

The fast iterated bootstrap

Russell Davidson, Mirza Trokić

▶ **To cite this version:**

Russell Davidson, Mirza Trokić. The fast iterated bootstrap. Journal of Econometrics, 2020, 218 (2), pp.451-475. 10.1016/j.jeconom.2020.04.025 . hal-02965001

HAL Id: hal-02965001
https://amu.hal.science/hal-02965001

Submitted on 9 Feb 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

The Fast Iterated Bootstrap

by

Russell Davidson*

Department of Economics and CIREQ
McGill University
Montreal, Quebec, Canada
H3A 2T7

AMSE
Aix-Marseille Université
CNRS, EHESS, Centrale Marseille
13205 Marseille cedex 01, France

email: **russell.davidson@mcgill.ca**

and

Mirza Trokić

IHS Markit, Economics and Country Risk, EViews
Irvine, CA 92618, USA
mirza.trokic@ihsmarkit.com

Key words: bootstrap iteration, fast iterated bootstrap

JEL codes: C10, C12, C15, C63

Opinions, statements, estimates and projections in this article are solely those of the individual author(s). They do not necessarily reflect the opinions of IHS Markit or any of its affiliates. IHS Markit has no obligation to update, modify or amend this article, or otherwise to notify readers, in the event that any matter stated herein, or any opinion, projection, forecast, or estimate set forth herein, changes, or subsequently becomes inaccurate.

October 2019

* Corresponding author

Abstract

The standard forms of bootstrap iteration are very computationally demanding. As a result, there have been several attempts to alleviate the computational burden by use of approximations. In this paper, we extend the fast double bootstrap of Davidson and MacKinnon (2007) to higher orders of iteration, and provide algorithms for their implementation. The new methods make computational demands that increase only linearly with the level of iteration, unlike standard procedures, whose demands increase exponentially. In a series of simulation experiments, we show that the fast triple bootstrap improves on both the standard and fast double bootstraps, in the sense that it suffers from less size distortion under the null with no accompanying loss of power.

1. Introduction

Bootstrap iteration has been discussed in a wide variety of contexts from not long after the invention of the bootstrap by Efron (1979). The first article to mention the topic of the iterated bootstrap to our knowledge is Hall (1986), a paper followed quickly by two articles by Beran (1987) and (1988), in which the double bootstrap is introduced. An extensive theoretical discussion, with examples, is given in Hall and Martin (1988), in which the approaches of Hall and Beran are unified. Bootstrap iteration is also mentioned in Hall's (1992) influential book, in which Edgeworth expansions play a central role.

However, the computational burden of bootstrap iteration was then very heavy, and still is, although to a lesser extent. Consequently, a number of attempts have been made to lighten this burden. An early one is found in DiCiccio, Martin, and Young (1992). It makes no use of Edgeworth expansions, preferring instead other asymptotic arguments. Lee and Young (1995) develop procedures for bootstrap confidence intervals with an iterated bootstrap that does not involve any resampling, replacing it with analytic asymptotic approximations. Chan and Lee (2001) derive an algorithm for infinitely iterated bootstrap bias correction, by considering bootstrap iteration as a Markov process. In Lee and Young (2003), possibilities are investigated of weighted resampling to mimic the results of bootstrap iteration. Ho and Lee (2005) consider bootstrap iteration in conjunction with smoothing of the discrete distribution associated with conventional resampling.

Another attempt to alleviate the computation burden is the fast double bootstrap (FDB) exposited in Davidson and MacKinnon (2007). The technique had been used previously: Davidson and MacKinnon (2002) use it to improve the reliability of bootstrap tests of non-nested linear regression models; Omtzigt and Fachin (2006) show that it gives better size correction than the single bootstrap with or without Bartlett correction in the cointegrated VAR model; Lamarche (2004) investigates some of its numerical properties; Davidson (2006) uses it in developing bootstrap tests of cointegration with fractionally integrated time series. A more recent paper, Ouysse (2013) develops a version of the FDB for the purposes of bias correction. Giacomini, Politis, and White (2013) provide a formal analysis of the FDB, under the name of a "warp-speed" method, and give conditions under which the method is justified asymptotically. Most recently, Chang and Hall (2015) revert to the use of Edgeworth expansions to explore rates of convergence of the FDB, and show that it improves on the single bootstrap for the purpose of bias correction, but not for constructing bootstrap confidence intervals.

The starting point for this paper is found in Davidson (2010), in which the FDB is studied and compared with the standard double bootstrap and some other procedures aimed at improving bootstrap reliability. In the context of a much simplified unit root test, it is shown that estimating the distribution of the FDB P value leads to a reduction in the size distortion of the FDB, which is less distorted than the single bootstrap, and roughly comparable to the standard double bootstrap. The procedure is more or less a bootstrap of the FDB, which gives rise to an approximation to the bootstrapped double bootstrap, that is, the twice iterated, or triple, bootstrap. The FDB algorithm is not a nested bootstrap procedure. But when it is bootstrapped, what results is a nested bootstrap procedure, essentially as computationally intensive as the double bootstrap.

It is therefore interesting to see whether it is possible to make use of the approximations that simplify the original double bootstrap algorithm to the FDB, but to the bootstrapped FDB. One would then effectively obtain an approximation to the standard triple bootstrap, an approximation that can reasonably be termed the fast triple bootstrap (FTB). In this paper, after having discussed bootstrap iteration and developing some suitable notation, we show how fast versions of any order of iterated bootstrap can be defined. Further, whereas almost all of the papers cited above restrict attention to the conventional resampling bootstrap, or varieties of the block bootstrap, which also involve resampling, our development here is applicable more widely, and covers the parametric bootstrap, the wild bootstrap, and other bootstrap procedures quite generally.

In the next section, we set up notation for dealing with bootstrap iteration, and detail some assumptions made in the subsequent material. In section 3, bootstrap iteration to any order is defined formally. Then, in section 4, we review the theory of the FDB, and give algorithms for the computation of FDB P values and confidence intervals. Section 5 is where we pursue an approach analogous to that which leads to the FDB in order to motivate and define the fast triple bootstrap, for which a computational algorithm is detailed. Once we can see how to get as far as the FTB, the road is clear to defining arbitrary orders of fast iterated bootstraps; that is done in section 6. In section 7, we report the results of three sets of simulation experiments, one based on the experimental design of Davidson (2010) for a unit root test, the next on a test for ARCH effects, and the third on a test for serial correlation of the disturbances of a regression model. Section 8 concludes.

2. Concepts and Notations

We denote by $\mathbb{M}_0$ the set of data-generating processes (DGPs) that satisfy a null hypothesis we wish to test. The test statistic used is denoted by τ. In general, τ has a different distribution under the different DGPs in $\mathbb{M}_0$, and it certainly has a different distribution under DGPs in the model, $\mathbb{M}$ say, that represents the alternative hypothesis. Here $\mathbb{M}_0 \subset \mathbb{M}$. It is conventional to suppose that τ is defined as a random variable on some suitable probability space, on which we define a different probability measure for each different DGP.

Rather than using this approach, we define a probability space $(\Omega, \mathcal{F}, P)$, with just one probability measure, P. Then we treat the test statistic τ as a stochastic process with the model $\mathbb{M}$ as index set. We have

$$\tau \; : \; \mathbb{M} \times \Omega \to \mathbb{R}. \tag{1}$$

Since most of the discussion of the paper is couched in the language of simulation, the probability space can, for our present purposes, be taken to be that of a random number generator. A realisation of the test statistic is therefore written as $\tau(\mu, \omega)$, for some $\mu \in \mathbb{M}$ and $\omega \in \Omega$.

Since this is an unconventional way of proceeding, it is probably necessary to explain why it is useful for our present purposes. First, it accurately represents how one sets about doing a simulation experiment. The notation ω stands for the random (or pseudo-random)

numbers obtained from the computer's random number generator, and μ stands for the non-random computer program that transforms the raw random numbers into realisations of the statistic or estimator under study. It makes it easy to describe experiments in which one uses the *same* random numbers with different DGPs, as well as the operation used whenever the bootstrap is implemented by simulation, namely generating a large number of independent and identically distributed (IID) realisations from one and the same DGP. In addition, it will be seen later that our notation allows an accurate description of bootstrap iteration.

It may be thought that the approach nonetheless does violence to our conception of real-world DGPs. This point is taken up in Davidson (2015), where the thesis is that, by using models to describe the real world, we invent a sort of virtual reality which can, if we so wish, be implemented on a computer.

Assumption 1: measurability and continuity

For all $\mu \in \mathbb{M}$, the mapping

$$\tau_\mu : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}), \quad \tau_\mu(\omega) = \tau(\mu, \omega),$$

where $\mathcal{B}$ is the Borel sigma-algebra defined on the real line $\mathbb{R}$, is measurable. The probability measure induced on $\mathbb{R}$ by τ_μ from the measure P on Ω is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$, and has a density continuous on the interior of its support.

Assumption 2: approximate P value

For all $\mu \in \mathbb{M}$, the range of τ_μ is the $[0, 1]$ interval, and its density is nowhere zero in the interior of the interval. The statistic takes the form of an approximate P value, which leads to rejection when the statistic is too small.

Remarks:

Measurability in Assumption 1 imposes no meaningful restriction. Absolute continuity, on the other hand, is restrictive, but it makes for a considerable simplification of the theoretical analysis, and serves as a useful approximation when the actual distribution is discrete. The specification of the range of τ_μ in Assumption 2 is made purely for notational convenience, and, if Assumption 1 is satisfied, is completely unrestrictive. The requirement of a non-zero density is made so that the corresponding distribution and quantile functions are strictly increasing on $[0, 1]$.

Let $R_0 : [0, 1] \times \mathbb{M} \to [0, 1]$ be the cumulative distribution function (CDF) of τ under the DGP μ:

$$R_0(x, \mu) = P\big\{\omega \in \Omega : \tau(\mu, \omega) \leq x\big\}. \tag{2}$$

Let $Q_0 : [0, 1] \times \mathbb{M} \to [0, 1]$ be the corresponding quantile function. Since by Assumption 2 R_0 and Q_0 are continuous and strictly increasing, we have the relations

$$R_0\big(Q_0(x, \mu), \mu\big) = x = Q_0\big(R_0(x, \mu), \mu\big). \tag{3}$$

Suppose that we have a statistic computed from a data set that was generated by a DGP $\mu \in \mathbb{M}$. Denote this statistic as $t \equiv \tau(\mu, \omega)$. If μ belongs to the null-hypothesis model $\mathbb{M}_0$, the ideal P value, $R_0(t, \mu)$, is a drawing from $\mathrm{U}(0, 1)$, the uniform distribution on $[0, 1]$. Exact inference on the basis of the ideal P value is of course in general infeasible, because μ is unknown.

Remark:

If τ is a pivotal statistic with respect to $\mathbb{M}_0$, then the distribution of $\tau(\mu, \cdot)$ is the same for all $\mu \in \mathbb{M}_0$. In such a case, exact inference can be based on a Monte Carlo test; see for instance Dufour and Khalaf (2001).

The principle of the bootstrap is that, when we want to use some function or functional of an unknown DGP μ, we use an estimate of μ in its place. This estimate is the bootstrap DGP. Analogously to (1), define the DGP-valued stochastic process and set of random variables β_μ, $\mu \in \mathbb{M}$, as follows:

$$\beta \ : \ \mathbb{M} \times \Omega \to \mathbb{M}_0, \qquad \beta_\mu \ : \ (\Omega, \mathcal{F}) \to \mathbb{M}_0, \ \text{with } \beta_\mu(\omega) = \beta(\mu, \omega),$$

where, although the model $\mathbb{M}$ on the left-hand side may include the entire alternative hypothesis, we insist that the $\mathbb{M}_0$ on the right-hand side is the null model under test (the first golden rule of bootstrapping). The bootstrap DGP computed from the data set for which the statistic is $\tau(\mu, \omega)$ is then $b \equiv \beta(\mu, \omega)$, with the *same* realisation ω. The notation neatly encapsulates the fact that the distribution of statistics generated by the bootstrap DGP is *conditional* on the realised sample. The bootstrap statistic, or bootstrap P value, is then $R_0(t, b)$. Since we have assumed that the distribution of τ is absolutely continuous for all $\mu \in \mathbb{M}_0$, and since the bootstrap DGP $b \in \mathbb{M}_0$, the bootstrap statistic has an absolutely continuous distribution.

Unless the function R_0 is known analytically, simulation is necessary for the computation of the bootstrap statistic. We make the definition

$$\hat{R}_0(x, \mu) = \frac{1}{B} \sum_{j=1}^{B} \mathrm{I}\big(\tau(\mu, \omega_j^*) \leq x\big), \tag{4}$$

where the ω_j^* are IID drawings from the random number generator, and $\mathrm{I}(\cdot)$ is the indicator function, equal to one when the Boolean argument is true, and to zero when it is false. As $B \to \infty$, $\hat{R}_0(\alpha, \mu)$ tends almost surely to $R_0(\alpha, \mu)$ by the strong law of large numbers. Accordingly, we estimate the bootstrap statistic by $\hat{R}_0(t, b)$. However, in the theoretical discussion that follows, we imagine that B is infinite, so that the function R_0 is available.

3. Bootstrap Iteration

It is useful at this point to define a stochastic process that represents the bootstrap statistic. Since $t = \tau(\mu, \omega)$ and $b = \beta(\mu, \omega)$, the definition is

$$p_1 \ : \ \mathbb{M} \times \Omega \to \mathbb{R}, \quad p_1(\mu, \omega) = R_0\big(\tau(\mu, \omega), \beta(\mu, \omega)\big). \tag{5}$$

Since by absolute continuity R_0 is a continuous function, it follows that p_1 also has an absolutely continuous distribution. We denote the continuous CDF of $p_1(\mu, \omega)$ by $R_1(\cdot, \mu)$.

The random variable $R_1\big(p_1(\mu, \omega), \mu\big)$ is, by construction, distributed as U(0,1). But, as with $R_0\big(\tau(\mu, \omega), \mu\big)$, which is also distributed as U(0,1), this fact is not enough to allow exact inference, because the actual μ that generates the data is unknown, except of course in the context of a simulation experiment. However, the bootstrap principle can again be applied, and the unknown μ replaced by the estimate b, which can be computed from observed data. This leads to the double bootstrap, of which the P value, for realisations t and b, can be written as

$$R_1\big(R_0(t, b), b\big). \tag{6}$$

This is just the bootstrap estimate of the probability mass in the distribution of the single bootstrap statistic to the left of the realisation $R_0(t, b)$. Expressed as a random variable, this double bootstrap P value is

$$p_2(\mu, \omega) \equiv R_1\big(R_0(\tau(\mu, \omega), \beta(\mu, \omega)), \beta(\mu, \omega)\big) \tag{7}$$

If we write the right-hand side above as $R_1\big(p_1(\mu, \omega), b(\mu, \omega)\big)$, the analogy with the definition (5) of $p_1(\mu, \omega)$ is complete. This demonstrates that the double bootstrap effectively bootstraps the single bootstrap P value.

From that observation, it is clear that we can define iterated bootstraps as follows.

Definition 1: iterated bootstrap

For $k = 0, 1, 2, \ldots$, the k^{th} order bootstrap P value and its distribution are given by recurrence relation

$$R_k(\alpha, \mu) = P\big\{\omega \in \Omega : p_k(\mu, \omega) \leq \alpha\big\}, \tag{8}$$

$$p_{k+1}(\mu, \omega) = R_k\big(p_k(\mu, \omega), \beta(\mu, \omega)\big), \tag{9}$$

where we initialise the recurrence by the definition $p_0(\mu, \omega) = \tau(\mu, \omega)$.

Thus $p_{k+1}(\mu, \omega)$ is the bootstrap P value obtained by bootstrapping the k^{th} order P value $p_k(\mu, \omega)$. It estimates the probability mass in the distribution of the k^{th} order P value to the left of its realisation.

In all discussions about bootstrap iteration, it is implicitly or explicitly assumed that iteration leads to more reliable inference. In a wide variety of circumstances, it has been possible to demonstrate that asymptotic refinements accrue with bootstrap iteration. As

the implementation of the bootstrap involves no asymptotics, it is desirable to analyse the situation in finite samples.

We wish to consider a measure defined on the product space $[0,1] \times \mathbb{M}_0$, the space in which the two random variables $\tau(\mu, \omega)$ and $\beta(\mu, \omega)$ are jointly realised. The component $[0,1]$ of the product space can be equipped with the Borel σ-algebra $\mathcal{B}$, as in Assumption 1 but restricted to the interval $[0,1]$ on account of Assumption 2, but it is less obvious how to associate a σ-algebra to the model $\mathbb{M}$. The easiest way to do so is to introduce an L^∞ metric on the space of CDFs of random variables defined on $[0,1]$.

Definition 2: metric on $\mathbb{M}$

For any two DGPs μ and ν in $\mathbb{M}$, let the distance between them be defined as

$$d(\mu, \nu) = \sup_{x \in [0,1]} \big| R_0(x, \mu) - R_0(x, \nu) \big|. \tag{10}$$

The choice of the function R_0 in the definition of the metric is motivated by the fact that the different behaviour of the function when evaluated at a DGP μ and when evaluated at the bootstrap DGP $\beta(\mu, \omega)$ is what accounts for the bootstrap discrepancy, that is, the difference between the rejection probability (RP) of a bootstrap test and the nominal significance level.

The metric d can be used to define the σ-algebra of Borel sets on $\mathbb{M}$, which we write as $\mathcal{B}(\mathbb{M})$, generated by the open sets in the topology defined by the metric. Then we can define the product σ-algebra $\mathcal{B}[0,1] \times \mathcal{B}(\mathbb{M})$ on the product space $[0,1] \times \mathbb{M}_0$, thus giving the measurable space $\mathcal{M} \equiv \big([0,1] \times \mathbb{M}_0, \mathcal{B}[0,1] \times \mathcal{B}(\mathbb{M}_0)\big)$.

A DGP $\mu \in \mathbb{M}_0$ defines an induced measure P_μ on $\mathcal{M}$ as follows: for any Borel set $C \in \mathcal{B}[0,1] \times \mathcal{B}(\mathbb{M}_0)$,

$$P_\mu(C) = P\big\{\omega \in \Omega : \big(\tau(\mu, \omega), \beta(\mu, \omega)\big) \in C\big\}. \tag{11}$$

The measure P_μ has a projection measure B_μ defined on $\mathcal{B}(\mathbb{M}_0)$: for $B \in \mathcal{B}(\mathbb{M}_0)$,

$$B_\mu(B) = P\big\{\omega \in \Omega : \beta(\mu, \omega) \in B\big\}.$$

Assumption 3: existence of carrier measure

There exists a measure $m : \mathcal{B}(\mathbb{M}_0) \to [0,1]$ such that, for all $\mu \in \mathbb{M}_0$, the measure B_μ is absolutely continuous with respect to m.

The product measure M is defined naturally on $\mathcal{M}$ as $\mathcal{L}[0,1] \times m$, where $\mathcal{L}[0,1]$ is Lebesgue measure on $[0,1]$. For $A \in \mathcal{B}[0,1]$ and $B \in \mathcal{B}(\mathbb{M}_0)$, we have $M(A \times B) = \mathcal{L}(A)m(B)$.

Assumption 4: existence of density

The probability measure P_μ has a density $f_\mu : [0,1] \times \mathbb{M}_0 \to \overline{R}_+$, such that, if $C \in \mathcal{B}[0,1] \times \mathcal{B}(\mathbb{M}_0)$,

$$P_\mu(C) = \int_0^1 \int_{\mathbb{M}_0} \mathrm{I}\big((t,b) \in C\big) f_\mu(t,b)\, \mathrm{d}t\, m(\mathrm{d}b).$$

Assumption 5: bounded derivatives

For all $\mu \in \mathbb{M}$, there exists $r > 0$ such that both the density $R_0'(\cdot,\mu)$ and the quantile density function $Q_0'(\cdot,\mu)$ (in the terminology of Parzen (1979)) satisfy the following inequalities for all $x \in [0,1]$:

$$1 - r < R_0'(x,\mu) < 1 + r \quad \text{and} \quad 1 - r < Q_0'(x,\mu) < 1 + r.$$

This mild assumption quantifies the part of Assumption 2 which requires a nonzero density everywhere, and also rules out points at which the density may become infinite.

4. The Fast Double Bootstrap

The fast double bootstrap (FDB) of Davidson and MacKinnon (2007) is based on two approximations. The first is to assume that, for any $\mu \in \mathbb{M}_0$, the random variables τ_μ and β_μ are independent. The assumption is of course false except in special circumstances, but it holds asymptotically in many commonly encountered situations. By definition,

$$R_1(x,\mu) = P\big\{\omega \in \Omega : p_1(\mu,\omega) \le x\big\} = \mathrm{E}\big[\mathrm{I}\big(R_0(\tau(\mu,\omega),\beta(\mu,\omega)) \le x\big)\big]. \tag{12}$$

Use of the relation (3) between R_0 and Q_0 lets us write (12) as

$$R_1(x,\mu) = \mathrm{E}\big[\mathrm{I}\big(\tau(\mu,\omega) \le Q_0(x,\beta(\mu,\omega))\big)\big]$$

If τ_μ and β_μ are treated as though they were independent, then we have

$$\begin{aligned} R_1(x,\mu) &= \mathrm{E}\Big[\mathrm{E}\big[\mathrm{I}\big(\tau(\mu,\omega) \le Q_0(x,\beta(\mu,\omega))\big) \,|\, \beta(\mu,\omega)\big]\Big] \\ &= \mathrm{E}\big[R_0\big(Q_0(x,\beta(\mu,\omega)),\mu\big)\big] \end{aligned} \tag{13}$$

where the last step follows from the "Independence Lemma"– Lemma A.3 in Schilling and Partzsch (2012). Since in general τ_μ and β_μ are not independent, (13) is taken as an approximation.

Assumption FDB1: first FDB approximation

There exists a (small) quantity $\eta > 0$ such that, for all $\mu \in \mathbb{M}_0$, and for all $x \in [0,1]$,

$$\big|R_1(x,\mu) - \mathrm{E}\big[R_0\big(Q_0(x,\beta(\mu,\omega)),\mu\big)\big]\big| < \eta. \tag{14}$$

Consider now two identical probability spaces $(\Omega_1, \mathcal{F}_1, P_1)$ and $(\Omega_2, \mathcal{F}_2, P_2)$, and their product space $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \times \mathcal{F}_2, P_1 \times P_2)$. Define the stochastic process

$$\tau^1 \ : \ \mathbb{M} \times (\Omega_1 \times \Omega_2) \to \mathbb{R}$$

by the formula

$$\tau^1(\mu, \omega_1, \omega_2) = \tau\big(\beta(\mu, \omega_1), \omega_2\big). \tag{15}$$

Thus $\tau^1(\mu, \omega_1, \omega_2)$ can be thought of as a realisation of the bootstrap statistic when the underlying DGP is μ. We denote the CDF of τ^1 under μ by $R^1(\cdot, \mu)$. Thus

$$\begin{aligned} R^1(x, \mu) &= (P_1 \times P_2)\big\{(\omega_1, \omega_2) \in \Omega_1 \times \Omega_2 : \tau\big(\beta(\mu, \omega_1), \omega_2\big) \le x\big\} \\ &= \mathrm{E}\big[\mathrm{I}\big(\tau(\beta(\mu, \omega_1), \omega_2) \le x\big)\big] \\ &= \mathrm{E}\Big[\mathrm{E}\big[\mathrm{I}\big(\tau(\beta(\mu, \omega_1), \omega_2) \le x\big) \,|\, \mathcal{F}_1\big]\Big] \\ &= \mathrm{E}\big[R_0\big(x, \beta(\mu, \omega_1)\big)\big]. \end{aligned} \tag{16}$$

In the third line above, $\mathcal{F}_1$ denotes the σ-algebra generated by deterministic functions of ω_1, and the last step follows from the Independence Lemma.

We make the following assumption:

Assumption FDB2: Close bootstrap DGP

There exists a (small) quantity $\delta \ge 0$ such that, for all $\mu \in \mathbb{M}_0$ and for all $x \in [0, 1]$,

$$\big|R^1(x, \mu) - R_0(x, \mu)\big| < \delta \text{ and } \big|Q^1(x, \mu) - Q_0(x, \mu)\big| < \delta, \tag{17}$$

where $Q^1(\cdot, \mu)$ is the quantile function inverse to the CDF $R^1(\cdot, \mu)$. Further,

$$\big|\mathrm{E}\big[Q_0(x, \beta(\mu, \omega))\big] - Q_0(x, \mu)\big| < \delta.$$

The quantity δ measures the extent to which the bootstrap DGP can differ on average from the true DGP. We can now show that

Lemma FDB: second FDB approximation

Let r be as defined in Assumption 5, and δ be as defined in Assumption FDB2. Then for all $\mu \in \mathbb{M}_0$ and for all $x \in [0, 1]$,

$$\big|\mathrm{E}\big[R_0\big(Q_0(x, \beta(\mu, \omega)), \mu\big)\big] - R_0\big(Q^1(x, \mu), \mu\big)\big| < 4r\delta. \tag{18}$$

Proof : in Appendix A. ■

Note :

The proof proceeds by demonstrating the following approximations:

$$\begin{aligned} \mathrm{E}\big[R_0\big(Q_0(x,\beta(\mu,\omega)),\mu\big)\big] &\approx R_0\big(\mathrm{E}\big[Q_0(x,\beta(\mu,\omega))\big],\mu\big), \quad \text{and} \\ \mathrm{E}\big[Q_0(x,\beta(\mu,\omega))\big] &\approx Q^1(x,\mu), \end{aligned} \tag{19}$$

The approximation that underlines the FDB follows from Assumption FDB1, Assumption FDB2, and Lemma FDB.

Theorem FDB: FDB approximation

Make the definition:

$$R_1^f(x,\mu) = R_0\big(Q^1(x,\mu),\mu\big). \tag{20}$$

Then, for all $\mu \in \mathbb{M}_0$ and for all $x \in [0,1]$,

$$\big|R_1(x,\mu) - R_1^f(x,\mu)\big| < \eta + 3r\delta \tag{21}$$

where η is defined in Assumption FDB1, and δ in Assumption FDB2.

Proof : The result follows immediately from (14) and (18). ■

Remarks :

The quantity δ defined in Assumption FDB2 is clearly less than 1, since all the functions, R_0, Q_0, R^1, Q^1, have values in the interval $[0,1]$. But when the bootstrap DGP is in some sense close to the true DGP, δ should be small. Again, if the quantity r of Assumption 5 is small, then the density $R_0'(x,\mu)$ is close to 1 for all $x \in [0,1]$ and $\mu \in \mathbb{M}_0$, and the value of $R_0(x,\mu)$ is close to x. This is the case if the statistic is approximately pivotal for $\mathbb{M}_0$. If, even for a statistic that is close to being pivotal, it is the case that, for all $\mu \in \mathbb{M}_0$, $R_0(x,\mu)$ is far from x, then one can use the CDF of an arbitrary DGP $\mu_0 \in \mathbb{M}_0$ to transform the statistic so that it has the uniform distribution under μ_0, and a distribution close to uniform for all other DGPs in $\mathbb{M}_0$.

If the realisations of $\tau(\mu,\omega)$ and $\beta(\mu,\omega)$ are respectively t and b, then, by analogy with the double-bootstrap P value (6), the FDB P value is

$$p_2^f \equiv R_1^f\big(R_0(t,b),b\big) = R_0\big(Q^1(R_0(t,b),b),b\big). \tag{22}$$

Before detailing the algorithm for estimating the FDB P value, it may be helpful to develop a link between the finite-sample approach used in this paper and asymptotic arguments like those used by Beran (1988). The next theorem derives the rates at which the small quantities η and δ tend to zero as negative powers of the sample size.

Theorem FDBasy: asymptotic behaviour of the FDB

Under the assumptions of the asymptotic construction used in Beran (1988) for a test statistic that is asymptotically pivotal, with the further assumption that τ_μ and β_μ are asymptotically independent, both η and δ tend to zero at least as fast as $n^{-3/2}$.

Proof : in Appendix A. ■

The algorithm for estimating the FDB P value by simulation is as follows.

Algorithm FDB

1. From the data set under analysis, compute the realised statistic t and the bootstrap DGP b.
2. Draw B bootstrap samples and compute B bootstrap statistics $t_j^* = \tau(b, \omega_j^*)$, $j = 1, \ldots, B$, and B iterated bootstrap DGPs $b_j^* = \beta(b, \omega_j^*)$, for B independent realisations ω_j^* from $(\Omega, \mathcal{F}, P)$.
3. Compute B second-level bootstrap statistics $t_j^{1*} = \tau(b_j^*, \omega_j^{**})$, where the ω_j^{**} are further independent realisations from $(\Omega, \mathcal{F}, P)$, and sort them in increasing order.
4. Compute the estimated first-level bootstrap P value $\hat{p}_1(t, b)$ as the proportion of the t_j^* smaller than t.
5. Obtain the estimate $\hat{Q}^1\big(\hat{p}_1(t, b), b\big)$ as the order statistic of the t_j^{1*} of rank $\lceil B\hat{p}_1(t, b)\rceil$. Denote this estimate as $\hat{Q}^{1*}$.
6. The estimated FDB P value is the proportion of the t_j^* that are smaller than $\hat{Q}^{1*}$.

The above algorithm produces the FDB P value for a test of some specific hypothesis. The FDB can also be used to construct a confidence interval for a scalar parameter, θ say, that is defined within a model $\mathbb{M}$. The idea is conventional: for a confidence interval at confidence level $1 - \alpha$, the FDB hypothesis test is inverted in order to find the limit(s) of the interval. Some care is needed here, since it is no longer convenient to work with a test statistic in approximate P value form.

If a one-sided interval is desired, open to the right say, the lower limit of the interval can be found by inverting a hypothesis test that rejects to the right; similarly for the upper limit of an interval open to the left. For a two-sided interval, the lower (upper) limit is found by inverting a test that rejects to the right (left).

Let the FDB P value used to test the hypothesis that $\theta = \theta_0$, with rejection to the relevant side, be denoted as $p_2^f(\theta_0)$. The corresponding limit of the confidence interval is then the solution of the equation $p_2^f(\theta_0) = \alpha$, or, explicitly, the equation

$$R_0\big(Q^1(p_1(\theta_0), b), b\big) = \alpha, \tag{23}$$

where $p_1(\theta_0)$ is the first-level bootstrap P value.

The following algorithm produces an FDB confidence interval based on an asymptotic t statistic.

Algorithm FDB confidence interval

1. From the data set under analysis, compute a point estimate $\hat{\theta}$ of the parameter of interest, along with a suitable standard error $\hat{\sigma}_\theta$, and a bootstrap DGP b for which the true value of the parameter is $\hat{\theta}$. The test statistic for testing the hypothesis that $\theta = \theta_0$ is $\tau(\theta_0) \equiv (\hat{\theta} - \theta_0)/\hat{\sigma}_\theta$.

2. Draw B bootstrap samples and compute B bootstrap statistics $t_j^* = (\theta^* - \hat{\theta})/(\sigma_\theta^*)_j$, where θ^* and $(\sigma_\theta^*)_j$ are respectively the point estimate and standard error from the j^{th} bootstrap sample, $j = 1, \ldots, B$. Also compute B iterated bootstrap DGPs b_j^*, as in Algorithm FDB.

3. Compute the second-level bootstrap statistics $t_j^{1*} = (\theta^{**} - \theta^*)/(\sigma_\theta^{**})_j$, where θ^{**} and $(\sigma_\theta^{**})_j$ are computed from the sample generated by b_j^*, and sort the t_j^{1*} in increasing order.

4. Consider first the upper limit of the confidence interval, found by inverting a test that rejects to the left. Then $\hat{p}_1(\theta_0) \equiv B^{-1} \sum_{j=1}^{B} \mathrm{I}\big(t_j^* < \tau(\theta_0)\big)$ is the single bootstrap P value for the hypothesis that $\theta = \theta_0$. For the test that rejects to the right, the sign of the inequality is changed.

5. Estimate the quantile $Q^1(p_1(\theta_0), b)$ needed for (23) as the order statistic of the t_j^{1*} with rank $\lceil B\hat{p}_1(\theta_0) \rceil$. Denote this order statistic by $\hat{Q}^1(\hat{p}_1(\theta_0))$.

6. Estimate the left-hand side of (23) by $B^{-1} \sum_{j=1}^{B} \big(t_j^* < \hat{Q}^1(\theta_0)\big)$. This is our estimate $\hat{p}_2^f(\theta_0) = \hat{R}_1^f\big(\hat{p}_1(\theta_0)\big)$.

7. If only the upper limit is needed, the equation $\hat{p}_2^f(\theta_0) = \alpha$ is to be solved. For an equal-tailed two-sided interval, α should be replaced on the right-hand side of the equation by $\alpha/2$.

8. For a lower limit, $\hat{p}_1(\theta_0)$ is redefined in step 4 for a test that rejects to the right. Then in step 5, the quantile to be used has rank $\lceil B\big(1 - \hat{p}_1(\theta_0)\big) \rceil$, and the sign of the inequality in step 6 is changed.

For steps 7 and 8, a direct method is to use a one-dimensional root finder, such as bisection or *regula falsi*. Alternatively, we may take the estimate of equation (23) apart by using a chain of equivalent equations. For an upper limit and a right-hand side of α, first, make the definition $\hat{R}_0(x) = B^{-1} \sum_{j=1}^{B} \mathrm{I}(t_j^* < x)$. Then we have

$$\begin{aligned}
\hat{p}_2^f(\theta) = \alpha &\iff \hat{R}_1^f\big(\hat{p}_1(\theta)\big) = \alpha \iff \hat{R}_0\Big(\hat{Q}^1\big(\hat{R}_0(\tau(\theta))\big)\Big) = \alpha \\
&\iff \hat{Q}^1\big(\hat{R}_0(\tau(\theta))\big) = \hat{Q}_0(\alpha) \iff \hat{R}_0\big(\tau(\theta)\big) = \hat{R}^1\big(\hat{Q}_0(\alpha)\big) \\
&\iff \tau(\theta) = \hat{Q}_0\Big(\hat{R}^1\big(\hat{Q}_0(\alpha)\big)\Big).
\end{aligned}$$

Here the function $\hat{Q}_0(p)$ is defined as the order statistic of rank $\lceil pB \rceil$ of the t_j^*.

Remarks:

The two methods of solving the equation for the limit of a confidence interval do not give numerically identical results in general. This is just a consequence of the discreteness of the estimated bootstrap distribution, which means that there is a range of values of θ for which the P value is equal to any given α.

If an interval is wanted that is symmetric about the point estimate $\hat{\theta}$, then the test statistic $\tau(\theta_0)$ should be replaced by the negative of its absolute value, so that the test rejects to the left. The right-hand side of the equation to be solved is $\alpha/2$.

Chang and Hall (2015) show, using Edgeworth expansions, that the FDB does not improve asymptotically on the single bootstrap for the construction of confidence intervals. Thus Algorithm FDB confidence interval may not be as reliable as Algorithm FDB for P values. If this is a concern, the same reliability as that given by FDB P values can be had by solving, not (23), but rather the equation

$$R_0\big(Q^1(p_1(\theta_0), b(\theta_0)), b(\theta_0)\big) = \alpha,$$

where $b(\theta_0)$ is a bootstrap DGP for which the true value of the parameter θ is θ_0. For full details of this method, see Davidson and MacKinnon (2010).

5. The Fast Triple Bootstrap

In order to study the distribution of the FDB P value (22), we wish to evaluate the expression

$$P(p_2^f \leq x) = \mathrm{E}\Big[\mathrm{I}\Big(R_0\big(Q^1\big(R_0(\tau(\mu,\omega),\beta(\mu,\omega)),\beta(\mu,\omega)\big),\beta(\mu,\omega)\big) \leq x\Big)\Big], \tag{24}$$

which is the probability, under the DGP μ, that the FDB P value is less than x. The inequality that is the argument of the indicator above is equivalent to several other inequalities, as follows:

$$\begin{aligned} & Q^1\big(R_0(\tau(\mu,\omega),\beta(\mu,\omega)),\beta(\mu,\omega)\big) \leq Q_0(x,\beta(\mu,\omega)) \\ \iff & R_0(\tau(\mu,\omega),\beta(\mu,\omega)) \leq R^1\big(Q_0(x,\beta(\mu,\omega)),\beta(\mu,\omega)\big) \\ \iff & \tau(\mu,\omega) \leq Q_0\big(R^1\big(Q_0(x,\beta(\mu,\omega)),\beta(\mu,\omega)\big),\beta(\mu,\omega)\big). \end{aligned}$$

At this point, we can again invoke an approximation that would be exact if τ_μ and β_μ were independent. The final inequality above separates $\tau(\mu,\omega)$ from $\beta(\mu,\omega)$ on the left- and right-hand sides respectively, and so the expectation of the indicator of that inequality is approximated by

$$\mathrm{E}\Big[R_0\big(Q_0\big(R^1(Q_0(x,\beta(\mu,\omega)),\beta(\mu,\omega)),\beta(\mu,\omega)\big),\mu\big)\Big]. \tag{25}$$

We can see from Assumption FDB1 that the difference between (24) and (25) is smaller than η in absolute value.

It is quite possible to estimate (25) for given x and μ by simulation. The function $R_0(\cdot,\mu)$ can be estimated as in (4). But, for the other needed functions, $Q_0(\cdot,\beta(\mu,\omega))$ and $R^1(\cdot,\beta(\mu,\omega))$, which both depend on ω, a second, inner, loop is necessary for each iteration of the main simulation loop. Thus the straightforward estimation procedure is about as computationally intensive as the double bootstrap.

However, we can make a further approximation in the spirit of (18), the second of the approximations that lead to the FDB. As in the proof of Lemma FDB, we can eliminate all mention of the bootstrap DGP $\beta(\mu, \omega)$, the presence of which requires the inner loop. By the same sort of reasoning as used in that proof (note also the approximations in (19)), we can see that the expression (25) differs from

$$R_0\Big(\mathrm{E}\Big[Q_0\big(\mathrm{E}[R^1(\mathrm{E}[Q_0(x, \beta(\mu, \omega))], \beta(\mu, \omega))], \beta(\mu, \omega)\big)\Big], \mu\Big)$$

by a quantity proportional to $r\delta$, where r and δ are defined in Assumption 5 and Assumption FDB2 respectively. Then we may approximate $\mathrm{E}[Q_0(x, \beta(\mu, \omega))]$ by $Q^1(x, \mu)$, giving

$$R_0\Big(\mathrm{E}\Big[Q_0\big(\mathrm{E}[R^1(Q^1(x, \mu), \beta(\mu, \omega))], \beta(\mu, \omega)\big)\Big], \mu\Big) \tag{26}$$

Next, analogously to (15), define the random variable

$$\tau^2(\mu, \omega_1, \omega_2, \omega_3) = \tau(\beta(\beta(\mu, \omega_1), \omega_2), \omega_3),$$

which can be thought of as a realisation of the second-order bootstrap statistic. The CDF of τ^2 under μ, denoted by $R^2(\cdot, \mu)$ is given by

$$\begin{aligned}
&= \mathrm{E}\big[\mathrm{I}\big(\tau^2(\mu, \omega_1, \omega_2, \omega_3) \leq x\big)\big] \\
&= \mathrm{E}\big[\mathrm{E}\big[\mathrm{I}\big(\tau(\beta(\beta(\mu, \omega_1), \omega_2), \omega_3) \leq x\big) \,|\, \mathcal{F}_{12}\big]\big] \\
&= \mathrm{E}\big[R_0\big(x, \beta(\beta(\mu, \omega_1), \omega_2)\big)\big] \\
&= \mathrm{E}\big[R^1\big(x, \beta(\mu, \omega_1)\big)\big],
\end{aligned} \tag{27}$$

where $\mathcal{F}_{12}$ denotes the product sigma-algebra defined on the probability space of ω_1 and ω_2. The third equality follows from the Independence Lemma and the definition of R_0, the fourth from the relation (16). It follows that (26) is equal to

$$R_0\Big(\mathrm{E}\Big[Q_0\big(R^2(Q^1(x, \mu), \mu)], \beta(\mu, \omega)\big)\Big], \mu\Big) \tag{28}$$

Finally, the following expression, in which the bootstrap DGP $\beta(\mu, \cdot)$ no longer appears, differs from (28) by a quantity proportional to $r\delta$:

$$R_0\Big(Q^1\big(R^2(Q^1(x, \mu), \mu), \mu\big), \mu\Big) \tag{29}$$

The result that underlies the fast triple bootstrap is as follows.

Theorem FTB: FTB approximation

Make the definition

$$R_2^f(x, \mu) = R_0\Big(Q^1\big(R^2(Q^1(x, \mu), \mu), \mu\big), \mu\Big). \tag{30}$$

Then, for all $\mu \in \mathbb{M}_0$ and for all $x \in [0,1]$, the CDF of the FDB P value, p_2^f, is approximated by $R_2^f(x,\mu)$, with error equal to η (given in Assumption FDB1) plus a small quantity proportional to $r\delta$.

Proof : The result follows immediately from (25) and (29). ■

The theoretical FDB P value (22) is the approximation $R_1^f(x,\mu)$ to $R_1(x,\mu)$, as defined in (20), evaluated with x set equal to the first-level bootstrap P value, and μ replaced by the bootstrap DGP. The theoretical fast triple bootstrap (FTB) P value is formed analogously from (30) by setting x equal to the FDB P value, and again replacing μ by the (first-level) bootstrap DGP, according to the bootstrap principle. The result is

$$p_3^f(\mu,\omega) \equiv R_0\big(Q^1\big(R^2(Q^1(p_2^f(\mu,\omega),\beta(\mu,\omega)),\beta(\mu,\omega)),,\beta(\mu,\omega)\big),\beta(\mu,\omega)\big), \tag{31}$$

with p_2^f given by (22). The simulation estimate, which must be expressed as a function of the observed statistic t and bootstrap DGP b, is

$$\hat{p}_3^f(t,b) = \hat{R}_0\big(\hat{Q}^1\big(\hat{R}^2(\hat{Q}^1(\hat{p}_2^f(t,b),b),b),b\big),b\big), \tag{32}$$

with $\hat{p}_2^f(t,b)$ given by Algorithm FDB.

Here is the algorithm for the FTB P value.

Algorithm FTB

1. From the data set under analysis, compute the realised statistic t and the bootstrap DGP b.
2. Draw B bootstrap samples and compute B bootstrap statistics $t_j^* = \tau(b,\omega_j^*)$, $j = 1,\ldots,B$, and B iterated bootstrap DGPs $b_j^* = \beta(b,\omega_j^*)$.
3. Compute B second-level bootstrap statistics $t_j^{1*} = \tau(b_j^*,\omega_j^{**})$, and sort them in increasing order. At the same time, compute the corresponding second-level bootstrap DGPs $b_j^{**} = \beta(b_j^*,\omega_j^{**})$.
4. Compute B third-level bootstrap statistics $t_j^{2*} = \tau(b_j^{**},\omega_j^{***})$.
5. Compute the estimated first-level bootstrap P value $\hat{p}_1(t,b)$, as the proportion of the t_j^* smaller than t.
6. Obtain the estimate $\hat{Q}^{1*} \equiv \hat{Q}^1\big(\hat{p}_1(t,b),b\big)$ as the order statistic of the t_j^{1*} of rank $\lceil B\hat{p}_1(t,b)\rceil$.
7. Compute the estimated FDB P value $\hat{p}_2^f(t,b)$ as the proportion of the t_j^* smaller than $\hat{Q}^{1*}$.
8. Compute $\hat{Q}^{1**} \equiv \hat{Q}^1\big(\hat{p}_2^f(t,b),b\big)$ as the order statistic of the t_j^{1*} of rank $\lceil B\hat{p}_2^f(t,b)\rceil$.
9. Compute $\hat{R}^{2*} \equiv \hat{R}^2\big(\hat{Q}^1(\hat{p}_2^f(t,b),b),b\big)$ as the proportion of the t_j^{2*} smaller than $\hat{Q}^{1**}$.

10. Compute $\hat{Q}^{1***} \equiv \hat{Q}^1\big(\hat{R}^2\big(\hat{Q}^1(\hat{p}_2^f(t,b),b),b\big),b\big)$ as the order statistic of the t_j^{1*} of rank $\lceil B\hat{R}^{2*}\rceil$.

11. Compute $\hat{p}_3^f(t,b)$ as the proportion of the t_j^* smaller than $\hat{Q}^{1***}$.

Remark:

It can be hoped that, if the FDB improves on the single bootstrap, the FTB should improve on the FDB to a similar extent, since the approximations used in deriving the theoretical FTB P value (31) hold with the same accuracy as those used to obtain p_2^f. That this is so in circumstances in which the FDB works reasonably well emerges from the simulations reported in Section 7.

6. Fast Higher-Order Bootstraps

The ideas that lead to the FDB and FTB P values can obviously be extended to higher orders. For the FDB, we approximate the distribution of the first-level bootstrap P value $p_1(\mu,\omega)$, and evaluate it at the computed first-level P value $p_1(t,b)$ and the bootstrap DGP b. For the FTB, we approximate the distribution of the FDB P value $p_2^f(\mu,\omega)$ and evaluate it at the computed FDB P value $p_2^f(t,b)$ and b. For a fast quadruple bootstrap, we wish to approximate the distribution of the FTB P value $p_3^f(\mu,\omega)$ and evaluate it at the computed FTB P value $p_3^f(t,b)$ and b. And so on.

The approximate CDFs R_1^f and R_2^f are given explicitly by (20) and (30). We define higher-order approximate CDFs and fast higher-order bootstrap P values recursively, as follows:

$$R_k^f(x,\mu) \approx \mathrm{E}\big[\mathrm{I}\big(p_k^f(\mu,\omega) \le x\big)\big], \text{ and} \tag{33}$$

$$p_{k+1}^f(\mu,\omega) = R_k^f\big(p_k^f(\mu,\omega), b(\mu,\omega)\big), \tag{34}$$

where the exact nature of the approximation in (33) above will now be made explicit.

For the following discussion, the notation t stands for $\tau(\mu,\omega)$, and b stands for $\beta(\mu,\omega)$. We need not distinguish between the random variables and their realisations. Thus $p_2^f(t,b)$ means $p_2^f(\mu,\omega)$, and so forth. By (20), the approximate CDF $R_1^f(x,\mu)$ of $p_1(t,b)$ is $R_0\big(Q^1(x,\mu),\mu\big)\big)$. By (30), $R_2^f(x,\mu)$ is $R_0\big(Q^1\big(R^2(Q^1(x,\mu),\mu),\mu\big),\mu\big)$. As we will see, the pattern for R_k^f is the composition of 2^k functions with R and Q alternating. In order to see how to determine what these functions are, we consider explicitly the case of the fast quadruple bootstrap, which will let us describe the iterative procedure necessary for the explicit expression of R_k^f for general k.

The explicit expression of $p_3^f(t,b)$ is given by (31) and (32), and for present purposes it can be written as

$$p_3^f(t,b) = R_0(Q^1(R^2(Q^1(p_2^f(t,b),b),b),b),b). \tag{35}$$

In order to approximate its distribution, we define the random variable

$$\tau^3(\mu, \omega_1, \omega_2, \omega_3, \omega_4) = \tau(\beta(\beta(\beta(\mu, \omega_1), \omega_2), \omega_3), \omega_4).$$

Its CDF is readily seen to be

$$R^3(x, \mu) = \mathrm{E}\big[R^2(x, \beta(\mu, \omega))\big].$$

The corresponding quantile function, $Q^3(x, \mu)$ is such that

$$\mathrm{E}\big[Q^2(x, \beta(\mu, \omega))\big] = Q^3(x, \mu)$$

with error proportional to $r\delta$. By extension, it is obvious how to define τ^k, R^k and Q^k.

The CDF of $p_3^f(t, b)$ evaluated at x is $\mathrm{E}\big[\mathrm{I}\big(p_3^f(t, b) < x\big)\big]$. By use of (35), we see that the inequality in the indicator here can be expressed as

$$p_2^f(t, b) < R^1(Q^2(R^1(Q_0(x, b), b), b), b). \tag{36}$$

The probability that this inequality is satisfied under μ, conditional on b, can be approximated by the approximate CDF R_2^f of $p_2^f(t, b)$ evaluated at the right-hand side of (36) and μ. By (30), this is

$$R_0(Q^1(R^2(Q^1(R^1(Q^2(R^1(Q_0(x, b), b), b), b), \mu), \mu), \mu), \mu). \tag{37}$$

An argument by now familiar shows that the unconditional expectation of this conditional probability can be approximated by

$$\begin{aligned} R_3^f(x, \mu) \equiv &\mathrm{E}\big[R_0(Q^1(R^2(Q^1(R^1(Q^2(R^1(Q^1(x, \mu), b), b), b), \mu), \mu), \mu), \mu)\big] = \dots \\ &\dots = R_0(Q^1(R^2(Q^1(R^2(Q^3(R^2(Q^1(x, \mu), \mu), \mu), \mu), \mu), \mu), \mu), \mu). \end{aligned} \tag{38}$$

If we compare the sequence of functions in the expression (38) of $R_3^f(x, \mu)$ and the expression (30) of $R_2^f(x, \mu)$, we see that there are exactly twice as many in the former as in the latter. This arises because the inequality (36) has the four functions of p_2^f, and they are then the final four in (37), preceded by the four functions of R_2^f. We may observe that in (37) the composition of the final four functions is the inverse of the composition of the first four. When we get to (38), the indices of the final four functions have all been raised by 1.

By the definition (34), the fourth-level P value $p_4^f(t, b)$ is $R_3^f\big(p_3^f(t, b), b\big)$. When we approximate the CDF of $p_4^f(t, b)$ by the function $R_4^f(x, \mu)$, it is not hard to check that we have, first, the eight functions in (38), followed by the inverse of their composition with indices raised by 1, that is, the composition of the sixteen functions that we write in sequence as follows:

$$R_0 Q^1 R^2 Q^1 R^2 Q^3 R^2 Q^1 R^2 Q^3 R^4 Q^3 R^2 Q^3 R^2 Q^1.$$

Although the way in which we have arrived at this sequence of functions is easy enough to describe, the explicit structure seems not to be expressible in closed form other than by actually working it out.

The following algorithm provides an equivalent but slightly easier way to derive the sequence of functions in R^f_{k+1} from the sequence in R^f_k.

Algorithm FkB:

1. Divide the sequence S_k of the 2^k functions in R^f_k into two sequences A_k and B_k, of length 2^{k-1} each, such that $S_k = A_k B_k$.
2. Obtain the sequence C_k of functions the composition of which is the inverse of the composition of the functions in the sequence B_k.
3. Obtain a new sequence D_k by incrementing the indices of the elements of the sequence C_k by 1.
4. The sequence S_{k+1} used to define R^f_{k+1} is $A_k B_k D_k B_k$, of length 2^{k+1}.

7. Illustrations

In this section, we present the results of various simulation experiments designed to see to what extent the fast double and triple bootstraps can improve the reliability of inference. The first experiments, which deal with a test for a unit root, make use of a parametric bootstrap that makes the distributions of bootstrap statistics absolutely continuous. The other two sets of experiments, one dealing with a test for an ARCH effect, the other a test for serial correlation of the disturbances in a regression model, would normally use resampling bootstraps, which lead to bootstrap statistics with discrete distributions, in violation of Assumption 1. Since resampling is far and away the most frequently used form of bootstrapping in empirical work, it is highly desirable to see if the fast iterated bootstraps suffer noticeably from the discrete distributions induced by resampling. However, for the ARCH test, we also report results obtained by using a smoothed distribution for the bootstrap disturbances, in order that all the assumptions in the earlier part of the paper should hold.

7.1 Testing for a unit root

There are well-known difficulties in testing for a unit root in a series obtained by summing a stationary series that is an MA(1) process with a parameter θ close to -1. Unless special precautions are taken, augmented Dickey-Fuller (ADF) tests can suffer from severe overrejection under the null hypothesis of a unit root, on account of the near cancellation of the unit root by the MA component in the driving stationary series. We may cite Schwert (1989) and Perron and Ng (1996) in this regard.

Over the last decade or so, various bootstrap techniques have been proposed as a means of improving size distortions in the unit-root testing literature. One bootstrap technique designed to deal with autoregressive models is the so called sieve bootstrap first proposed

by Bühlmann (1997). Simulation evidence demonstrates that this bootstrap approach has a certain appeal in reducing size distortions. Here it is appropriate to cite Psaradakis (2001), Chang and Park (2002), and Palm, Smeekes, and Urbain (2008) as evidence of this fact. These papers show the ability of bootstrapped ADF statistics to outperform their asymptotic counterparts. However, despite the ability of the sieve bootstrap to reduce size distortions in certain cases, the gain is really unimpressive in the case considered in our simulations, in which the MA component has a parameter close to -1. In this regard, Richard (2009) applies several variations of the FDB to show that size distortions can be significantly reduced by imposing certain linear restrictions on the truncation parameter of the bootstrap ADF regression and the truncation parameter of the bootstrap sieve.

In Davidson (2010), it is shown that, under the assumption that the MA(1) process has Gaussian innovations, parametric FDB tests can significantly reduce the size distortion of an ADF test. It is argued that bootstrapping the FDB should reduce the distortions still further. This is of course very computationally demanding. In what follows, we show that comparable results may be obtained far more economically with the fast triple bootstrap.

The model studied in this section may be summarised as follows:

$$y_t = \rho y_{t-1} + v_t \tag{39}$$

$$v_t = u_t + \theta u_{t-1}, \qquad u_t \sim \text{NID}(0, \sigma^2), \qquad t = 1, \dots, n. \tag{40}$$

The observed series is y_t, and the null hypothesis of a unit root sets $\rho = 1$. Under that hypothesis, $v_t = \Delta y_t$, where Δ is the first-difference operator. We may write (40) in vector notation using the lag operator L, as follows:

$$\boldsymbol{v} = (1 + \theta \mathrm{L})\boldsymbol{u}, \quad \text{or} \quad \boldsymbol{u} = \boldsymbol{v} - R(\mathrm{L})\boldsymbol{v},$$

where we define $R(\mathrm{L}) = \theta(1 + \theta \mathrm{L})^{-1}\mathrm{L}$. Davidson (2010) demonstrates that θ may be estimated by a two-stage nonlinear least squares regression using the model:

$$\boldsymbol{v} = \phi \boldsymbol{e}_1 + R(\mathrm{L})(\boldsymbol{v} - \phi \boldsymbol{e}_1) + \boldsymbol{u}, \tag{41}$$

where $\boldsymbol{e}_1$ is the vector with first element equal to 1 and all others zero. In the first stage, ϕ is set to zero and a preliminary estimate of θ is obtained. Then an estimate of ϕ is given by $\boldsymbol{s}^\top(1 - R(\mathrm{L}))\boldsymbol{u}/\boldsymbol{s}^\top\boldsymbol{s}$, where $\boldsymbol{s} = (1 - R(\mathrm{L}))\boldsymbol{e}_1$. Finally, this estimate of ϕ is used as a known constant in (41), which is re-estimated to obtain the second-stage estimate of θ.

Testing for a unit root in (39) proceeds by computing an ADF statistic using the ADF regression:

$$\Delta y_t = \beta_0 + \beta_1 y_{t-1} + \sum_{i=1}^{p} \gamma_i \Delta y_{t-i} + \text{residual}. \tag{42}$$

The statistic is the standard t-statistic for the null hypothesis $\beta_1 = 0$ when (42) is estimated by ordinary least squares. The ADF statistic, although easy to compute, has a non-standard asymptotic distribution which is that of a functional of Brownian motion that depends on no nuisance parameters. The ADF statistic is thus an asymptotic pivot.

A suitable bootstrap DGP can be expressed as follows:

$$\boldsymbol{u}^* \sim \mathrm{N}(\mathbf{0}, \mathbf{I}), \quad \boldsymbol{v}^* = (1 + \hat{\theta}\mathrm{L})\boldsymbol{u}^*, \quad \boldsymbol{y}^* = (1 - \mathrm{L})^{-1}\boldsymbol{v}^*, \tag{43}$$

where $\hat{\theta}$ is obtained by the procedure outlined above, with $\boldsymbol{v} = \Delta\boldsymbol{y} = (1 - \mathrm{L})\boldsymbol{y}$. Note that, since the ADF statistic is scale-invariant, we can set $\sigma = 1$. Thus the bootstrap DGP is completely determined by one single parameter, the estimate $\hat{\theta}$.

It is easy to check that Assumptions 1–5 are satisfied by this setup. For Assumption 1, observe that the ADF statistic is a continuous function of the data, and has a continuous density under the null hypothesis. Assumption 2 can be satisfied by using the asymptotic distributions in MacKinnon (1994) and MacKinnon (1996) to convert the raw statistics to approximate P values. Since bootstrap DGPs are completely characterised by the single parameter θ, Lebesgue measure on $[-1, 1]$ serves as the carrier measure required by Assumption 3. The density prescribed by Assumption 4 can be taken to be the joint density of the ADF statistic in approximate P value form and the estimate of θ described above – this density certainly exists since these two quantities are bounded and continuous functions of the data. Assumption 5 is trivially satisfied given the other assumptions; what is not clear just yet is to what extent the quantity r is in fact small. This point, the approximation in Assumption FDB1, and those in Lemma FDB and Theorem FTB are taken up in the simulation study reported in Appendix B. What emerges from the study is that the approximations of the paper that motivate, and, if satisfied, justify the fast iterated bootstraps are mostly satisfactory for the model treated in this subsection. However, when the standard single, double, or triple bootstraps perform poorly, especially as θ approaches -1, then the performance of the fast iterated bootstraps, which approximate the standard iterated bootstraps quite well, cannot be expected to be any better.

In the simulation experiments we now discuss, the data were generated by the model defined by (39) and (40), with $\rho = 1$ and for various values of θ close to -1. For any given estimate $\hat{\theta}$, the bootstrap DGP (43) was used. Second-level bootstrap DGPs were formulated by generating bootstrap data with $\hat{\theta}$ in (40), and then obtaining the estimate of θ for these data. This estimate, which we denote $\hat{\theta}^*$, then characterises a second-level bootstrap DGP, which can subsequently be used to generate second-level bootstrap data, used to obtain an estimate $\hat{\theta}^{**}$, which characterises a third-level bootstrap DGP. This experimental design is the same as the one used in Davidson (2010).

Figures 1 and 2 below show the errors in rejection probability (ERP), that is, the difference between the experimentally observed rejection rate and the nominal level α, of the ordinary single bootstrap, the FDB, and the FTB for varying values of the MA(1) parameter θ, for sample sizes $n = 50$ and 100, and for all nominal levels $0 \leq \alpha \leq 1$. Although for practical purposes one need not consider values of α greater than 0.1, using all possible values allows a clearer view of how the distributions of the various bootstrap P values differ from the ideal U(0,1) distribution. All graphs in Figure 1 were computed with $N = 25{,}000$ replications, $B = 4{,}999$ bootstrap repetitions, and ADF truncation parameter $p = 12$. The time required to complete the simulations using a computer cluster with 54 nodes with 4 CPUs each was roughly 16 hours and 22 hours for sample sizes $n = 50$ and $n = 100$ respectively.

We see that the FTB P values tend to have lower ERPs than their FDB and single bootstrap (BS) counterparts. As expected, the size correction is not very impressive for $\theta = -0.99$, but as θ increases away from $\theta = -1$ improvements can be drastic. What is interesting however is that when θ is very close to -1 as in Figure 1, the FTB ERP is slightly worse than the FDB ERP, which in turn is slightly worse than the single bootstrap ERP in the region of the nominal size which is most relevant to practitioners, namely $0 \leq \alpha \leq 0.10$; see Figure 1a. However, the effects are very small, and may be no more than artefacts due to simulation randomness.

An unexpected and somewhat strange feature of the results shown in the figures is that the distortions for $n = 50$ are often smaller than those for $n = 100$. This is a manifestation of the failure of any kind of inference in the close vicinity of $\theta = -1$. The distortions appear to be magnified when the sample size increases.

Figures 3 and 4 are like Figures 1 and 2, but they omit the case of the single bootstrap, and include results for the standard double and triple bootstraps. They show that the behaviour of the fast iterated bootstraps is very similar to that of the standard iterated bootstraps. The latter are slightly less distorted than the former, but the fast triple bootstrap usually outperforms the standard double bootstrap. Some caution is required in comparing the experimental results in Figures 3 and 4, because, for reasons of computational feasibility, the results for the standard iterated bootstraps were obtained with $N = 12{,}500$, $B = 99$, $B_1 = 119$, and $B_2 = 129$, in order to complete the simulations in a reasonable time frame. Even so, using all 216 CPUs of the cluster, it took over 24 hours to complete a simulation for a single value of θ with a sample size $n = 50$. But, even allowing for non-negligible simulation randomness, the ordering of the various procedures is clearly established.

Before ending this section, we must remark that the model considered here is much too restrictive to have any practical interest. We have studied it as a computationally feasible test bed for the fast and standard iterated bootstraps, and to demonstrate that going as far as the fast triple bootstrap does indeed have the possibility of giving improved reliability over the single and fast double bootstraps, although, since the η of Assumption FDB1 is not small, neither the fast nor the standard iterated bootstraps yield anything like perfect inference. An interesting piece of future work would be to see to what extent the results reported here agree, qualitatively at least, with using the fast double and triple bootstraps along with the non-parametric resampling bootstrap commonly used.

This model was used in a quite different context in Davidson (2016a) to investigate standard bootstrap iteration by making use of a discretised version of the model. It is shown there that the sequence of iterated bootstrap P values does converge eventually, but not necessarily to a limit useful for inference, at least in the neighbourhood of $\theta = -1$.

7.2 A test for ARCH

In Davidson and MacKinnon (2007), one of the examples used to show how the FDB improves on the single bootstrap is a test for ARCH disturbances in a linear regression model. Since the seminal work of Engle (1982), it has been recognized that serial dependence in

the variance of the disturbances of regression models using time-series data is a very common phenomenon. It is therefore usually advisable to test for the presence of such serial dependence.

Consider the linear regression model

$$\begin{aligned} y_t &= \boldsymbol{X}_t\boldsymbol{\beta} + u_t, \quad u_t = \sigma_t\varepsilon_t, \quad t = 1, \ldots, n, \\ \sigma_t^2 &= \sigma^2 + \gamma u_{t-1}^2 + \delta\sigma_{t-1}^2, \quad \varepsilon_t \sim \text{IID}(0,1). \end{aligned} \tag{44}$$

The disturbances of this model follow the GARCH(1,1) process introduced by Bollerslev (1986). The easiest way to test the null hypothesis that the u_t are IID in the model (44) is to run the regression

$$\hat{u}_t^2 = b_0 + b_1\hat{u}_{t-1}^2 + \text{residual}, \tag{45}$$

where $\hat{u}_t$ is the t^{th} residual from an OLS regression of y_t on $\boldsymbol{X}_t$. The null hypothesis that $\gamma = \delta = 0$ can be tested by testing the hypothesis that $b_1 = 0$. Besides the ordinary t statistic for b_1, a commonly used statistic is n times the centred R^2 of the regression, which has a limiting asymptotic distribution of χ_1^2 under the null hypothesis.

Since in general one is unwilling to make any restrictive assumptions about the distribution of the ε_t, a resampling bootstrap seems the best choice. As mentioned above, this is in violation of Assumption 1. Conventional resampling gives rise to a discrete distribution, although, for samples of reasonable size, it may be close enough to being continuous for the discreteness not to matter. This idea is tested by our also presenting results where the discrete distribution is smoothed, so that Assumption 1 is satisfied. The other assumptions, dealing with measurability, continuity, *etc.*, are satisfied because the test statistic, and any of the bootstrap DGPs, discrete or smoothed, are continuously differentiable functions of the data.

The experimental design is copied from Davidson and MacKinnon (2007). In all cases, $\boldsymbol{X}_t$ consists of a constant and two independent, standard normal random variates, since changing the number of regressors has only a modest effect on the finite-sample behaviour of the tests. The sample size takes on the values 40, 80, and 160, a small subset of the set of sample sizes studied by Davidson and MacKinnon. In order to have non-negligible ERPs, the ε_t are drawn from the χ_2^2 distribution, subsequently centred and rescaled to have variance 1. Without loss of generality, we set $\boldsymbol{\beta} = \boldsymbol{0}$ and $\sigma^2 = 1$, since the test statistic is invariant to changes in the values of these parameters.

The invariance means that we can use a straightforward, possibly smoothed, resampling bootstrap DGP, with the y_t^* in a bootstrap sample being IID drawings from the empirical distribution of the y_t. For iterated bootstraps, y_t^{**} is resampled from the y_t^*, and y_t^{***} is resampled from the y_t^{**}.

The experiments consisted of $N = 100{,}000$ replications with $B = 399$ bootstrap repetitions each. For each replication, P values were computed for the asymptotic test, the single bootstrap, the FDB, and the FTB. The results under the null are displayed in Figure 5a for a conventional resampling bootstrap. Observe that there is little difference in the ERPs

for $n = 80$ and $n = 160$. The ordering of the four procedures is quite clear. The single bootstrap improves considerably on the asymptotic test based on the χ^2_1 distribution, the FDB again considerably on the single bootstrap, and the FTB has no significant distortion at all for $n = 80$ and $n = 160$, and very little for $n = 40$.

In Figure 5b the results for $n = 40$ are shown again, alongside the results for the same setup, but with the discrete bootstrap distribution smoothed by use of a cumulative Epanechnikov kernel. The continuous curves are for the resampling bootstrap, the dashed curves for the smoothed bootstrap. The smoothed version of the empirical distribution function of the sample y_t, $t = 1, \dots, n$, is

$$F_{\text{sm}}(y) = \frac{1}{n} \sum_{t=1}^{n} K\big(h^{-1}(y_t - y)\big),$$

where n is the sample size, h is the bandwidth, chosen here rather arbitrarily as 0.05, and the cumulative kernel is

$$K(z) = \text{I}\big(|z| \leq \sqrt{5}\big)\big(\frac{3}{4\sqrt{5}}(z - z^3/15) + \frac{1}{2}\big) + \text{I}\big(z > \sqrt{5}\big).$$

A conventional resample is made, with typical element y^*_t, to which is added some resampling "noise", to yield $y^{\text{sm}}_t = y^*_t + hK^{-1}(p)$. For reference, we note that

$$K^{-1}(p) = 2\sqrt{5}\cos\Big(\tfrac{1}{3}\big(2\pi - \cos^{-1}(1 - 2p)\big)\Big).$$

It is quite evident from the Figure that there are no significant differences between the errors in rejection probability with the conventional and smoothed resampling bootstraps. For this reason, no results are provided for the smoothed bootstrap for the larger sample sizes.

One might wonder if the excellent behaviour of the FTB is bought at the cost of diminished power. That this is not the case is demonstrated by the experiments the results of which are shown in Figure 6. The simulated data were generated by the DGP of model (44) with $\sigma^2 = 1$, $\gamma = 0.3$, and $\delta = 0.3$. These values were chosen in order to have some power for $n = 40$ and not to have so much power for $n = 160$ for an ordering of the results to be unclear. In this figure, the experimental rejection rate is plotted as a function of the *nominal* level of the test; see Horowitz and Savin (2000) and Davidson and MacKinnon (2006) for discussion of why this makes more sense than attempting any sort of "size adjustment".

The under-rejection of the asymptotic test under the null is carried over under the DGP studied here, to the extent that, for $n = 40$, the rejection rate of the test is smaller than the nominal level for some values of the latter. This is no longer true for the larger values of n. The three bootstrap tests have very similar rejection rates, always greater than that of the asymptotic test, and with the FDB and FTB having slightly greater power than the single bootstrap test.

7.3 A test for serial correlation

Another of the examples of the good performance of the FDB found in Davidson and MacKinnon (2007) is given by the Durbin-Godfrey test for serial correlation of the disturbances in a linear regression model. The test was introduced in Durbin (1970) and Godfrey (1978). The model that serves as the alternative hypothesis for the test is the linear regression model

$$y_t = \boldsymbol{X}_t\boldsymbol{\beta} + \gamma y_{t-1} + u_t, \quad u_t = \rho u_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim \text{IID}(0, \sigma^2), \quad t = 1, \dots, n, \tag{46}$$

where $\boldsymbol{X}_t$ is a $1 \times k$ vector of observations on exogenous variables. The null hypothesis is that $\rho = 0$. Let the OLS residuals from running regression (46) be denoted $\hat{u}_t$. Then the Durbin-Godfrey (DG) test statistic is the t statistic for the coefficient of $\hat{u}_{t-1}$ in a regression of $\hat{u}_t$ on $\boldsymbol{X}_t$, y_{t-1}, and $\hat{u}_{t-1}$. It is asymptotically distributed as $\text{N}(0, 1)$ under the null hypothesis. Since this test can either overreject or underreject in finite samples, it is natural to use the bootstrap in an effort to improve its finite-sample properties.

For the bootstrap DGP, from running regression (46), we obtain estimates $\hat{\boldsymbol{\beta}}$, $\hat{\gamma}$, as well as the residuals $\hat{u}_t$. The semiparametric bootstrap DGP can be written as

$$y_t^* = \boldsymbol{X}_t\hat{\boldsymbol{\beta}} + \hat{\gamma} y_{t-1}^* + u_t^*, \tag{47}$$

where the u_t^* are obtained by resampling the residuals rescaled as $(n/(n-k-1))^{1/2}\hat{u}_t$. The initial value y_0^* is set equal to the actual pre-sample value y_0.

Our experimental design is similar to that in Davidson and MacKinnon (2007). We set $k = 6$, with the first regressor a constant, and the remaining five generated by independent, stationary AR(1) processes with normal innovations and parameter $\rho_x = -0.8$. The disturbances ε_t are normally distributed with $\sigma = 10$. We put $\boldsymbol{\beta} = \mathbf{0}$ and $\gamma = 0.75$. We look at sample sizes $n = 20$ and 40 under the null, and also at $n = 56$ in our study of power. These choices are sufficient for us to distinguish clearly the behaviour of the various testing procedures: asymptotic, single bootstrap, FDB, and FTB. As before, we used 10,000 replications each involving 399 bootstrap repetitions. Since it was apparent in our study of the test for ARCH that smoothing the bootstrap distribution has a quite negligible effect, no smoothing was done here.

Figure 7 shows the ERPs under the null. For $n = 40$, we do not show the ERP of the asymptotic test, because it is so great that including it in the graph would make the ERPs of the other tests indistinguishable. All the bootstrap procedures are much less distorted than the asymptotic test, and they have, once again, the same ordering as before, with the FTB the least distorted. Even for as small a sample size as 40, its observed rejection rate is never different from the nominal level by more than 0.01.

Power is illustrated in Figure 8. The data for these experiments were generated with the autocorrelation parameter ρ in (46) equal to 0.5, $\sigma = 1$, and the parameter ρ_x used in generating the regressors equal to 0.8. These values were chosen on the basis of the results in Davidson and MacKinnon (2007), where they give rise to significant differences in the

power of the single bootstrap and the FDB. For $n = 20$, there is no visible power at all; in fact there is some slight underrejection by all the tests. With $n = 40$, the asymptotic test rejects slightly more often than the bootstrap tests, and, for $n = 56$, all tests have very similar rejection rates. Once again, the superior performance of the fast iterated bootstrap tests does not entail reduced power.

8. Conclusions

The approximations that led to the fast double bootstrap have been extended not only to a fast triple bootstrap but to arbitrary levels of bootstrap iteration. Algorithms have been given for the implementation of the fast double and fast triple bootstraps, along with an algorithm for generating the expressions to be estimated by simulation for higher-level iterated bootstraps.

The approximations underlying the fast iterated bootstraps have been justified here only under two assumptions. The first is that the statistic being bootstrapped is approximately independent of the bootstrap DGP, for instance, if asymptotic independence holds. The second is that all bootstrap DGPs generate statistics of which the distributions are absolutely continuous with respect to Lebesgue measure on the real line. Although the second assumption is not satisfied if a resampling bootstrap is used, it is clear from one of our simulations that, even for a small sample size, rejection probabilities of bootstrap tests do not differ significantly when conventional, discrete, resampling is used rather than smoothed resampling of a sort that satisfies our second assumption. It was seen in two of the illustrative examples that going from the standard bootstrap to the fast double bootstrap and on to the fast triple bootstrap leads to progressively less size distortion.

In the experiments with a unit root test, a parametric bootstrap is used, so that the second assumption is satisfied, but the first decidedly is not. A detailed investigation of the extent to which the assumptions that lead to the FDB and FTB are violated was undertaken, from which it emerged that, even in the absence of approximate independence, the approximations underlying the fast procedures are quite good, except in a few cases. These cases are characterised by very poor performance of the ordinary, single, bootstrap. For the unit-root model considered here, at least, the fast procedures lose little accuracy relative to the standard iterated bootstraps, but cannot improve on the latter when their performance is poor.

When the assumption of approximate independence does not hold, the bootstrap can be expected to perform badly; see Davidson (2016b). Quite generally, then, the fast iterated bootstraps improve reliability of inference when the bootstrap is already reasonably reliable, but whether they do so in other cases is undetermined in general. As we saw in the example with a unit-root test, it is possible in unfavourable circumstances for the single bootstrap to out-perform both the FDB and the FTB. Further, since the experiments reported in Figures 3 and 4 show that the FDB and FTB perform similarly to the standard double and triple bootstraps, it appears that *any* sort of bootstrap iteration may yield very little advantage over the single bootstrap when the latter performs badly.

Overall, our experiments show that the FDB and FTB suffer from only a little more distortion than their standard counterparts. Further, the resampling examples demonstrate that improved behaviour under the null is not achieved at the cost of reduced power.

The few experiments we report about the standard triple bootstrap required enormous computing resources. With current technology, and indeed with technology likely to be available in the near future, experiments on standard iterated bootstraps beyond the triple would consume computing power well beyond that accessible to academic researchers. The fast variants are much less computationally intensive, and their demands increase only linearly with the level of iteration, while those of the standard iterated bootstraps increase exponentially. For a single bootstrap P value, one has to evaluate $1+B$ statistics and just one bootstrap DGP. For the level-k fast iterated bootstrap, one needs $1+kB$ statistics, and $1+(k-1)B$ DGPs. But for the level-k standard iterated bootstrap, if for simplicity we do not use different numbers of bootstraps at each level, the number of statistics needed is $(B^{k+1}-1)/(B-1)$ and the number of DGPs is $(B^k-1)/(B-1)$.

Interesting theoretical questions remain, to do with the convergence or otherwise of the sequence of iterated bootstrap P values, fast and standard. It would be good to have conditions that guarantee convergence or non-convergence. Where convergence does occur, is the limiting distribution the uniform distribution on $[0,1]$? With resampling, we know that it cannot be, since, with repeated resampling, eventually a stage is reached in which only one element is resampled. With absolutely continuous distributions, to what extent do the answers to these convergence questions depend on the assumptions, namely that η and δ are small, that underlie the FDB? We hope to clarify these and other points in future work.

Appendix A

Proof of Lemma FDB:

By the mean-value theorem, for any $x \in [0,1]$, there exists a random quantity ρ_1 such that

$$R_0(Q_0(x,\beta(\mu,\omega),\mu)) - R_0(Q_0(x,\mu),\mu) = (1+\rho_1)(Q_0(x,\beta(\mu,\omega)) - Q_0(x,\mu))$$

where, by Assumption 5, $|\rho_1| < r$ almost surely. On taking the expectation, and on noting that $R_0(Q_0(x,\mu),\mu) = x$, we see that

$$\mathrm{E}[R_0(Q_0(x,\beta(\mu,\omega)),\mu)] = x + (\mathrm{E}[(1+\rho_1)Q_0(x,\beta(\mu,\omega)] - Q_0(x,\mu)). \tag{A.1}$$

But

$$R_0(\mathrm{E}[Q_0(x,\beta(\mu,\omega))],\mu) = x + (1+\rho_2)(\mathrm{E}[Q_0(x,\beta(x,\mu))] - Q_0(x,\mu)) \tag{A.2}$$

with $|\rho_2| < r$. By subtracting (A.2) from (A.1), we find that

$$\mathrm{E}[R_0(Q_0(x,\beta(\mu,\omega)),\mu)] - R_0(\mathrm{E}[Q_0(x,\beta(\mu,\omega))],\mu) = (\mathrm{E}(\rho_1) - \rho_2)(\mathrm{E}[Q_0(x,\beta(x,\mu))] - Q_0(x,\mu)).$$

Then, by Assumption FDB2, $|\mathrm{E}[Q_0(x,\beta(\mu,\omega))] - Q_0(x,\mu)| < \delta$, so that

$$\mathrm{E}[R_0(Q_0(x,\beta(\mu,\omega)),\mu)] - R_0(\mathrm{E}[Q_0(x,\beta(\mu,\omega))],\mu) < 2r\delta. \tag{A.3}$$

Next, since $R^1(Q^1(x,\mu),\mu) = x$, it follows from (16) that $\mathrm{E}(R_0(Q^1(x,\mu),\beta(\mu,\omega))) = x$. Thus there exists ρ_3, with $|\rho_3| < r$ almost surely such that

$$\begin{aligned}
0 &= \mathrm{E}[R_0(Q^1(x,\mu),\beta(\mu,\omega)) - R_0(Q_0(x,\beta(\mu,\omega)),\beta(\mu,\omega))] \\
&= \mathrm{E}[(1+\rho_3)(Q^1(x,\mu) - Q_0(x,\beta(\mu,\omega))] \\
&= Q^1(x,\mu) - \mathrm{E}[(Q_0(x,\beta(\mu,\omega)] + \mathrm{E}[\rho_3(Q^1(x,\mu) - Q_0(x,\beta(\mu,\omega))],
\end{aligned}$$

whence, by Assumption FDB2 again, $Q^1(x,\mu) - \mathrm{E}[Q_0(x,\beta(\mu,\omega))] < r\delta$. But then

$$R_0(\mathrm{E}[Q_0(x,\beta(\mu,\omega))],\mu) - R_0(Q^1(x,\mu),\mu) = (1+\rho_3)(\mathrm{E}[Q_0(x,\beta(\mu,\omega))] - Q^1(x,\mu)) < 2r\delta. \tag{A.4}$$

Combining (A.3) and (A.4) yields the result of the lemma. ■

Proof of Theorem FDBasy:

To analyse the expression (12) for $R_1(x,\mu)$ further, we follow closely an argument given in section 2 of Davidson and MacKinnon (2006). First fix x, and then define two scalar random variables, $\rho(\mu,\omega)$ and $\kappa(\mu,\omega)$, the first a differentiable deterministic function of $\tau(\mu,\omega)$, the second of $\beta(\mu,\omega)$, as follows:

$$\rho(\mu,\omega) = R_0(\tau(\mu,\omega),\mu) \quad \text{and} \quad \kappa(\mu,\omega) = R_0(Q_0(x,\beta(\mu,\omega)),\mu) - x. \tag{A.5}$$

Then, by (12),

$$\begin{aligned}
R_1(x,\mu) &= \mathrm{E}[\mathrm{I}(\tau(\mu,\omega) \le Q_0(x,\beta(\mu,\omega))] = \mathrm{E}[\mathrm{I}(R_0(\tau(\mu,\omega),\mu) \le R_0(Q_0(x,\beta(\mu,\omega)),\mu))] \\
&= \mathrm{E}[\mathrm{I}(\rho(\mu,\omega) \le x + \kappa(\mu,\omega))]
\end{aligned} \tag{A.6}$$

Consider the joint distribution of $\rho(\mu,\omega)$ and $\kappa(\mu,\omega)$. By Assumptions 1 and 2, the marginal distribution of $\rho(\mu,\omega)$ is uniform on $[0,1]$. Use of the measure P_μ defined in (11) lets us define

$$G_\mu(p,q) = P\{(\rho(\mu,\omega) \le p) \wedge (\kappa(\mu,\omega) \le q)\}$$

and the corresponding density $g_\mu(p,q)$, of which the existence is guaranteed by Assumption 4. Since the marginal density of ρ is equal to 1 for $0 \le \rho \le 1$, the joint density $g_\mu(p,q)$ is equal to the conditional density $g_\mu(q \mid p)$. The integral of $g_\mu(q \mid p)$ with respect to q, denoted $G_\mu(q \mid p)$, is then the CDF of κ conditional on $\rho = p$.

If we evaluate the unconditional expectation (A.6) as the expectation of the expectation conditional on ρ, we see that $R_1(x,\mu)$ is

$$\begin{aligned}
\mathrm{E}\{\mathrm{E}[\mathrm{I}(\kappa \ge \rho - x) \mid \rho]\} &= 1 - \mathrm{E}\{\mathrm{E}[\mathrm{I}(\kappa < \rho - x) \mid \rho]\} \\
&= 1 - \mathrm{E}\{G_\mu(\rho - x \mid \rho)\} = 1 - \int_0^1 G_\mu(p - x \mid p)\,\mathrm{d}p.
\end{aligned}$$

Integration by parts then gives

$$R_1(x,\mu) = 1 - [p\; G_\mu(p - x \mid p)]_0^1 + \int_0^1 p\;\mathrm{d}G_\mu(p - x \mid p) = \int_0^1 p\;\mathrm{d}G_\mu(p - x \mid p),$$

and, by changing the integration variable by $p = x + q$, we see that

$$R_1(x,\mu) = \int_{-x}^{1-x} (x+q)\, \mathrm{d}G_\mu(q \mid x+q) = x + \int_{-x}^{1-x} q\, \mathrm{d}G_\mu(q \mid x+q). \tag{A.7}$$

On the other hand, $R_0(Q_0(x,\beta(\mu,\omega)),\mu) = x + \kappa$, and so

$$\mathrm{E}[R_0(Q_0(x,\beta(\mu,\omega)),\mu)] = x + \mathrm{E}(\kappa) = x + \int_{-x}^{1-x} q\, \mathrm{d}G_\mu(q), \tag{A.8}$$

where $G_\mu(q) = G_\mu(1,q)$ is the marginal CDF of κ.

It follows from the argument in section 3.1 of Beran (1988), Case 1, for which the limit distribution of the statistic τ as the sample size $n \to \infty$ is the same for all $\mu \in \mathbb{M}_0$, that, in our notation, $R_0(x,\mu) - R_0(x,\beta(\mu,\omega))$ is $O_p(n^{-1})$ or of lower order, presumably uniformly in $x \in [0,1]$ and $\mu \in \mathbb{M}_0$. By replacing x in this relation by $Q_0(x,\beta(\mu,\omega))$, we see that

$$R_0(Q_0(x,\beta(\mu,\omega)),\mu) - x = \kappa(\mu,\omega) = O_p(n^{-1}).$$

The integrals in the rightmost members of (A.7) and (A.8) are conditional and unconditional expectations, respectively, of κ, and so they too are $O_p(n^{-1})$ or of lower order. The difference between them is therefore also quite generally of that order as well. It follows from Assumption FDB1 that we may choose η of the same order.

In fact, the unconditional expectation of κ, that is, the integral in (A.8), is of lower order than κ itself, as shown in section 4 of Davidson and MacKinnon (1999). The reasoning is simple enough: from the definition (A.5) of κ, it is clear that $x + \kappa$ is the estimator of $R(Q_0(x,\mu),\mu) = x$. If this estimator is consistent, as it certainly is under Beran's assumptions, then the expectation of κ tends to zero as $n \to \infty$ faster than κ itself.

If $\tau(\mu,\omega)$ and $\beta(\mu,\omega)$ are asymptotically independent, as is supposed here, then $\rho(\mu,\omega)$ and $\kappa(\mu,\omega)$ are asymptotically independent as well. This implies that $G_\mu(q \mid x+q) - G_\mu(q)$ tends to zero as $n \to \infty$. In Beran's scheme, this means that $G_\mu(q \mid x+q) = G_\mu(q)(1 + O(n^{-1/2}))$. It follows that the difference between the rightmost members of (A.7) and (A.8) is of the order of $n^{-1/2}\kappa$ so that η can be chosen to be $O_p(n^{-3/2})$ or of lower order.

Regarding δ, we have a similar argument. From Assumption FDB2, δ bounds the absolute value of $R^1(x,\mu) - R_0(x,\mu)$ uniformly in $x \in [0,1]$ and $\mu \in \mathbb{M}_0$. Now by the definition (16), we have for any given x that

$$R^1(x,\mu) - R_0(x,\mu) = \mathrm{E}[R_0(x,\beta(\mu,\omega)) - R_0(x,\mu)].$$

By replacing x above by $Q_0(x,\beta(\mu,\omega))$, we see that the above expression is equal to

$$\mathrm{E}[x - R_0(Q_0(x,\beta(\mu,\omega)),\mu)] = -E(\kappa),$$

and we have just seen that $\mathrm{E}(\kappa)$ is of order $n^{-3/2}$ or lower. ■

Appendix B

The purpose of this Appendix is to see how good the various assumptions used to justify the fast double and triple bootstraps for the unit-root model are. We consider four cases: sample sizes $n = 50$ and 100, and the MA parameter $\theta = -0.90$ and -0.99. For Assumption FDB1, we compare $R_1(x, \mu)$ and $\mathrm{E}[R_0(Q_0(x, \beta(x, \mu)), \mu)]$. For the approximation in Lemma FDB, which relies on Assumption FDB2, we compare $\mathrm{E}[R_0(Q_0(x, \beta(x, \mu)), \mu)]$ and $R_0(Q^1(x, \mu), \mu)$. It is of course infeasible to attempt such a comparison analytically, but it is quite straightforward to do so by use of a simulation experiment, in which all three of the quantities concerned are estimated by simulation. The design of the experiment is as follows. For each choice of n and θ:

1. Draw $N = 10{,}000$ IID samples from the DGP μ given by (39) and (40).
2. For sample i, $i = 1, \ldots, N$, compute the test statistic using the test regression (42), and store these IID realisations $\tau(\mu, \omega_i)$. Further, compute the estimate $\hat{\theta}_i$, which defines a bootstrap DGP $\beta(\mu, \omega_i)$ for the same ω_i, and draw one second-level statistic from $\beta(\mu, \omega_i)$. Store these IID realisations $\tau(\beta(\mu, \omega_i), \omega_{ii})$, and sort them in increasing order.
3. For $x = 0.01$ by steps of 0.01 until $x = 0.50$, estimate the quantile $Q^1(x, \mu)$ by the $\lceil Nx \rceil$ order statistic of the $\tau(\beta(\mu, \omega_i), \omega_{ii})$; call it $q^1(x)$.
4. Compute the proportion of the $\tau(\mu, \omega_i)$, $i = 1, \ldots, N$, that are less than $q^1(x)$. This estimates $R_0(Q^1(x, \mu), \mu)$.
5. Draw $N_1 = 1{,}000$ samples from the DGP $\beta(\mu, \omega_i)$, for each compute a test statistic $\tau(\beta(\mu, \omega_i), \omega_{ik})$, $k = 1, \ldots, N_1$, and sort these statistics in increasing order.
6. For $x = 0.01$ by steps of 0.01 until $x = 0.50$, estimate the quantile $Q_0\big(x, \beta(\mu, \omega_i)\big)$ by the $\lceil N_1 x \rceil$ order statistic of the $\tau(\beta(\mu, \omega_i), \omega_{ik})$; call it $q_i(x)$.
7. Compute the proportion of the $\tau(\mu, \omega_j)$, $j = 1, \ldots, N$, that are less than $q_i(x)$; call this $r_i(x)$. This estimates $R_0(Q_0(x, \beta(\mu, \omega_i)), \mu)$.
8. Estimate $\mathrm{E}[R_0(Q_0(x, \beta(\mu, \omega)), \mu)]$ by the average of the $r_j(x)$.

Figure 9 displays graphs of the differences between x and the estimated $\mathrm{E}[R_0(Q_0(x, \beta(\mu, \omega)), \mu)]$, $R_0(Q^1(x, \mu), \mu)$, and $R_1(x, \mu)$ (computed earlier as part of the simulation experiments with results in Figures 1–4) for x running from 0 to 0.5, for the four cases considered. In these graphs, the curves labelled "approx1" plot the estimates of $\mathrm{E}[R_0(Q_0(x, \beta(\mu, \omega)), \mu)] - x$, while those labelled "approx2" plot the estimates of $R_0(Q^1(x, \mu), \mu) - x$. Both should be close to the corresponding curve labelled "bootstrap", which plots estimates of the bootstrap discrepancy $R_1(x, \mu) - x$. Only in the case with $n = 100$, $\theta = -0.90$, is there any sign that the approximations are other than very good indeed. It can be seen in the bottom right panel of Figure 1 that that is the case for which the distortions of all of the tests, asymptotic and bootstrap, are greatest. Note also that it is the "approx1" curve that is different from the other two: this implies that the quantity η in Assumption FDB1 is greater than in any of the other cases considered, which in turn implies that the assumption based on approximate independence of the statistic and the bootstrap DGP is substantially violated.

According to Theorem FTB, the CDF of the FDB P value, p_2^f, evaluated at $x \in [0, 1]$, is approximated by the expression

$$R\Big(Q^1(R^2(Q^1(x, \mu), \mu), \mu,), \mu\Big). \tag{30}$$

This expression can be estimated by simulation, according to the following algorithm, for a given choice of n and θ.

1. Draw $N = 10{,}000$ IID samples from the DGP μ given by (39) and (40).
2. For sample i, $i = 1, \ldots, N$, compute the test statistic using the test regression (42), and store these IID realisations $\tau(\mu, \omega_i)$. Compute the estimate $\hat{\theta}_i$, which defines a bootstrap DGP $\beta(\mu, \omega_i)$ for the same ω_i, and draw one second-level statistic from $\beta(\mu, \omega_i)$. Store these IID realisations $\tau(\beta(\mu, \omega_i), \omega_{ii})$, and sort them in increasing order. Also compute an estimate of θ for the DGP $\beta(\mu, \omega_i)$, thus defining the bootstrap DGP $\beta(\beta(\mu, \omega_i), \omega_{ii})$, from which draw a statistic $\tau(\beta(\beta(\mu, \omega_i), \omega_{ii}), \omega_{iii})$. Store these realisations.
3. For $x = 0.01$ by steps of 0.01 until 0.50, estimate $Q^1(x, \mu)$ as the $\lceil Nx \rceil$ order statistic of the $\tau(\beta(\mu, \omega_i), \omega_{ii})$; call it $q^1(x)$.
4. Estimate $R^2(Q^1(x.\mu), \mu)$ by the proportion of the $\tau(\beta(\beta(\mu, \omega_i), \omega_{ii}), \omega_{iii})$ less than $q^1(x)$; call it $r^2(x)$.
5. Estimate $Q^1(R^2(Q^1(x, \mu), \mu), \mu)$ as the $\lceil Nr^2(x) \rceil$ order statistic of the $\tau(\beta(\mu, \omega_i), \omega_{ii})$; call it $q^{11}(x)$.
6. Estimate $R(Q^1(R^2(Q^1(x, \mu), \mu), \mu,), \mu)$ by the proportion of the $\tau(\mu, \omega_i)$ less than $q^{11}(x)$.

In Figure 10, comparisons are displayed graphically of the discrepancy associated with the FDB P value and the approximation (30). For $\theta = -0.90$, the approximation is virtually perfect. For $\theta = -0.99$, it is still very good for levels in the range of interest when $n = 50$, but is much worse when $n = 100$. Once again, the poor performance of the bootstrap near $\theta = -1$ increases with the sample size. Overall, however, it is clear why the FTB does so well in most cases, as it provides a really good approximation to the true CDF of the FDB P value.

Acknowledgements

This research was supported by the Canada Research Chair program (Chair in Economics, McGill University), and by grants from the Social Sciences and Humanities Research Council of Canada and the Fonds de Recherche du Québec – Société et Culture. We are grateful to seminar participants at the Third French Econometric Conference, the Singapore Management University, the Hong Kong University of Science and Technology, and the University of Sydney, for useful comments. We also thank our discussant Matt Webb for insightful comments at the 2017 CESG. Our thanks go to three anonymous referees and a guest co-editor for insightful comments that helped us improve the paper.

References

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.

Bollerslev, T. (1986). "Generalized autoregressive conditional heteroskedasticity", *Journal of Econometrics* **31**, 307–327.

Bühlmann, P., (1997). "Sieve bootstrap for time series", *Bernoulli* **3**, 123–148.

Chan, K. Y. F. and S. M. S. Lee (2001). "An exact iterated bootstrap algorithm for small-sample bias reduction", *Computational Statistics and Data Analysis* **36**, 1–13.

Chang, J. and P. Hall (2015). "Double-bootstrap methods that use a single double-bootstrap simulation", *Biometrika*, doi:10.1093/biomet/asu060

Chang, Y. and J. Y. Park, (2002). "A sieve bootstrap for the test of a unit root", *Journal of Time Series Analysis* **24**, 379–400.

Davidson, J. (2006). "Alternative bootstrap procedures for testing cointegration in fractionally integrated processes", *Journal of Econometrics* **133**, 741–777.

Davidson, R. (2010). "Size distortion of bootstrap tests: Application to a unit root test", *Review of Economic Analysis* **2**, 169–193.

Davidson, R. (2015). "Computing, the bootstrap and economics", *Canadian Journal of Economics* **48**, 1195–1214.

Davidson, R. (2016a). "A Discrete Model for Bootstrap Iteration", *Journal of Econometrics*, **201**, 228–236, doi:10.1016/j.jeconom.2017.08.005

Davidson, R. (2016b). "Diagnostics for the Bootstrap and Fast Double Bootstrap", *Econometric Reviews*, **36**, 1021–1038, doi:10.1080/07474938.2017.1307918

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

Davidson, R. and J. G. MacKinnon (2002). "Fast double bootstrap tests of nonnested linear regression models", *Econometric Reviews* **21**, 419–429.

Davidson, R. and J. G. MacKinnon (2006). "The power of asymptotic and bootstrap tests", *Journal of Econometrics* **133**, 421–441.

Davidson, R. and J. G. MacKinnon (2007). "Improving the reliability of bootstrap tests with the fast double bootstrap", *Computational Statistics and Data Analysis* **51**, 3259–3281.

Davidson, R. and J. G. MacKinnon (2010). "Wild Bootstrap Tests for IV Regression", *Journal of Business & Economic Statistics*, **28**, 128–144.

DiCiccio, T. J., M. A. Martin, and G. A. Young (1992). "Analytical approximations for iterated bootstrap confidence intervals", *Statistics and Computing* **2**, 161–171.

Dufour, J.-M., and L. Khalaf (2001). "Monte Carlo test methods in econometrics", Ch. 23 in *A Companion to Econometric Theory*, ed. B. Baltagi, Oxford, Blackwell Publishers, pp. 494–519.

Durbin, J. (1970). "Testing for serial correlation in least-squares regression when some of the regressors are lagged dependent variables", *Econometrica* **38**, 410–421.

Efron, B. (1979). "Bootstrap methods: Another look at the jackknife", *Annals of Statistics* **7**, 1–26.

Engle, R. F. (1982). "Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation", *Econometrica* **50**, 987–1007.

Giacomini, R., D. N. Politis, and H. White (2013). "A warp-speed method for conducting Monte Carlo experiments involving bootstrap estimators", *Econometric Theory* **29**, 567–589.

Godfrey, L. G. (1978). "Testing against general autoregressive and moving average error models when the regressors include lagged dependent variables", *Econometrica* **46**, 1293–1301.

Hall, P. (1986). "On the bootstrap and confidence intervals", *Annals of Statistics* **14**, 1431–1452.

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

Hall, P. and M. A. Martin (1988). "On bootstrap resampling and iteration", *Biometrika* **75**, 661–671.

Ho, Y. H. S. and S. M. S. Lee (2005). "Iterated smoothed bootstrap confidence intervals for population quantiles", *Annals of Statistics* **33**, 437–462.

Horowitz, J. L. and N. E. Savin (2000). "Empirically relevant critical values for hypothesis tests", *Journal of Econometrics* **95**, 375–389.

Lamarche, J.-F. (2004). "The numerical performance of fast bootstrap procedures", *Computational Economics* **23**, 379–389.

Lee, S. M. S. and G. A. Young (1995). "Asymptotic iterated bootstrap confidence intervals", *Annals of Statistics* **23**, 1301–1330.

Lee, S. M. S. and G. A. Young (2003). "Prepivoting by weighted bootstrap iteration", *Biometrika* **90**, 393–410.

MacKinnon, J. G. (1994). "Approximate asymptotic distribution functions for unit root and cointegration tests," *Journal of Business and Economic Statistics*, **12**, 167–76.

MacKinnon, J. G. (1996). "Numerical distribution functions for unit root and cointegration tests", *Journal of Applied Econometrics*, **11**, 601–18.

Omtzigt, P., and S. Fachin (2006). "The size and power of bootstrap and Bartlett-corrected tests of hypotheses on the cointegrating vectors", *Econometric Reviews* **25**, 41–60.

Ouysse, R. (2013). "A fast iterated bootstrap bootstrap procedure for approximating the small-sample bias", *Communications in Statistics – Simulation and Computation*, **42**, 1472–1494, doi:10.1080/03610918.2012.667473

Palm, F. C., S. Smeekes, and J.-P. Urbain (2008). "Bootstrap unit root tests: comparison and extensions", *Journal of Time Series Analysis* **29**, 371–401.

Parzen, E. (1979). "Nonparametric Statistical Data Modeling", *Journal of the American Statistical Association*, **79**, 105–121.

Perron, P. and S. Ng (1996). "Useful modifications to unit root tests with dependent errors and their local asymptotic properties", *Review of Economic Studies* **63**, 435–465.

Psaradakis, Z. (2001). "Bootstrap tests for an autoregressive unit root in the presence of weakly dependent errors", *Journal of Time Series Analysis* **22**, 577–594.

Richard, P. (2009). "Modified fast double sieve bootstraps for ADF tests", *Computational Statistics & Data Analysis* **53**, 4490–4499.

Schilling, R. L. and L. Partzsch (2012). *Brownian Motion: An introduction to stochastic processes*, De Gruyter, Berlin/Boston.

Schwert, G. W. (1989). "Testing for unit roots: a Monte Carlo investigation", *Journal of Business and Economic Statistics* **7**, 147–159.

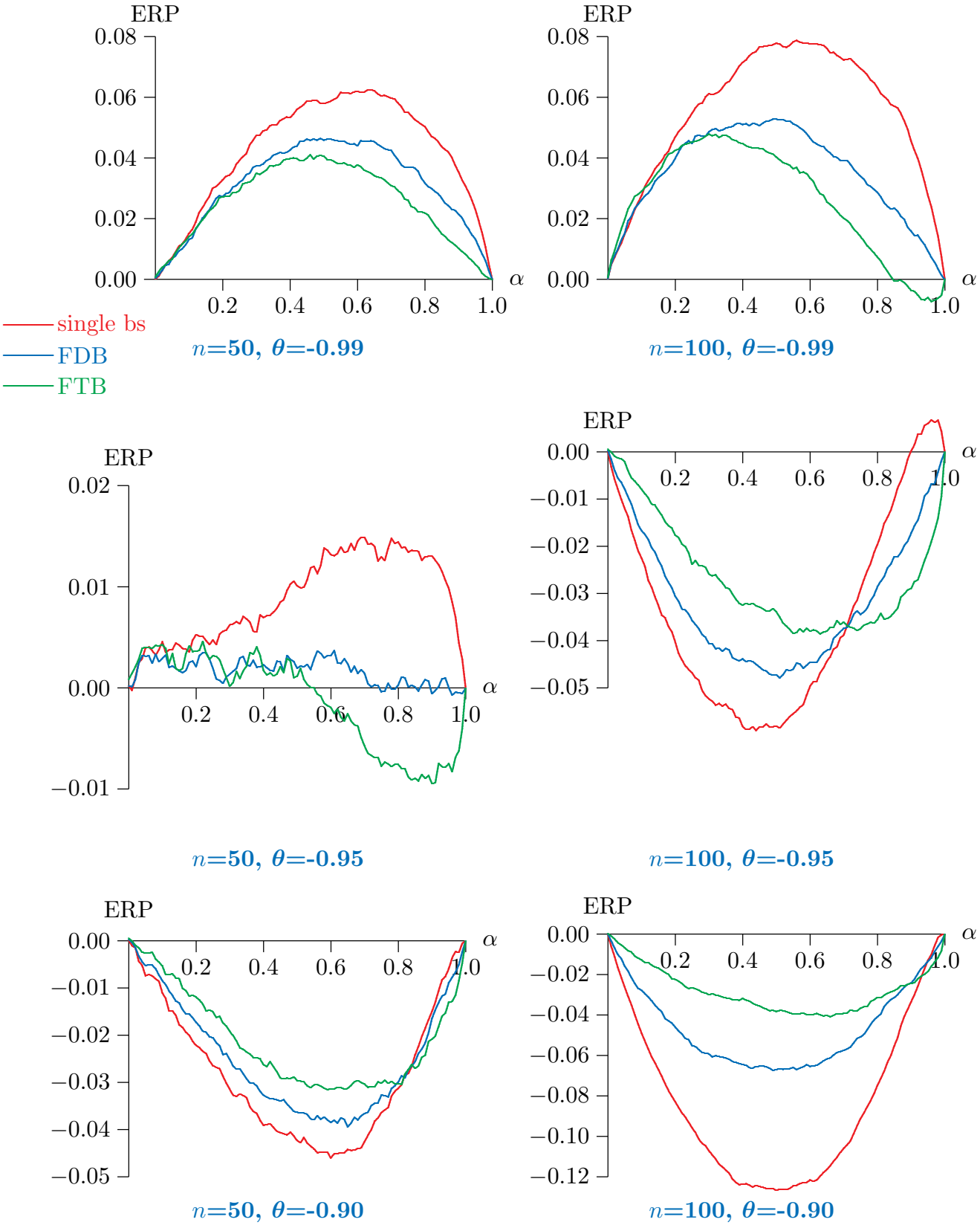


Figure 1: ERPs of unit root test

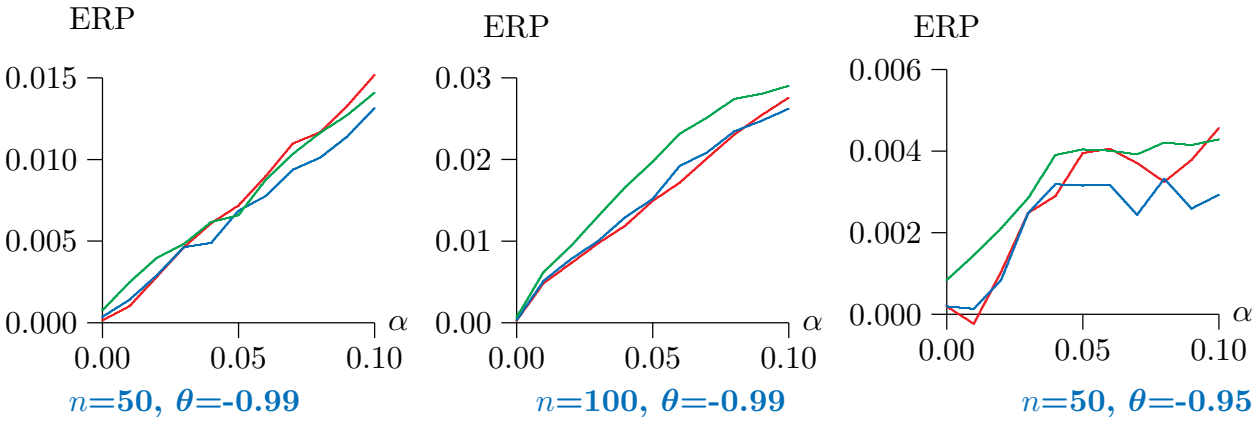


Figure 1a: Restricted view for some difficult cases

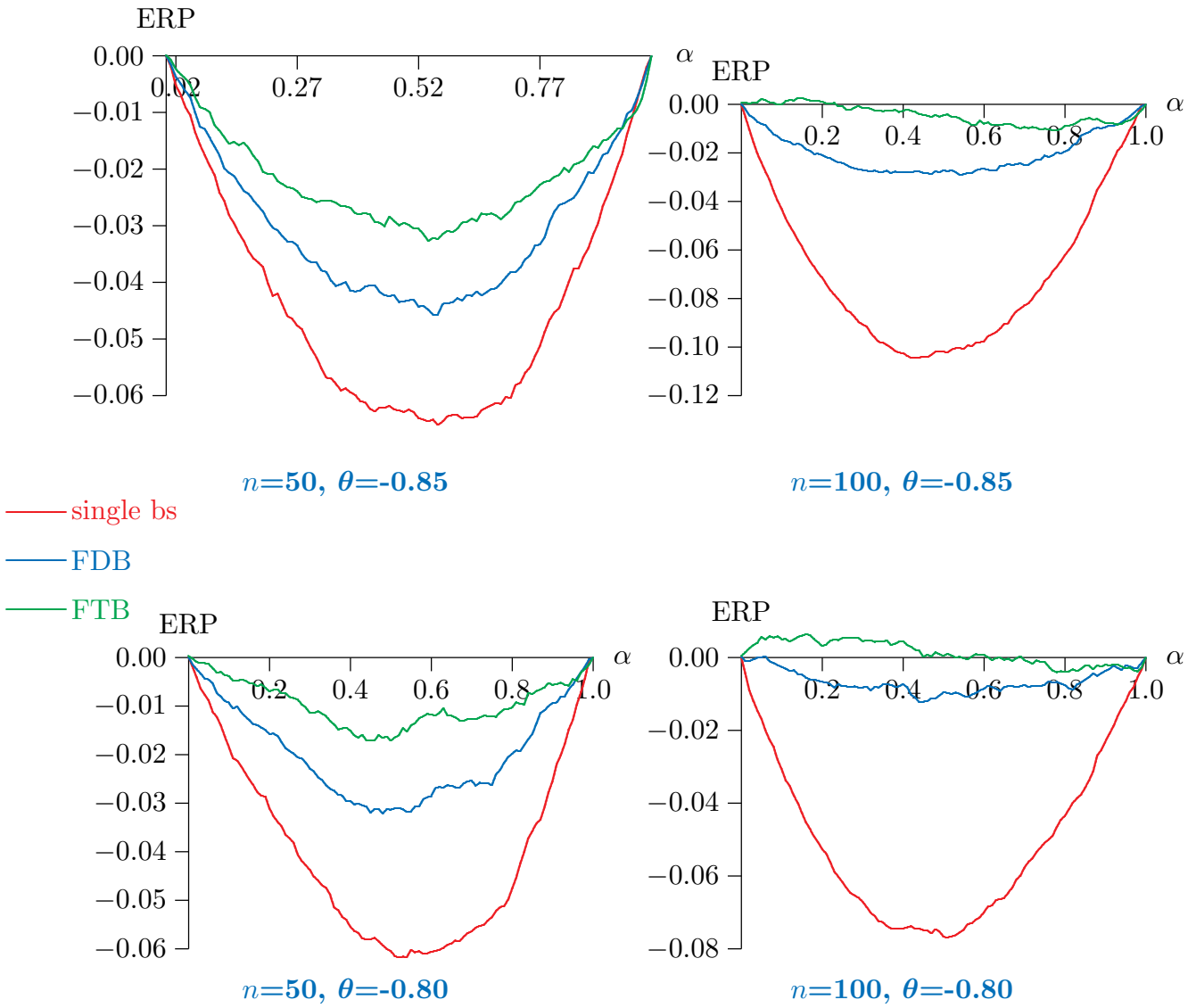


Figure 2: ERPs of unit root test

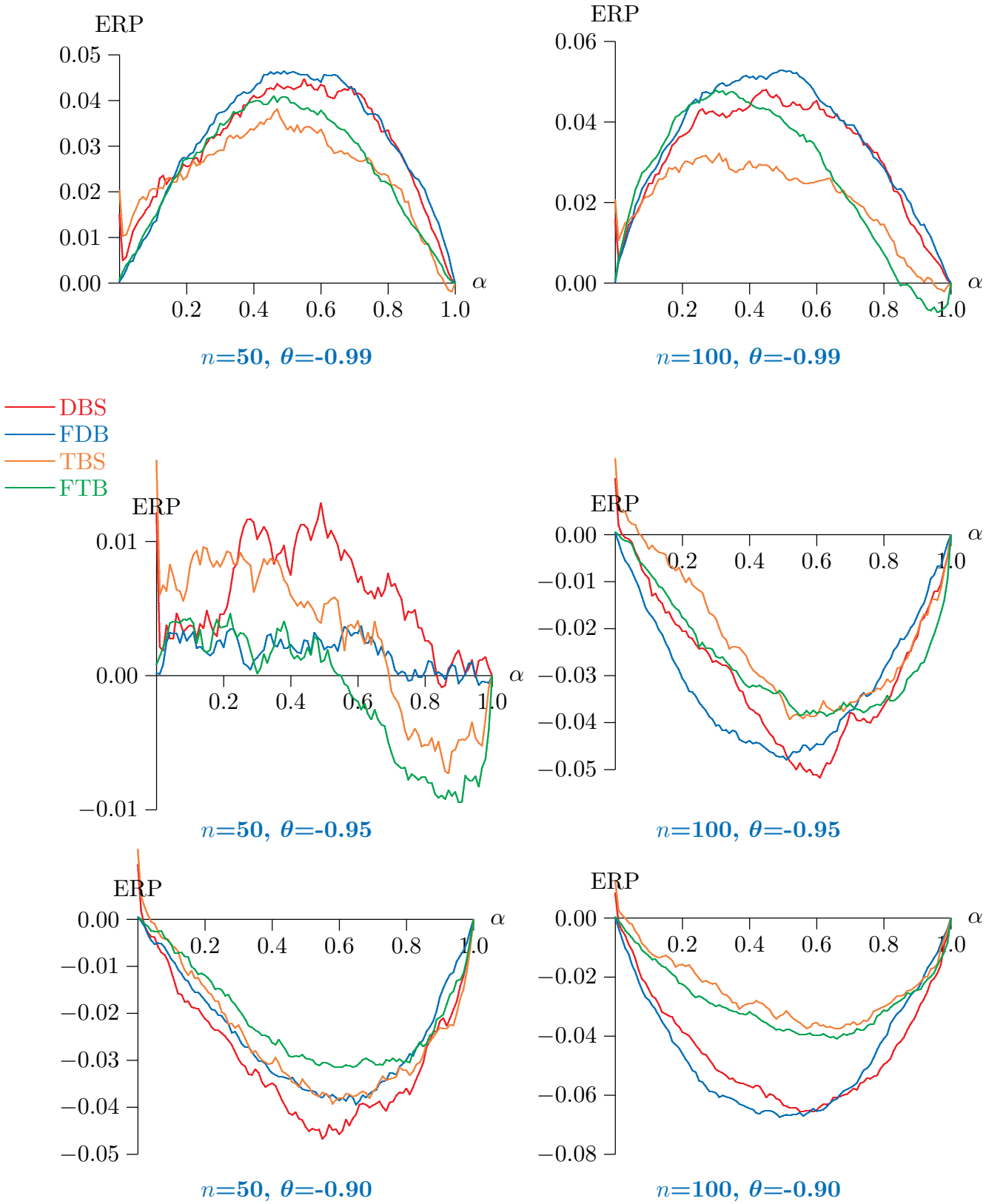
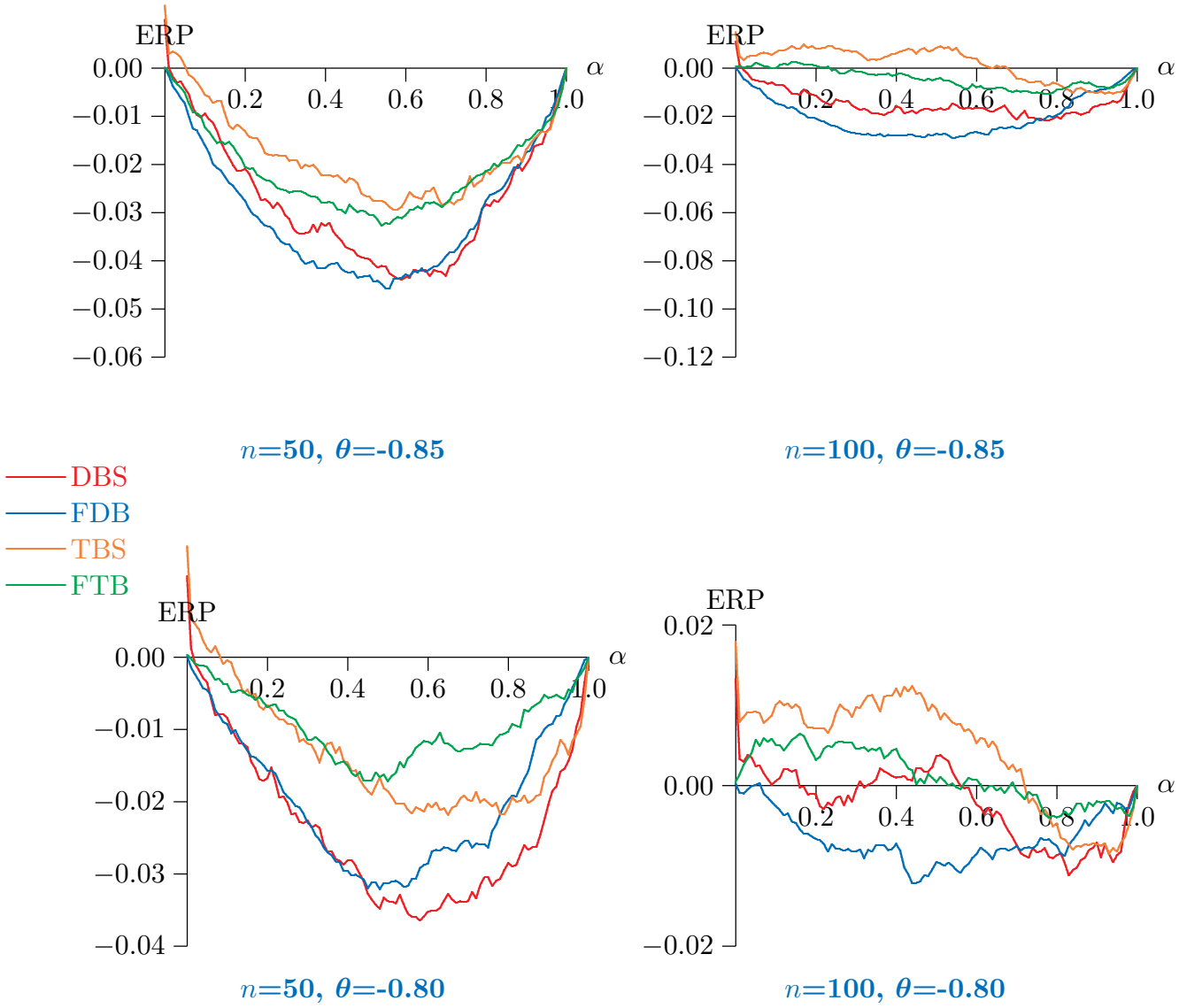


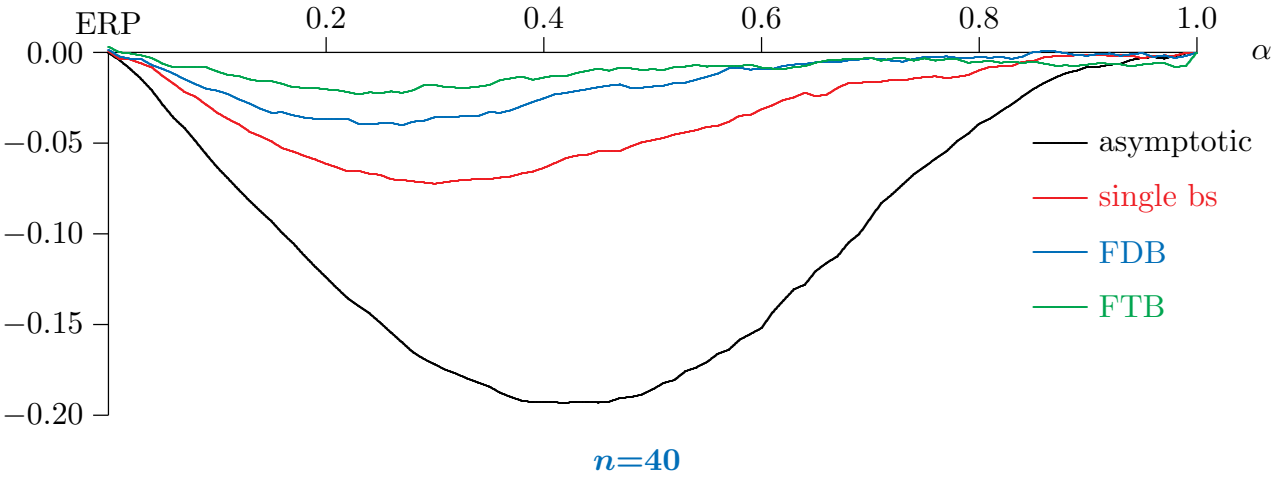
Figure 3: ERPs of unit root test



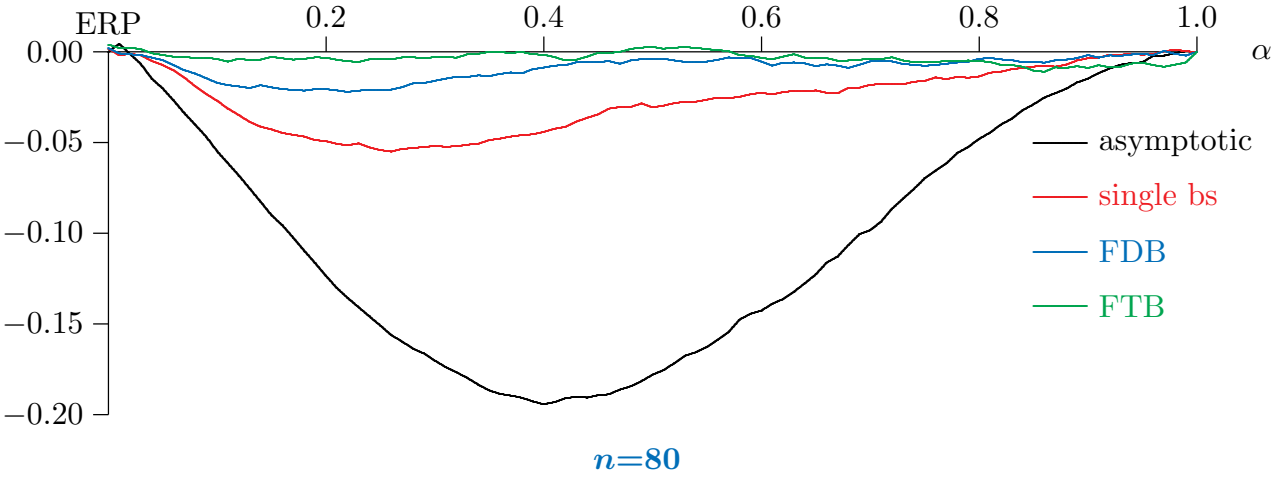
n=50, θ=-0.85 n=100, θ=-0.85

n=50, θ=-0.80 n=100, θ=-0.80

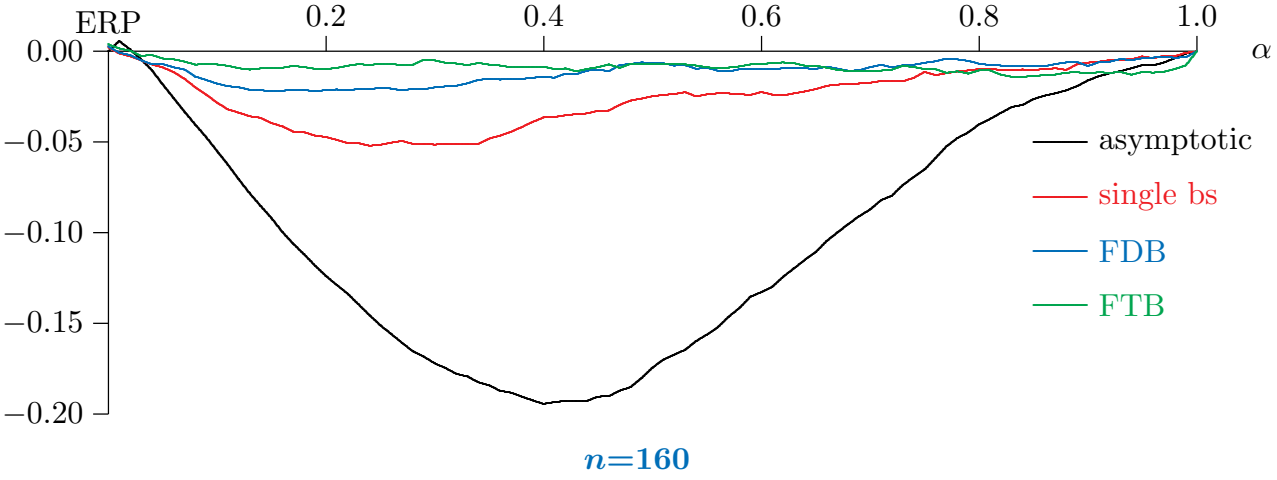
Figure 4: ERPs of unit root test



***n*=40**



***n*=80**



***n*=160**

Figure 5a: ERPs of test for ARCH with ordinary resampling

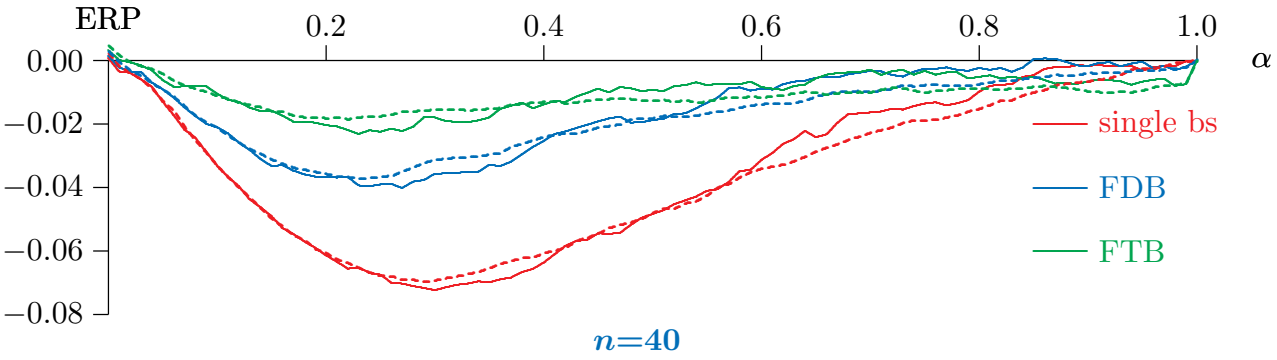


Figure 5b: ERPs for ordinary and smoothed resampling

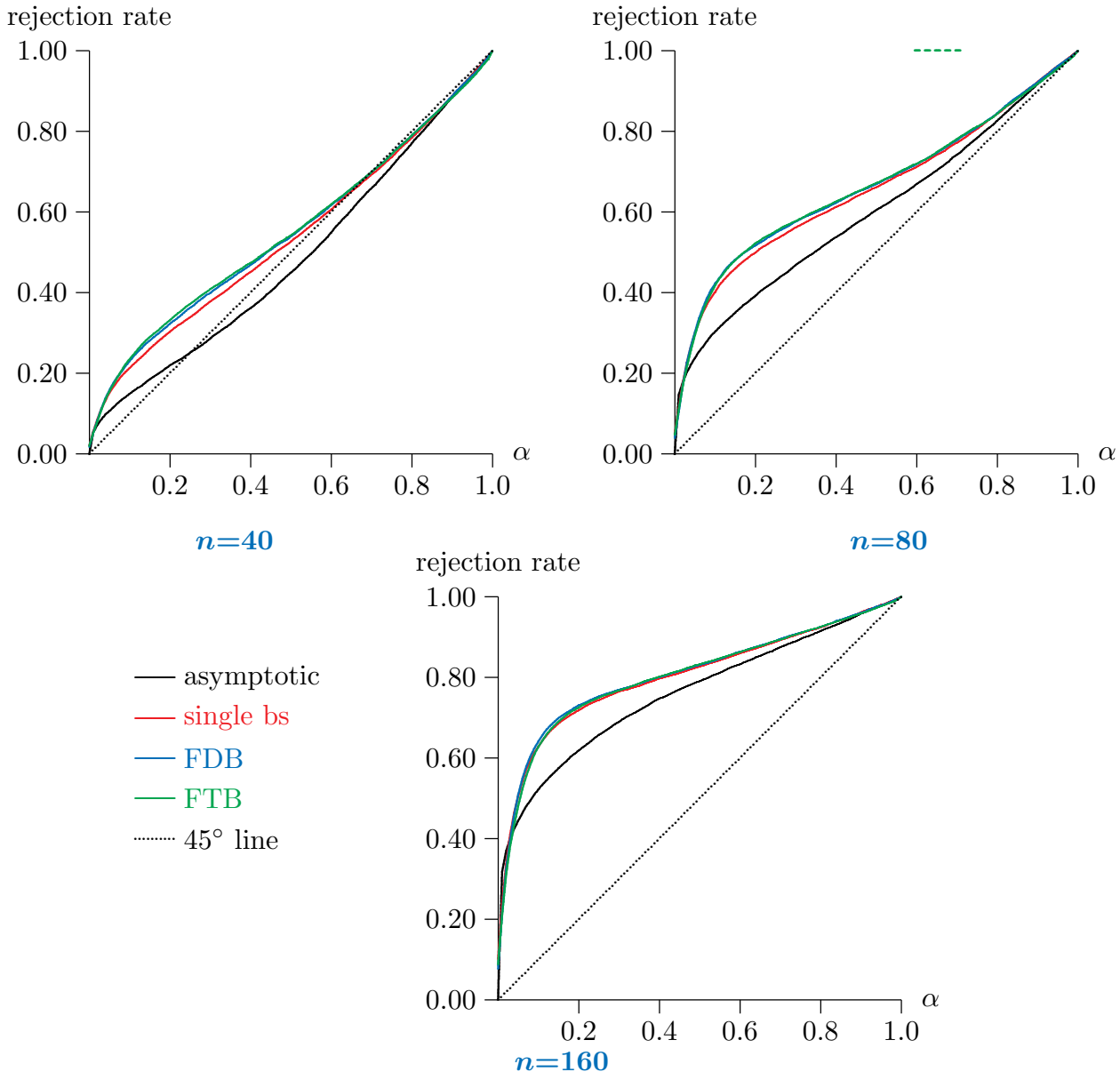
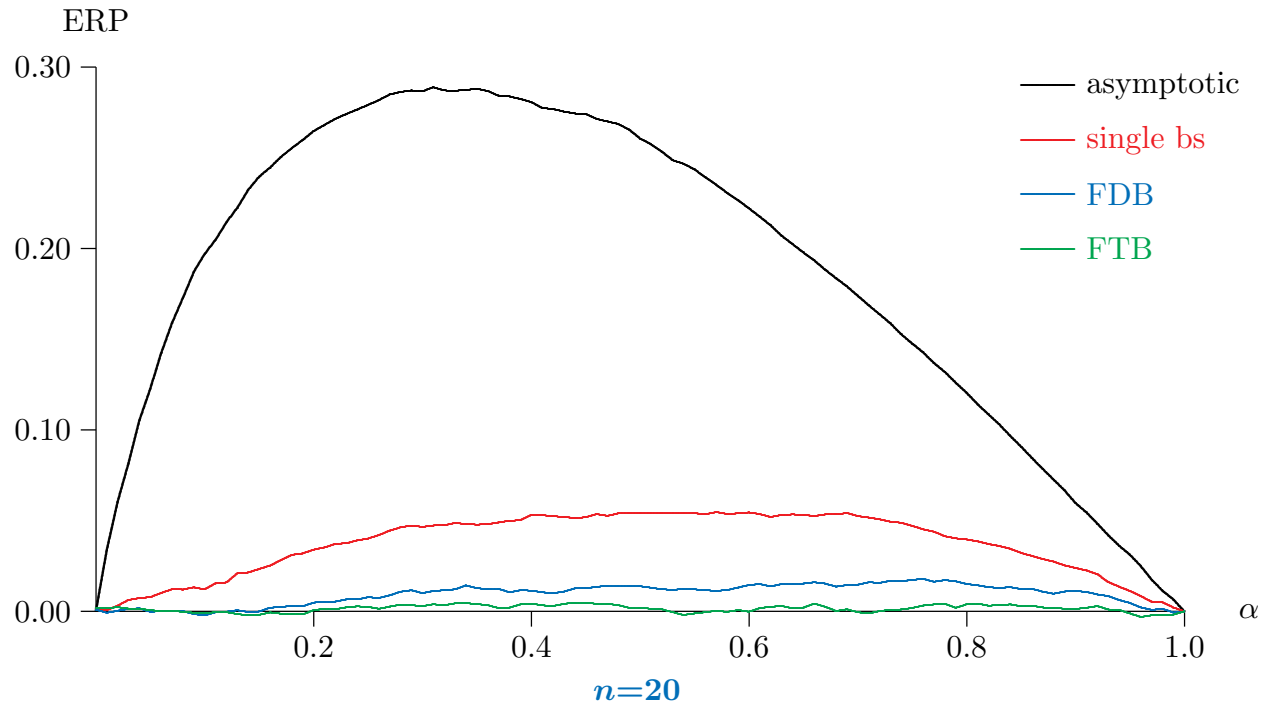


Figure 6: ARCH test rejection rate, α=1, γ=δ=0.3



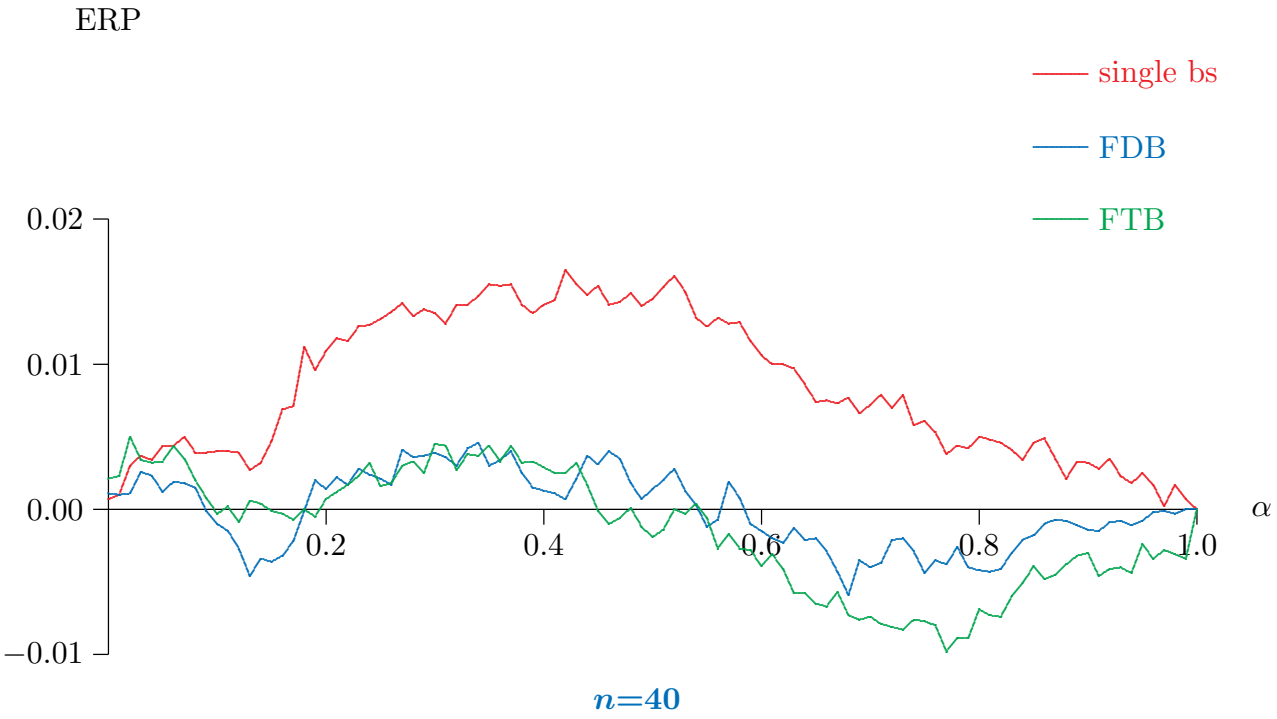
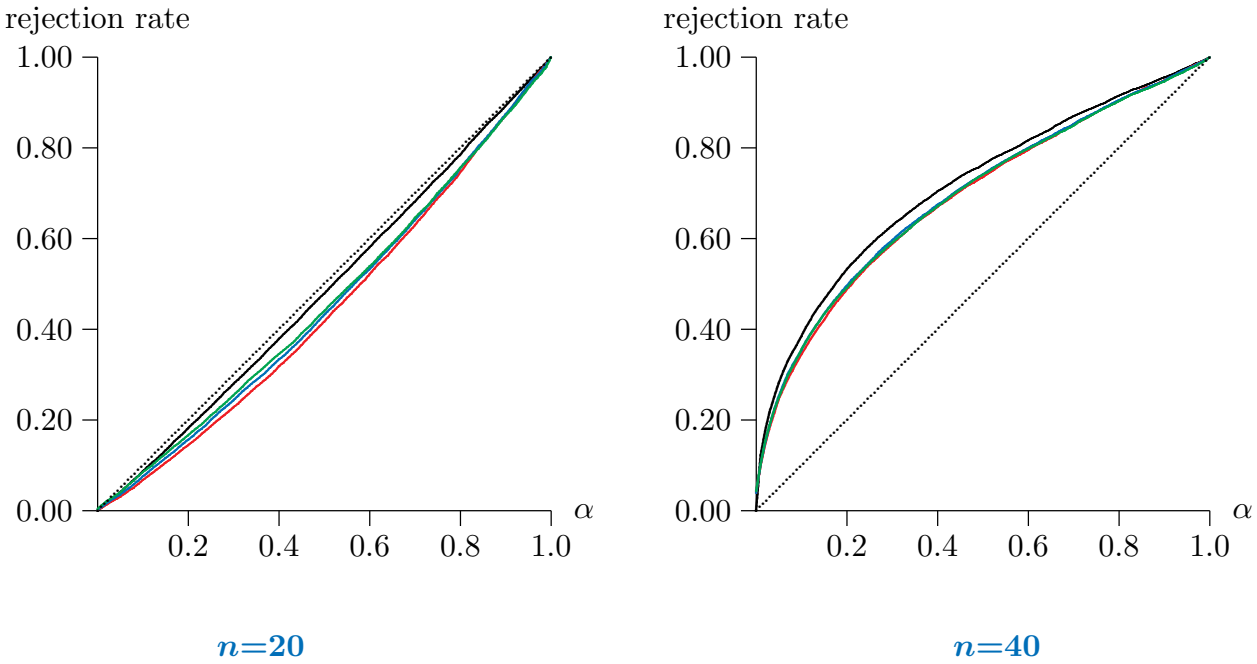
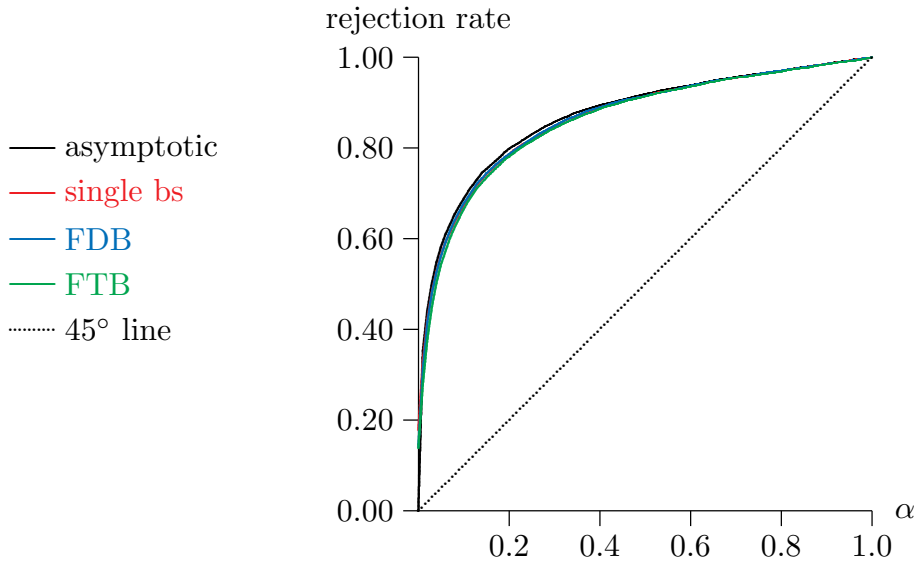


Figure 7: ERPs of Durbin-Godfrey test

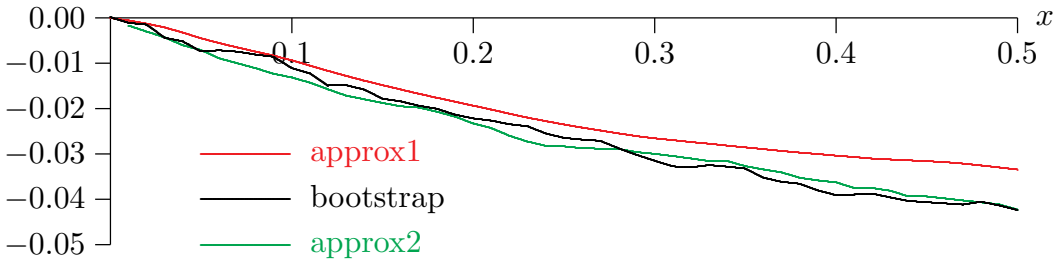


***n*=20** ***n*=40**

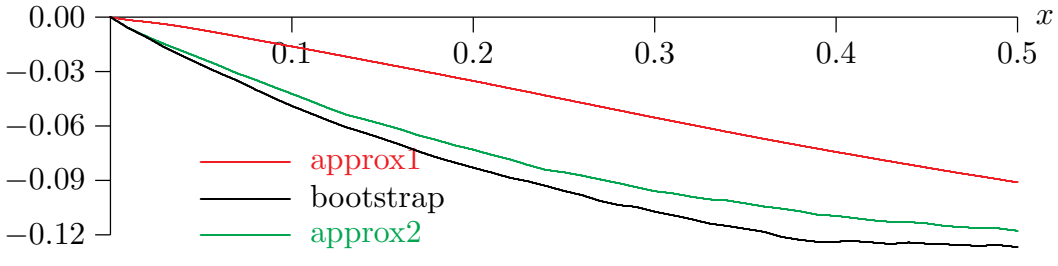


***n*=56**

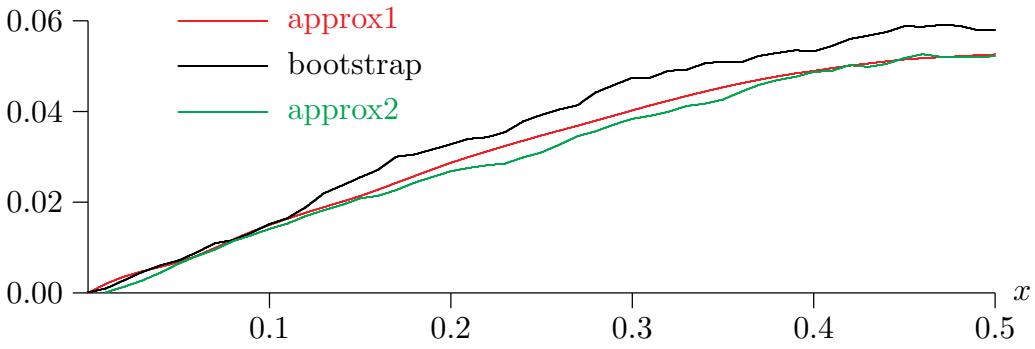
Figure 8: Durbin-Godfrey test rejection rate, ρ=0.5



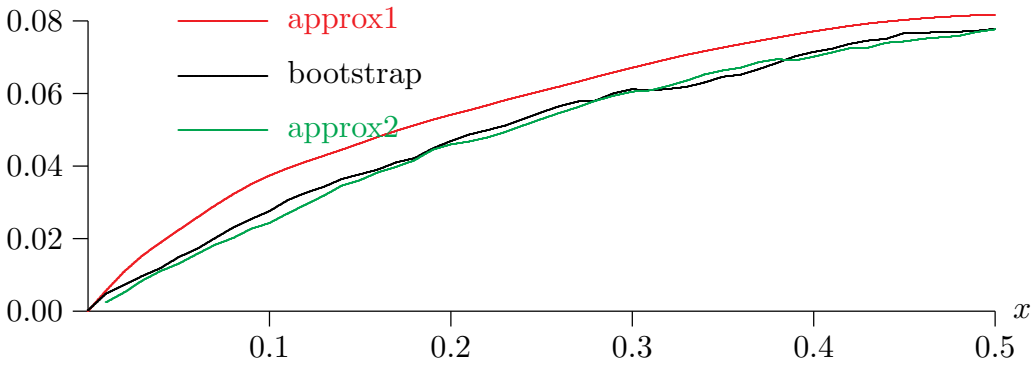
n=50, θ = -0.90



n=100, θ = -0.90

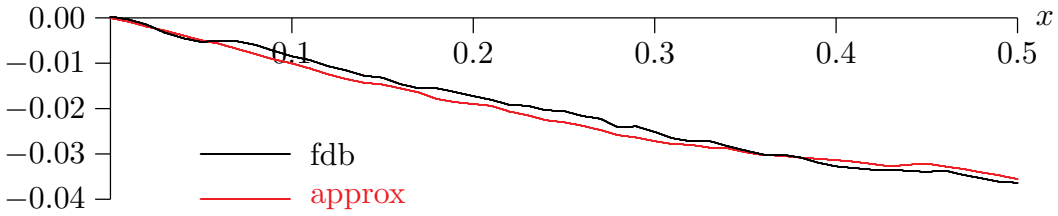


n=50, θ = -0.99

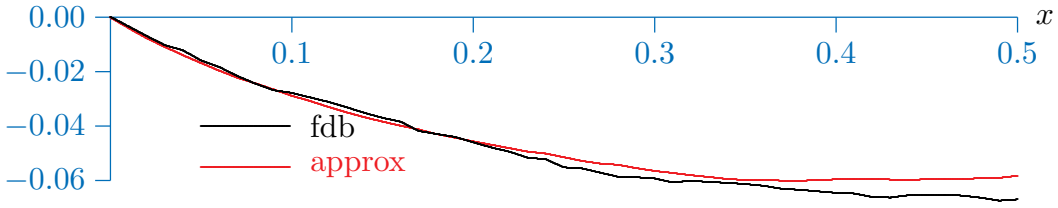


n=100, θ = -0.99

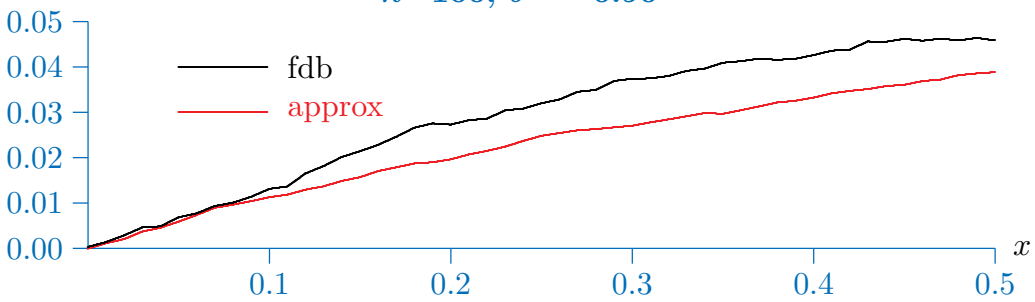
Figure 9: FDB approximations



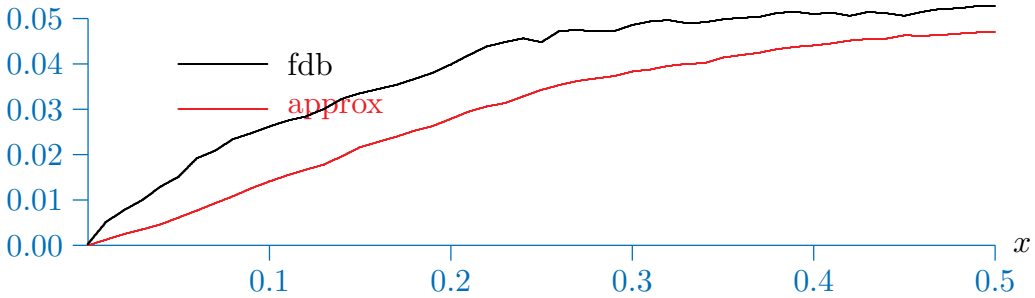
n=50, θ = -0.90



n=100, θ = -0.90



n=50, θ = -0.99



n=100, θ = -0.99

Figure 10: FTB approximation