{ and } \begin{pmatrix} \varepsilon_t \\ \omega_t \end{pmatrix} \sim \text{IIDN}(0, E) \equiv \text{IIDN}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_\varepsilon^2 & \sigma_{\varepsilon\omega} \\ \sigma_{\varepsilon\omega} & \sigma_\omega^2 \end{pmatrix}\right).$$

It is assumed that  $E > 0$ , a positive definite matrix. In the simulations, we set

$$E = \begin{pmatrix} 1 & \lambda \\ \lambda & 1 \end{pmatrix} \text{ and } \Phi = \begin{pmatrix} \rho & 0 \\ 0 & \theta \end{pmatrix},$$

with  $\rho$ ,  $\theta$  and  $\lambda$  as free parameters. In all cases, the true cointegrating parameter is chosen as  $\beta_0 = 2$  and the sample size is  $n = 50$ . For each parameter combination, a total number of MC samples are simulated. For each of the MC samples, the number of bootstrap replication is NB.

## 5. Bootstrapping a cointegrating regression model with exogenous regressors

We will now consider a case with exogenous regressors but serial correlation in both  $u_t$  and  $v_t$ . We assume  $\lambda = 0$ ,  $\rho = \{0.0, 0.3, 0.6, 0.9, 0.95\}$ , and  $\theta = \{0.0, 0.3, 0.6, 0.9, 0.95\}$ . We consider hypothesis test of  $H: \beta = \beta_0$  with a nominal size of 5%.

In the case of classical regression models when the errors follow an AR(1) structure, it was noted that standard estimates of the AR(1) coefficient would be biased towards zero in finite samples and hence transformations based on these estimates would only partially correct for the problem. See King and Giles (1984) and the references therein. As a consequence, if the explanatory variables are strongly trended and the degree of serial correlation is large, the generalized least squares (GLS) standard errors for the coefficient estimates will be substantially underestimated and hypothesis tests on the coefficients lead to over-rejection of the null. To correct for these size distortion, bootstrap procedures have been used, for instance, by Veall (1986) and Rayner (1991).

We investigate this problem in the case of cointegrating regressions, with exogenous regressors and AR(1) errors. In this case, as shown in Phillips and Park (1988), the OLS and GLS estimators are asymptotically equivalent. The conventional  $t$ -statistic modified, if multiplied by square root of  $(1 - \rho)/(1 + \rho)$ , has a limiting normal distribution.

The simulation results presented in Table 1 show that using the asymptotic  $\chi^2$ -test, both the OLS and GLS estimators tend to over reject the null of  $\beta = \beta_0$  when the AR(1) coefficient  $\rho$  gets higher. The GLS estimates of  $\hat{\beta}$  are downward biased as in the conventional regressions with AR(1) errors. See Table 2 for related results. In fact, the size distortions in the case of GLS are even worse than those with the OLS. When  $\rho$  is close to zero, the nominal and the empirical sizes are approximately the same.

### 5.1. The recursive bootstrap

Since we have autocorrelated errors, we might consider the recursive bootstrap method for this model. As discussed in Section 2, Rayner (1991) used this method with some success in correcting the distortions in the test size in the case of a regression model with AR(1) errors. He got better results with the bootstrap using percentile- $t$  whereas Veall (1986) obtained no improvement over asymptotic theory because he used the percentile method.

In the case of cointegrating regressions with exogenous regressors, we used the recursive bootstrap method to get empirical critical values, which are compared with the asymptotic critical values. The bootstrap test sizes were obtained by using two different bootstrap methods: the percentile method and the percentile- $t$  method. Not surprisingly, the results from the percentile method were not much better than those from the asymptotic theory and hence for the

Table 1

Empirical test size for OLS and GLS based on the asymptotic procedures ( $H_0: \beta = \beta_0$  nominal size: 5%)

OLS		$\rho$				
		0.0	0.3	0.6	0.9	0.95
$\theta$	0.0	0.060	0.060	0.078	0.096	0.168
	0.3	0.062	0.072	0.080	0.122	0.162
	0.6	0.066	0.072	0.080	0.126	0.170
	0.9	0.066	0.070	0.108	0.164	0.220
	0.95	0.062	0.068	0.102	0.190	0.242
GLS						
$\theta$	0.0	0.054	0.072	0.090	0.166	0.212
	0.3	0.058	0.076	0.090	0.176	0.228
	0.6	0.058	0.074	0.104	0.196	0.244
	0.9	0.060	0.070	0.118	0.232	0.324
	0.95	0.062	0.064	0.112	0.246	0.356

Note: MC = 500, where MC is the number of Monte Carlo samples.

Table 2

Mean of the GLS estimates of  $\rho$

$\hat{\rho}$		$\rho$				
		0.0	0.3	0.6	0.9	0.95
$\theta$	0.0	-0.032	0.236	0.500	0.742	0.771
	0.3	-0.033	0.235	0.500	0.745	0.777
	0.6	-0.034	0.234	0.500	0.746	0.779
	0.9	-0.036	0.232	0.496	0.740	0.773
	0.95	-0.039	0.228	0.492	0.737	0.768

Note: MC = 500.

sake of brevity we will not present the details. The results for the bootstrap- $t$  are presented in Table 3.

As can be readily seen, the size distortion problem has been corrected except at extreme values of  $\theta$  and  $\rho$ . In fact the test statistics based on OLS are better than those of GLS. Bickel and Freedman (1983) propose rescaling residuals by a factor  $\sqrt{n/(n-m)}$  which is adopted later in many bootstrap applications. ( $n$  is the number of observations and  $m$  the number of regressors.) The rescaling is

Table 3  
Empirical test size based on the percentile- $t$  bootstrap method ( $H_0: \beta = \beta_0$ , nominal size: 5%)

OLS		$\rho$				
		0.0	0.3	0.6	0.9	0.95
$\theta$	0.0	0.038	0.034	0.046	0.066	0.092
	0.3	0.044	0.038	0.048	0.052	0.078
	0.6	0.046	0.052	0.050	0.044	0.084
	0.9	0.044	0.046	0.056	0.088	0.098
	0.95	0.038	0.044	0.062	0.092	0.122
GLS						
$\theta$	0.0	0.034	0.044	0.056	0.056	0.102
	0.3	0.038	0.044	0.058	0.060	0.102
	0.6	0.050	0.060	0.060	0.068	0.114
	0.9	0.046	0.046	0.060	0.104	0.156
	0.95	0.044	0.048	0.070	0.112	0.164

Note: MC = 500, NB = 500, where NB is the number of bootstrap samples from each Monte Carlo sample.

due to the fact that the residuals tend to be smaller than the true regression errors. In our case, rescaling the residuals produced almost the same results. Thus the corresponding simulation results are not reported.

## 6. Block bootstrap methods applied to asymptotic procedures

The proceeding discussion referred to bootstrap methods applied to OLS and GLS estimators. We will now discuss bootstrap methods applied to two asymptotic procedures that correct for the endogeneity and serial correlation problems: Phillips and Hansen's FMOLS and Johansen's MLVECM. Both suggest asymptotically valid tests to test  $H_0: \beta = \beta_0$ . It is now well known that these tests have substantial size distortions in finite samples. They tend to over reject the null. We want to investigate how the moving block and stationary bootstrap procedures perform in correcting these size distortions.

The DGP considered is the same as in section 4 but we drop the assumption of exogeneity of the regressors. Instead we assume  $\theta = 0$ , that is no serial correlation in the errors in equation (6). This is done to save on computations as well as ease of presentation. The DGP we assume is as in section 4 with  $\theta = 0$ ,  $\lambda = \{-0.5, 0.0, 0.5\}$  and  $\rho = \{0.0, 0.5, 0.9\}$ . We assume MC = 500, NB = 200.

In the generation of the bootstrap samples the residual method is used. As discussed in section 3, the results from bootstrapping the data (which were

worse) were discarded. In the residual based method the asymptotically efficient procedures under consideration are used to estimate the model and then the residuals are resampled. However, as discussed in section 3 the residuals resampled were obtained under the null (sampling scheme  $S_3$ ), that is,  $\tilde{u}_t = y_t - \beta_0 x_t$ ,  $\hat{v}_t = \Delta x_t$  and the pairs  $(\tilde{u}_t, \hat{v}_t)$  were resampled. The procedure for the FMOLS and the moving block bootstrap is as follows:

1. Estimate (3) using the original sample by FMOLS. Calculate the associated  $t$ -statistic  $\hat{t}$  for testing the null  $H_0: \beta = \beta_0$ .
2. Calculate the regression residuals under the null. Form moving block pairs  $\{\tilde{u}_t, \dots, \tilde{u}_{t+k-1}, \hat{v}_t, \dots, \hat{v}_{t+k-1}\}$  from  $t$  to  $t+k-1$  and  $t = 1, \dots, n-k+1$ , where  $\tilde{u}_t = y_t - \beta_0 x_t$ ,  $\hat{v}_t = \Delta x_t$ .
3. Draw blocks  $\{u_t^*, \dots, u_{t+k-1}^*, v_t^*, \dots, v_{t+k-1}^*\}$  randomly with replacement from the residual moving block pairs. Construct  $x_t^* = x_{t-1}^* + v_t^*$  (with initial value set to zero) and  $y_t^* = \beta_0 x_t^* + u_t^*$ .
4. Estimate (3) using the bootstrap sample by FMOLS. Calculate the associated  $t$ -statistic for testing the null  $H_0: \beta = \beta_0$ . This  $t$ -statistic is called  $t_t^*$ .
5. Repeat step 3 and step 4 NB times. Obtain the bootstrap parameter distribution  $\beta_1^*, \beta_2^*, \dots, \beta_{NB}^*$  and the distribution of the  $t$ -statistic  $t_1^*, t_2^*, \dots, t_{NB}^*$ .
6. The 2.5% quantiles (at the two ends) of the bootstrap  $t$ -statistics,  $t_L^*$  and  $t_H^*$ , are obtained. Reject the null if  $\hat{t} > t_H^*$  or  $\hat{t} < t_L^*$ .

The procedure is similar for MLVECM and for the stationary bootstrap. The results for FMOLS and MLVECM are presented in Tables 4 and 5, where the mean, standard deviation, and root mean squared errors of the bias term and the empirical test sizes are reported.

The moving block bootstrap is based on a block length of 10. Thus with a sample size of 50 there will be 41 blocks available for resampling. For the stationary bootstrap since the average block length is  $1/p$  where  $p$  is the parameter in the geometric distribution we assume  $p = 0.1$  to make the results comparable to those of the moving block bootstrap procedure. The results presented in Tables 4 and 5 suggest the following:

(i) Hypothesis tests based on the asymptotic distributions have substantial distortions and the two bootstrap procedures correct these distortions. The empirical test sizes based on the two bootstrap procedures are closer to the nominal 5% significance level.

(ii) In the case of FMOLS, the MEAN, STDE and RMSE are all reduced to a certain degree and thus the bootstrap procedures considered can also be used for bias correction.

(iii) For the MLVECM procedure, however, this is not true. The bias correction is not feasible because of the high variance in the estimates. But the correction for size distortion in the test statistics can still be implemented and it is as good as in the case of FMOLS.



Table 4

Comparison of MBB and SB based on FMOLS procedure with the asymptotic results from FMOLS ( $H_0: \beta = \beta_0$ , nominal size: 5%)

$\lambda$	$\rho$		MEAN	STDE	RMSE	SIZE
- 0.5	0.9	FMOLS	- 0.184	0.341	0.387	0.430
		MBB	- 0.111	0.316	0.335	0.064
		SB	- 0.100	0.297	0.313	0.064
	0.5	FMOLS	- 0.047	0.135	0.143	0.236
		MBB	- 0.005	0.119	0.119	0.038
		SB	- 0.005	0.115	0.115	0.020
	0.0	FMOLS	- 0.035	0.085	0.092	0.218
		MBB	- 0.003	0.071	0.071	0.030
		SB	- 0.003	0.069	0.069	0.030
0.0	0.9	FMOLS	- 0.018	0.411	0.411	0.396
		MBB	- 0.015	0.374	0.375	0.062
		SB	- 0.012	0.349	0.349	0.064
	0.5	FMOLS	0.002	0.145	0.145	0.214
		MBB	0.001	0.128	0.128	0.040
		SB	0.002	0.120	0.120	0.026
	0.0	FMOLS	0.003	0.081	0.081	0.174
		MBB	0.001	0.071	0.071	0.040
		SB	0.002	0.067	0.067	0.030
0.5	0.9	FMOLS	0.172	0.407	0.442	0.404
		MBB	0.097	0.372	0.385	0.068
		SB	0.090	0.347	0.359	0.056
	0.5	FMOLS	0.051	0.144	0.152	0.242
		MBB	0.004	0.125	0.125	0.028
		SB	0.006	0.121	0.121	0.026
	0.0	FMOLS	0.038	0.085	0.093	0.238
		MBB	0.003	0.070	0.070	0.030
		SB	0.004	0.069	0.069	0.024

Note: (1) MBB indicates moving block bootstrap based on FMOLS. Each block has length 10.

(2) SB indicates stationary bootstrap based on FMOLS. The probability  $p = 0.1$ .

(3) MC = 500, NB = 200.

(4) Note, however, that the means and standard deviations in this table and the next may be estimates of non-existent population moments. See Sargan (1982).

(iv) The stationary bootstrap produces similar results compared to those of the moving block bootstrap. However further simulation results indicate that the moving block bootstrap is sensitive to the selection of the block length  $k$ , while the stationary bootstrap is less sensitive to the selection of the probability  $p$ .

Table 5

Comparison of MBB and SB based on MLVECM procedure with the asymptotic results from MLVECM ( $H_0: \beta = \beta_0$ , nominal size: 5%)

$\lambda$	$\rho$		MEAN	STDE	RMSE	SIZE
-0.5	0.9	MLVECM	0.673	11.299	11.319	0.594
		MBB	0.727	11.802	11.824	0.076
		SB	0.895	12.710	12.741	0.074
	0.5	MLVECM	-0.012	1.938	1.938	0.408
		MBB	-0.007	2.057	2.057	0.048
		SB	-0.038	2.027	2.027	0.044
	0.0	MLVECM	-0.019	0.590	0.591	0.340
		MBB	-0.024	0.627	0.628	0.034
		SB	-0.232	4.280	4.286	0.020
0.0	0.9	MLVECM	-0.332	3.984	3.998	0.636
		MBB	-0.536	5.655	5.680	0.116
		SB	0.881	47.655	47.663	0.106
	0.5	MLVECM	-0.482	9.470	9.482	0.458
		MBB	-0.397	9.584	9.592	0.052
		SB	-0.583	9.804	9.821	0.044
	0.0	MLVECM	-0.046	0.771	0.773	0.360
		MBB	-0.056	0.837	0.839	0.038
		SB	0.073	4.488	4.489	0.024
0.5	0.9	MLVECM	-0.306	8.590	8.596	0.610
		MBB	0.071	10.419	10.419	0.104
		SB	-0.656	10.317	10.338	0.096
	0.5	MLVECM	0.419	9.060	9.070	0.406
		MBB	1.353	23.453	23.492	0.052
		SB	-9.881	229.05	229.26	0.040
	0.0	MLVECM	-0.176	4.193	4.197	0.328
		MBB	-0.185	4.177	4.181	0.040
		SB	-0.160	4.109	4.112	0.038

Note: (1) MBB indicates moving block bootstrap based on MLVECM. Each block has length 10.

(2) SB indicates stationary bootstrap based on MLVECM. The probability  $p = 0.1$ .

(3) MC = 500, NB = 200.

## 7. Conclusions

There are several results in this paper that would be useful guides for empirical research in this area. There are two problems researchers face:

- (i) The choice of an appropriate estimation method for a cointegrated system.
- (ii) The appropriate method of inference, once an estimation method is chosen.

For (i) there are several Monte Carlo studies. For (ii) the inference is usually based on asymptotic theory.

This paper is not concerned with issue (i). Regarding (ii), the paper argues that one should consider estimates for which bootstrap-based correction has been applied and one should use bootstrap based critical values rather than the critical values from the asymptotic distributions.

Some other questions that are often raised relate to the problems of

- (1) how to generate the bootstrap samples and
- (2) what particular bootstrap method should be used (recursive, moving block, stationary).

Regarding (1) there is first the issue of bootstrapping the data vs. bootstrapping the residuals. The paper argues that in the case of unit root processes and cointegrating regressions, one should not bootstrap the data. One should use bootstrapping residuals.

The next question is ‘What residuals?’ – OLS residuals or restricted residuals based on the null hypothesis to be tested? The paper argues that first, the residuals used and the pseudo data generation should be consistent with the test statistic used. Second, although three procedures are available (described as  $(S_1, T_1)$ ,  $(S_2, T_2)$ ,  $(S_3, T_2)$  where  $S$  denotes sampling scheme and  $T$  denotes test statistic), in the case of unit root and cointegration models,  $(S_1, T_1)$  has invalid limiting distribution and only  $(S_2, T_2)$  and  $(S_3, T_2)$  should be used. In particular,  $(S_3, T_2)$  is superior to  $(S_2, T_2)$  based on the performance in Monte Carlo studies.

Regarding issue (2), the bootstrap method used, the paper evaluates recursive, moving block and stationary bootstrap methods. If the serial correlation structure in the model is misspecified, the recursive bootstrap gives poor results.

The theoretical justification for the moving block bootstrap is extremely (almost impossibly) complicated and has not been attempted here. However, the simulation results presented suggest its superiority over asymptotic inference. Also, the simulations suggest that the stationary bootstrap method works even better. So, despite the lack of theoretical basis, these methods have been found to significantly improve on asymptotic inference and hence are strongly suggested for use by empirical researchers.

The results we have presented suggest that the substantial size distortions of the asymptotic tests of significance can be corrected by properly defined bootstrap methods. However, this improvement is of no use if there is also a substantial loss of power. We have investigated this problem. The detailed power curves are not presented here for the sake of brevity. In all cases, we found that there was loss of power but this was small compared with the considerable improvement in the size distortions. However, the results presented in this paper do not constitute an exhaustive study and should be treated as suggestive. It is also important to check the relative performance of bootstrap vs. asymptotic inference for larger sample sizes, say 100 and 200, and the sensitivity of the stationary

bootstrap to choice of  $p$ . Clearly if there is no serial correlation in the errors, the optimal choice of  $p$  is  $p = 1$ . So, we also have to analyze the consequences of specification errors. Given that the simulation studies of bootstrap methods are very time consuming, we do not have the results of an exhaustive study yet. Meanwhile, we have presented some guidelines on how bootstrap procedures should be used in the analysis of cointegrated systems.

## References

- Banerjee, A., J. Dolado, D.F. Hendry, and G.W. Smith, 1986, Exploring equilibrium relationships in econometrics through static models: Some Monte Carlo evidence, *Oxford Bulletin of Economics and Statistics* 48, 253–277.
- Basawa, I.V., A.K. Mallik, W.P., McCormick, and R.L. Taylor, 1991a, Bootstrapping unstable first order autoregressive processes, *Annals of Statistics* 19, 1098–1101.
- Basawa, I.V., A.K. Mallik, W.P. McCormick, and R.L. Taylor, 1991b, Bootstrapping test of significance and sequential bootstrap estimation for unstable first order autoregressive processes, *Communications in Statistics -Theory and Methods* 20, 1015–1026.
- Beran, R., 1987, Prepivoting to reduce level error of confidence sets, *Biometrika* 74, 457–468.
- Beran, R., 1988, Prepivoting test statistics: A bootstrap view of asymptotic refinements, *Journal of the American Statistical Association* 83, 687–697.
- Bickel, P.J. and D.A. Freedman, 1983, Bootstrapping regression models with many parameters, in: P. Bickel, Doksum, K. and Hodges, J.L. Belmont, eds., *A Festschrift for Erich Lehmann*, (Wadsworth, California) pp. 28–48.
- Carlstein, E., 1986, The use of subsample values for estimating the variance of a general statistic from a stationary sequence, *Annals of Statistics* 14, 1171–1179.
- Davison, A.C. and P. Hall, 1993, On studentizing and blocking methods for implementing the bootstrap with dependent data, *Australian Journal of Statistics* 35, 212–224.
- Efron, B., 1979, Bootstrap methods: another look at the jackknife, *The Annals of Statistics* 7, 1–26.
- Efron, B., 1981, Censored data and the bootstrap, *Journal of American Statistics Association* 76, 312–319.
- Efron, B. and G. Gong, 1983, A leisurely look at the bootstrap, jackknife and cross-validation, *The American Statistician* 37, 36–48.
- Efron, B. and R. Tibshirani, 1986, Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy, *Statistical Science* 1, 54–77.
- Engle, R. and C.W.J. Granger, 1987, Cointegration and error correction: representations, estimation and testing, *Econometrica* 55, 252–276.
- Ferretti, N. and J. Romo, 1994, Unit root bootstrap tests for AR(1) models, working paper, Division of Economics, Universidad Carlos III de Madrid.
- Freedman, D.A., 1981a, Bootstrapping regression models, *The Annals of Statistics* 9, 1218–1228.
- Freedman, D.A., 1981b, Bootstrapping regression models, *The Annals of Statistics* 9, 1229–1238.
- Freedman, D.A. and S.C. Peters, 1984, Bootstrapping a regression equation: some empirical results, *Journal of the American Statistical Association* 79, 97–106.
- Gonzalo, J., 1994, Five alternative methods of estimating long-run equilibrium relationships, *Journal of Econometrics* 60, 203–233.
- Granger, C.W.J., 1981, Some properties of time series data and their use in econometric model specification, *Journal of Econometrics* 16, 121–130.
- Hall, P. 1988, Theoretical comparison of bootstrap confidence intervals, *Annals of Statistics* 16, 927–953.

- Hall, P. (1992), *The Bootstrap and Edgeworth Expansion*. (Springer, New York).
- Hall, P. and J.L. Horowitz, 1993, Corrections and blocking rules for the block bootstrap with dependent data, Working Paper #93-11, Department of Economics, University of Iowa.
- Hall, P. and J.L. Horowitz, 1994, Bootstrap critical values for tests based on generalized-method-of-moments estimators, manuscript, Department of Economics, University of Iowa.
- Hall, P. and S.R. Wilson, 1991, Two guidelines for bootstrap hypothesis testing, *Biometrics* 47, 757–762.
- Hansen, B.E. and P.C.B. Phillips, 1990, Estimation and inference in models of cointegrating: A simulation study, *Advances in Econometrics* 8, 225–248.
- Hargreaves, C., 1994, A review of methods of estimating cointegrating relationships, in: C.Hargreaves, ed., *Nonstationary Time Series Analysis and Cointegration*, 87–131, Oxford.
- Hartigan, J.A., 1986, Comment on the paper by Efron and Tibshirani, *Statistical Science* 1, 75–76.
- Hinkley, D.V., 1988, Bootstrap method, *Journal of the Royal Statistical Society, Series B* 50, 321–337.
- Inder, B., 1993, Estimating long-run relationships in economics, *Journal of Econometrics* 57, 53–68.
- Jeong, J. and G.S. Maddala, 1993, A perspective on application of bootstrap methods in econometrics, in: G.S. Maddala, C.R. Rao and H.D. Vinod, eds., *Handbook of Statistics Vol. 11: Econometrics*, (North-Holland, Amsterdam) 573–610.
- Johansen, S., 1988, Statistical analysis of cointegration vectors, *Journal of Economic Dynamics and Control* 12, 231–255.
- Johansen, S., 1991, Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregression models, *Econometrica* 59, 1551–1580.
- King, M.L. and D.E.A. Giles, 1984, Autocorrelation pretesting in the linear model: Estimation, testing and prediction, *Journal of Econometrics* 23, 35–48.
- Kiviet, J.F., 1984, Bootstrap inference in lagged dependent variable models, Working paper, University of Amsterdam.
- Künsch, H.R., 1989, The jackknife and the bootstrap for general stationary observations, *The Annals of Statistics* 17, 1217–1241.
- Li, H. and G.S. Maddala, 1995, Estimating long-run economic relationships in economics: Some comments and further results, Department of Economics, The Ohio State University.
- Liu, R.Y. and K. Singh, 1992, Moving blocks jackknife and bootstrap capture weak dependence, in: R. LePage, and L. Billard, eds., *Exploring the Limits of Bootstrap*, (Wiley, New York) 225–248.
- Nankervis, J.C. and N.E. Savin, 1994, The level and power of the bootstrap  $t$ -test in the AR(1) model with trend, manuscript, Department of Economics, University of Surrey and University of Iowa.
- Park, J.Y., 1992, Canonical cointegrating regressions, *Econometrica* 60, 119–143.
- Phillips, P.C.B. and B.E. Hansen, 1990, Statistical inference in instrumental variables regression with  $I(1)$  process, *Review of Economic Studies* 57, 99–125.
- Phillips, P.C.B. and M. Loretan, 1991, Estimating long-run economic equilibria, *Review of Economic Studies* 58, 407–436.
- Phillips, P.C.B. and J.Y. Park, 1988, Asymptotic equivalence of ordinary least squares and generalized least squares in regressions with integrated regressors, *Journal of the American Statistical Association* 83, 111–115.
- Politis, D.N. and J.P. Romano, 1994, The stationary bootstrap, *Journal of American Statistical Association* 89, 1303–1313.
- Rayner, R.K., 1990, Bootstrapping  $p$ -values and power in the first-order autoregression: A Monte Carlo investigation, *Journal of Business and Economic Statistics* 8, 251–263.
- Rayner, R.K., 1991, Resampling methods for tests in regression models with autocorrelated errors, *Economics Letters* 36, 281–284.
- Sargan, J.D., 1982, On Monte Carlo estimates of moments that are infinite, in: R.L. Basman and George F. Rhodes, Jr., eds., *Advances in Econometrics Vol. 1* (JAI Press, London) 267–299.

- Shea, G.S., 1989a, A re-examination of excess rational price approximations and excess volatility in the stock market, in: R.M.C. Guimaraces et al., eds., *A Reappraisal of the Efficiency of Financial markets* (Springer, Berlin).
- Shea, G.S., 1989b, Ex-post rational price approximations and the empirical reliability of the present-value relation, *Journal of Applied Econometrics* 4, 139–159.
- van Giersbergen, N.P.A. and J.F. Kiviet, 1993, How to implement bootstrap hypothesis testing in regression models, Paper presented at Fourth Meeting of the European Conference Series in Quantitative Economics and Econometrics: (EC<sup>2</sup>), December 1993, Oxford.
- Veall, M.R., 1986, Bootstrapping regression estimators under first-order serial correlation, *Economics Letters* 21, 41–44.
- Veall, M.R., 1987, Bootstrapping the probability distribution of peak electricity demand, *International Economic Review* 28, 203–212.