L'Actualité économique

Maximum Non-Extensive Entropy Block Bootstrap for Non-stationary Processes

Michele Bergamelli, Jan Novotný and Giovanni Urga

Volume 91, Number 1-2, March–June 2015

Identification, Simulation and Finite-Sample Inference

URI: https://id.erudit.org/iderudit/1036916ar
DOI: https://doi.org/10.7202/1036916ar

See table of contents

Publisher(s)

HEC Montréal

ISSN

0001-771X (print)
1710-3991 (digital)

Cite this article

Bergamelli, M., Novotný, J. & Urga, G. (2015). Maximum Non-Extensive Entropy Block Bootstrap for Non-stationary Processes. *L'Actualité économique*, *91*(1-2), 115–139. https://doi.org/10.7202/1036916ar

Article abstract

In this paper, we propose a novel entropy-based resampling scheme valid for non-stationary data. In particular, we identify the reason for the failure of the original entropy-based algorithm of Vinod and López-de Lacalle (2009) to be the perfect rank correlation between the actual and bootstrapped time series. We propose the Maximum Entropy Block Bootstrap which preserves the rank correlation locally. Further, we also introduce the Maximum non-extensive Entropy Block Bootstrap to allow for fat tail behaviour in time series. Finally, we show the optimal finite sample properties of the proposed methods via a Monte Carlo analysis where we bootstrap the distribution of the Dickey-Fuller test.

L'Actualité économique, Revue d'analyse économique, vol. 91, nos 1-2, mars-juin 2015

MAXIMUM NON-EXTENSIVE ENTROPY BLOCK BOOTSTRAP FOR NON-STATIONARY PROCESSES*

Michele BERGAMELLI
Cass Business School
City University London
United Kingdom
Michele.Bergamelli.2@cass.city.ac.uk

Jan NOVOTNÝ
Cass Business School
City University London
United Kingdom
CERGE-EI
Czech Republic
Jan.Novotny.1@city.ac.uk

Giovanni URGA
Cass Business School
City University London
United Kingdom
Bergamo University
Italy
G.Urga@city.ac.uk

Abstract–In this paper, we propose a novel entropy-based resampling scheme valid for non-stationary data. In particular, we identify the reason for the failure of the original entropy-based algorithm of Vinod and López-de Lacalle (2009) to be the perfect rank correlation between the actual and bootstrapped time series. We propose the Maximum

* We wish to thank participants in the 15th OxMetrics Conference (Cass Business School, 4-5 September, 2014), in particular to Guillaume Chevillon for useful comments. Special thanks to Russell Davidson and Lynda Khalaf for very useful comments on a previous version of the paper. We are greatly in debt with the Editor, Marie-Claude Beaulieu, and an anonymous referee for providing us with very insightful comments and suggestions that greatly helped to improve the paper. The usual disclaimer applies. Michele Bergamelli acknowledges financial support from the "PhD Scholarship in Memory of Ana Timberlake", while Jan Novotny acknowledges financial support from the Centre for Econometric Analysis and GAČR grant 14-27047S.

Entropy Block Bootstrap which preserves the rank correlation locally. Further, we also introduce the Maximum non-extensive Entropy Block Bootstrap to allow for fat tail behaviour in time series. Finally, we show the optimal finite sample properties of the proposed methods via a Monte Carlo analysis where we bootstrap the distribution of the Dickey-Fuller test.

Introduction

Since the seminal contribution of Efron (1979). for *identically and independent distributed* (i.i.d.) data, the bootstrap method has been extended to handle more complex data structures. In particular, time-series data fail to satisfy the i.i.d. assumption because both the data distribution might well change over time and the observations are far from being independent. In order to preserve the dependence structure, Künsch (1989) proposes the non-parametric *block* bootstrap technique which involves re-sampling blocks of data rather than individual observations. In the same spirit, Buhlmann (1997) develops a parametric alternative usually called *sieve* bootstrap which circumvents the dependence structure in the data by first fitting an AR(p) process — where p grows with the sample size T — and then resampling from supposedly i.i.d. residuals. Finally, another widely employed resampling scheme, aimed at dealing with heteroskedasticity, is the *wild* bootstrap method developed by Wu (1986) (see also Mammen, 1993). Other resampling schemes, or variations on the above mentioned ones, are however available and we refer the interested reader to Politis, Romano, and Wolf (1999) for an overview.

The above methods are designed to work with dependence structures that die away over time, i.e. with stationary processes. Yet, in economics and finance, we frequently study relationships that involve integrated processes and most typically integrated processes of order one. In this case, we are interested in bootstrapping the distribution of the statistical test to assess the unit root hypothesis. To this purpose, it is possible to follow two approaches: the first one consists in resampling the first differences of the data (which can exhibit their own dependence structure), the second one consists of resampling directly from the original non-differenced data. The first approach is by far the most widespread as after first-differencing it is possible to apply resampling schemes valid for stationary data (see Palm, Smeekes, and Urbain, 2007 for an overview of the alternative methods). On the contrary, resampling directly from non-stationary data seems to be still rather uncommon practice as it requires resampling schemes designed to mimic the stochastic trend exhibited by random walk processes. The major benefit of resampling directly from the original dataset is that we can ignore the dependence structure in the first differences. The only valid method providing consistency of the bootstrap procedure available in the statistical literature is the Continuous Path Block Bootstrap of Paparoditis and Politis (2001). In the econometric literature, the method is largely unexplored and although we have a theoretical validation of its consistency (Phillips, 2010), its finite sample behaviour has not been explored yet.

Recently, Vinod and López-de Lacalle (2009) have proposed a new bootstrap method based on the entropy concept allowing for any degree of persistence in the time series, including the unit root case. As pointed out for instance by Davidson and Monticini (2014), however, this resampling scheme has internal flaws, which cause significant size distortions.

In this paper, we show that this resampling scheme has significant size distortions when used to bootstrap the distribution of a test statistic under the null of unit root and we propose a correct entropy-based resampling scheme valid for non-stationary data. In particular, we identify the reason for the failure of the original entropy-based algorithm of Vinod and López-de Lacalle (2009) to be the perfect rank correlation. We therefore relax the rank constraint and preserve the rank correlation locally; this is our proposed Maximum Entropy Block Bootstrap. Further, we employ the notion of non-extensive entropy allowing for power-law behaviour leading to the general Maximum non-extensive Entropy Block Bootstrap. Finally, we compare the finite sample properties of our proposed methods with respect to the existing ones via a Monte Carlo analysis where we bootstrap the distribution of the Dickey-Fuller test.

The remainder of the paper is organized as follows: In Section 1, we briefly introduce the Maximum Entropy Bootstrap and explore the reasons beyond its failure. In Section 2, we propose a correction and introduce the notion of the Maximum Entropy Block Bootstrap. In Section 3, we evaluate its finite sample properties for the unit root test case via an extensive Monte Carlo simulation. In Section 4, we generalize the concept of entropy to non-extensive one providing the definition of the Maximum non-extensive Entropy Block Bootstrap. The last section concludes.

1. Maximum Entropy Bootstrap

The Maximum Entropy Bootstrap (MEB), introduced by Vinod and López-de Lacalle (2009), is a fully non-parametric bootstrap technique designed to resample from time series with any level of persistence. The method is based on the maximum information entropy principle, which works as follows: the probability distribution to find a system in a given state conditional on the prior data is such that the information entropy is maximized. This principle allows one to avoid any functional assumptions on the probability distribution function.

Let f(x) be a probability density function to find the system in a state x, then the Shannon entropy, H_S, is defined as:

$$H_S = E\left[-\log\left(f(x)\right)\right]. \tag{1}$$

The maximum entropy principle leads to a probability distribution function which satisfies the following optimization problem

$$f = \arg\max_{f'} E\left[-\log\left(f'(x)\right)\right]. \tag{2}$$

There are several solutions to (2). First, for the system with finite and bounded support, the probability density function is the *uniform distribution*. Second, for the system with half-infinite support and finite means, the probability density function is the *exponential distribution*. Third, for the system with infinite support and given mean and standard deviation, the *normal distribution* is the one which maximizes the entropy.

In order to use the Maximum Entropy principle to construct a bootstrap method valid for time series with any level of persistence, we have to ensure that the full set of prior information is correctly taken into account, and that persistence is preserved. Vinod and López-de Lacalle (2009) propose the MEB procedure as a solution to address both points; they suggest to impose the mass preserving and the mean preserving constraints to incorporate all the prior information, and to impose the perfect rank correlation to keep the memory of the system, respectively.

1.1 *The MEB Algorithm*

Let us consider an observed time series $X = x_1,\ldots,x_T$ and denote the associated order statistics as $x_{(t)}$. Further, let us assume we know the support of the order statistics $x_{(t)} \in \left[x_{(0)}, x_{(T+1)}\right]$. We define the midpoints z_t as

$$z_t = \frac{1}{2}\left(x_{(t)} + x_{(t+1)}\right),\ t \in \{1,\ldots,T-1\},$$

$$z_0 = x_{(0)},$$

$$z_T = x_{(T+1)}.$$

Using the midpoints, we define T half-open intervals $I_t = \left(z_{t-1}, z_t\right]$ around each observation. The maximum entropy density function is the solution to (2) with two additional constraints:

a. The *mass preserving constraint* imposed on the density function states that, on average, $1/T$ of the mass of the density function lies in each of the intervals I_t.

b. The *mean preserving constraint* states that

$$\sum_{t=1}^{T} x_t = \sum_{t=1}^{T} x_{(t)} = \sum_{t=1}^{T} m_t,$$

where m_t is the mean of f over the interval I_t.

The constrained solutions are given by the following choice of the density function

$$f(x) = \frac{1}{z_1 - z_0}, \qquad x \in I_1, \qquad m_1 = \frac{3x_{(1)}}{4} + \frac{x_{(2)}}{4}, \tag{3}$$

$$f(x) = \frac{1}{z_k - z_{k-1}}, \quad x \in I_k\, k \in \{2,\ldots,T-1\}, \qquad m_k = \frac{x_{(k-1)}}{4} + \frac{x_{(k)}}{2} + \frac{x_{(k+1)}}{4}, \tag{4}$$

$$f(x) = \frac{1}{z_T - z_{T-1}}, \qquad x \in I_T, \qquad m_T = \frac{x_{(T-1)}}{4} + \frac{3x_{(T)}}{4}. \tag{5}$$

The mass of the distribution f over the intervals I_1 and I_T depends on the choice of the $x_{(0)}$ and $x_{(T+1)}$ respectively. We can therefore impose the alternative constraints to the one in (4), which would in fact define the $x_{(0)}$ and $x_{(T+1)}$, respectively. On the other hand, Vinod and López-de Lacalle (2009) algorithm sets $x_{(0)}$ and $x_{(T+1)}$ based on some distributional properties of the sample path. Then m_1 and m_T are implied by this choice.

To create a single realization $X \to X^*$, the MEB is based on the following algorithm:

Step 1 We create the order statistics $x_{(t)}$ based on the empirical data set x_t and define the support of the order statistics $\left[x_{(0)}, x_{(T+1)}\right]$ where:

$$x_{(0)} = x_{(1)} - d_{trm},$$

$$x_{(T+1)} = x_{(T)} + d_{trm},$$

with $d_{trm} = E_{trim}\left[\left|x_{(t)} - x_{(t-1)}\right|\right]$ being the trimmed mean of the distances between the consecutive sorted observations.

Step 2 We define a (T x 2) sorting matrix, $\mathbf{S}_1$, and place the index set $t = \{1, \dots, T\}$ in the first column and the observed time series x_t in the second column.

Step 3 We sort the matrix $\mathbf{S}_1$ with respect to the second column, x_t, and define the order statistics $x_{(t)}$ We then define the midpoints z_t and the half-open intervals I_t.

Step 4 We draw T uniform pseudo-random numbers $p_s \sim U[0,1]$, with $s \in \{1, \dots, T\}$ and assign the range $R_t = \left(\frac{t}{T}, \frac{t+1}{T}\right]$ for $t \in \{0, T-1\}$ wherein each p_s falls.

Step 5 We match each R_t with I_t and using the density function defined in (3) - (5), we draw the new set $\tilde{x}_t^*$.

Step 6 We define a corresponding (T x 2) sorting matrix $\mathbf{S}_2$, analogous to $\mathbf{S}_1$. We sort the T elements $\tilde{x}_t^*$ in an increasing order of the magnitude to form the ordering statistics $x_{(t)}^*$.

Step 7 We replace the second column of $\mathbf{S}_1$, the order statistics $x_{(t)}$, by the second column of $\mathbf{S}_2$, the order statistics $x_{(t)}^*$ of the newly generated set. We sort the $x_{(t)}^*$ based on the first column of $\mathbf{S}_1$, and thus recover x_t^*. The set x_t^* represents a resampled set of observations x_t.

The MEB algorithm can be iteratively employed to approximate the distribution of the desired statistic. The method thus combines the maximum entropy principle with the perfect rank correlation to resample from the observed data.

1.2 *The Maximum Entropy Bootstrap to Assess the Unit Root Hypothesis*

Let us consider a standard AR(1) process frequently used in the econometric analysis

$$x_t = \rho \cdot x_{t-1} + \varepsilon_t, \qquad t = 1,\ldots,T, \tag{6}$$

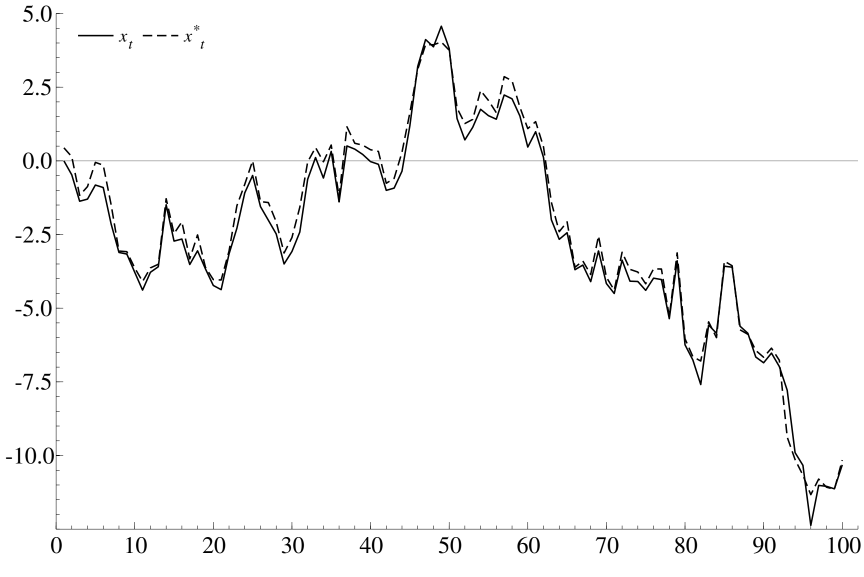
where $\varepsilon_t \sim i.i.d.N(0,1)$. In this paper we focus on the unit root case, i.e., $\rho = 1$.

Figure 1 reports the plot of the sample path generated by (6) with $\rho = 1$ and the bootstrapped path obtained using the Maximum Entropy algorithm proposed above. The figure shows that the MEB algorithm provides a close replication of the original time series.

We employ the MEB to assess the rejection frequency of the test with null hypothesis $\rho = 1$ in (6) against the alternative of $|\rho| < 1$. In Figure 2, we plot the empirical rejection frequencies to assess the quantile of the bootstrap distribution of the *t*-statistic. We generate 1000 series from the data generating process in (6) and for each replication, we create 299 bootstrap samples. We consider $T = 100$ and initial point to be set at $x_0 = 0$.

FIGURE 1

THE MEB SAMPLE PATH REPLICATION



NOTE : The figure reports the sample path of X and its bootstrapped counterpart X^* obtained using the MEB algorithm. The true data generating process is given as $x_t = x_{t-1} + \varepsilon_t$, with $\varepsilon_t \sim N(0,1)$, $x_0 = 0$, and $T = 100$.

We report the Q-Q plot for the MEB of the test under the null $H_0 : \rho = 1$ against the alternative $H_A : |\rho| < 1$. In our case, the Q-Q plot deviates from the 45^0 line, where the MEB has very low size and thus suggests the presence of significant flaws in the MEB suggested by Vinod and López-de Lacalle (2009).

1.2.1 *Why the Maximum Entropy Bootstrap Fails?*

The MEB test may fail for two main reasons, as suggested by Davidson and Monticini (2014):

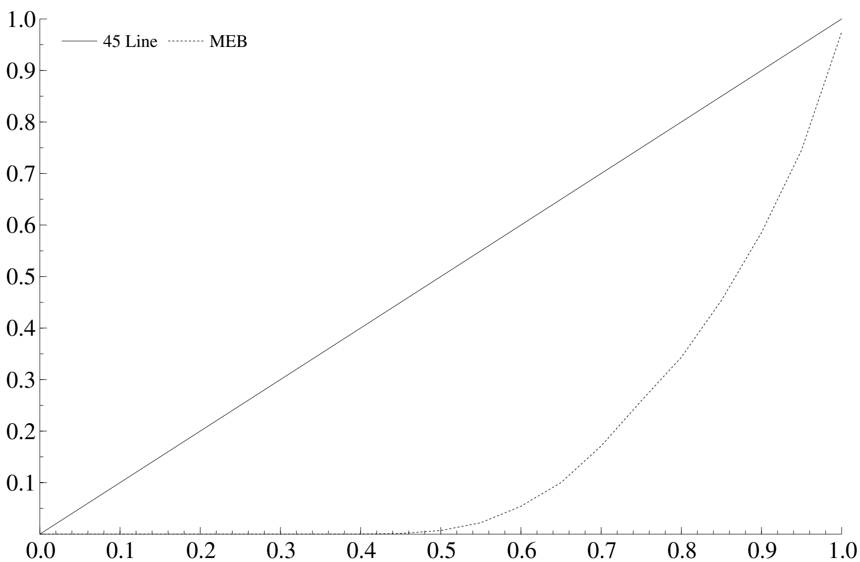
1. The distribution of the MEB statistic is on average more dispersed than that of the statistic itself. Thus, the mass of the bootstrap distribution to the right of the statistic is too large and so the *p*-value. Analogously, when the *t*-statistics is small, the *p*-value is small.

2. For each replication, the bootstrap statistics are positively correlated with the original statistic meaning that the bootstrap does not provide an independent draw of a sample.

In order to identify the cause of the MEB failure, we first plot the distribution of the *t*-statistic using the same setup we employ to generate the sample path.

FIGURE 2

EMPIRICAL REJECTION FREQUENCIES OF THE MEB



NOTE: The figure reports the empirical rejection frequencies against the nominal levels of the test H_0: $\rho = 1$ against H_A : $|\rho| < 1$ with distribution approximated by MEB.

Figure 3 reports the plot of the distribution of the original t-statistic for 1000 replications under H_0: $\rho = 1$. Further, for each realization of the data generating process, we store the value of one boostrapped statistic to obtain the following series $\{\tau_i^*\}_{i=1}^{1000}$ and we plot its distribution. The figure shows that the two distributions are very close to each other and hence the first explanation does not apply. Therefore, we check whether the second one may explain the MEB failure. In order to study the relationship between the two statistics, we regress the bootstrapped t-statistics obtained using the MEB algorithm, τ_i^*, onto the t-statistics obtained simulating from the data generating process. The resulting estimates are

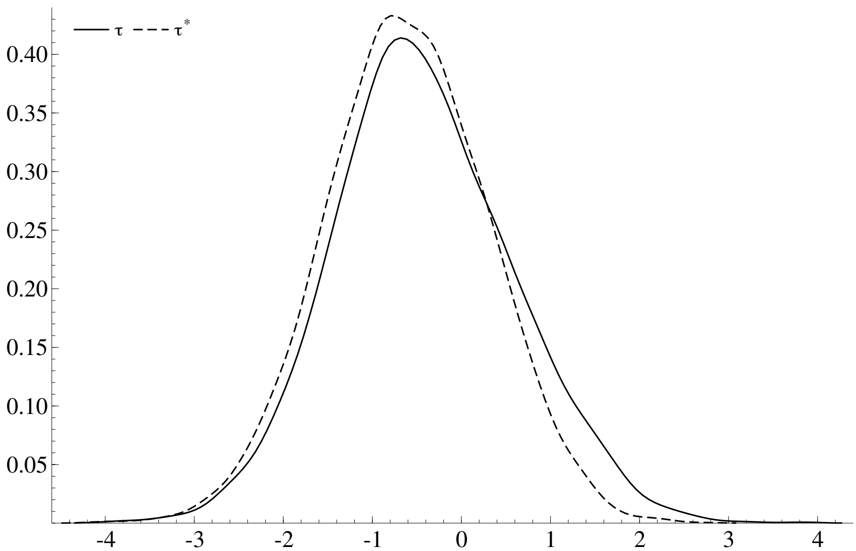
$$\tau_i^* = \underset{(0.00326)}{-0.2452} + \underset{(0.00299)}{0.8728}\, \tau_i, \quad i = 1,\ldots,1000, \tag{7}$$

with an adjusted $R^2 = 0.9445$, suggesting that the samples obtained using the MEB mimics too closely the sample they are drawn from, as Figure 1 also shows. Thus, there is a strong positive correlation between the two t-statistics which explains the severe undersize of the unit-root test when bootstrapping its distribution using the MEB algorithm.

Another way to visualize why things go wrong is to plot the joint density of the Monte Carlo statistic and the bootstrapped counterpart. Figure 4 reports the

FIGURE 3

DISTRIBUTION OF THE MEB STATISTIC UNDER H_0: $\rho = 1$



NOTE: The figure reports the distribution of the original t-statistic and one corresponding bootstrapped replication under H_0: $\rho = 1$.

kernel based joint density of (τ_i, τ_j^*) (gaussian kernel) and the related contour plot which resembles an ellipse with the semi-major axis sloping upwards. This is symptomatic of the high correlation between the bootstrapped statistics and the corresponding Monte Carlo draws.

It is worth stressing that the maximum entropy principle itself is not the cause of the MEB failure. Rather, the perfect rank correlation between the true data and the bootstrapped draws is causing the problems. In the next section, we propose a novel procedure to overcome the issue with the perfect rank correlation.

2. A New Procedure: The Maximum Entropy Block Bootstrap

We introduce the Maximum Entropy Block Bootstrap (MEBB), which preserves the perfect rank correlation *locally*, i.e. within each block, but not across the entire sample path. In order to ensure consistency of the bootstrap procedure, we apply the algorithm on the time series obtained as a partial sum of the demeaned residuals as in Paparoditis and Politis (2001)[1]. In addition, it is free of tail trimmings.

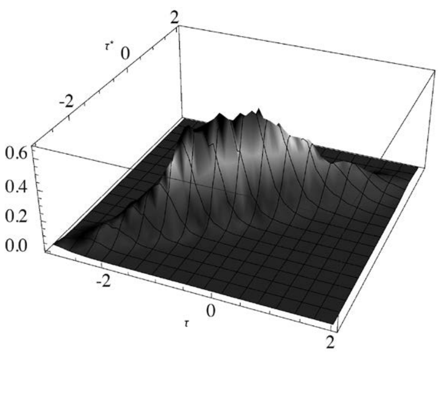
2.1 *The MEBB Algorithm*

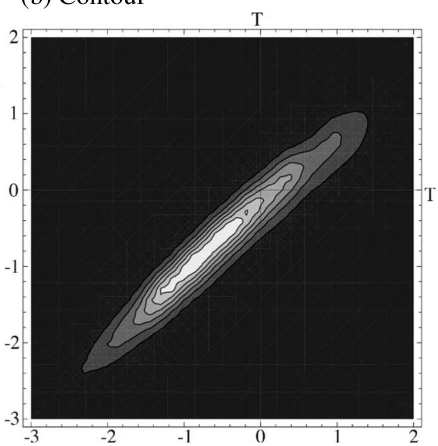
We propose to break the perfect rank correlation locally such that the MEB algorithm is employed block-wise. In each block, the perfect rank correlation is preserved, while over the entire sample path, the rank correlation between the

FIGURE 4

Joint Kernel Density

(a) Gaussian Kernel (b) Contour





1. Phillips (2010) shows that this initial step is not needed to achieve consistency under the null hypothesis; however, this is necessary to ensure the consistency of the bootstrap under the alternative hypothesis.

entire data generating process and the bootstrapped sample is unrestricted. We define the MEBB algorithm as follows:

Step A We choose the block length $\ell < T$ and let $i_0, i_1, \ldots, i_{k-1}$ i.i.d. uniform random numbers on the set $[1,2,\ldots,T-\ell]$ where $k = \lfloor T/\ell \rfloor$, (number of blocks).

Step B For each i_j, with $j = 0,\ldots,k-1$, we get the subset of the original time series $X^{(j)} = \left\{ x_{i_j}, x_{i_j+1}, \ldots, x_{i_j+\ell-1} \right\}$.

Step C We apply the MEB algorithm corresponding to Steps 1-7 in Section 2 for each subset $X^{(j)}$ separately and generate $X^{*(j)} = \left\{ x^*_{i_j}, x^*_{i_j+1}, \ldots, x^*_{i_j+\ell-1} \right\}$.

Step D We recover the bootstrapped sample path by sewing the $X^{*(1)}, \ldots, X^{*(k)}$ such that $x^*_{i_{j+1}} - x^*_{i_j+\ell-1}$ is set in a way to correspond to the difference $x_{i_{j+1}} - x_{i_j+\ell-1}$.

Step E If the length of the bootstrapped sample path is exceeding T, we take the first T values.

In addition, when employing the MEB, we omit the trimming in the algorithm. In such a case, Step 1 in the MEB algorithm is modified as follows:

Step 1* We create an order statistics $x_{(t)}$ based on the empirical data set x_t and define the support of the empirical data to be $[-\infty, \infty]$.

The constrained solution for the maximum entropy distribution on the half-open interval $[0,\infty)$ is given by $f = \lambda e^{-\lambda x}$ with mass at $1/\lambda$. Therefore, the solution analogous to (3) and (5) for the intervals $I_1 \equiv (-\infty, z_1]$ and $I_T \equiv [z_T, \infty)$, respectively, is given by

$$f_{I_1}(x) = \beta_1 \lambda_1 e^{-\lambda_1 (z_1 - x)}, \qquad x \in I_1, \qquad m_1 = \frac{3x_{(1)}}{4} + \frac{x_{(2)}}{4}, \tag{8}$$

$$f_{I_T}(x) = \beta_T \lambda_T e^{-\lambda_T (x - z_T)}, \qquad x \in I_T, \qquad m_T = \frac{x_{(T-1)}}{4} + \frac{3x_{(T)}}{4}, \tag{9}$$

where the parameters λ_1 and λ_T are set such that the mean preserving constraints imposed on m_1 and m_T are satisfied, respectively, while β_1 and β_T assures that the mass preserving constraints hold. Namely,

$$\beta_1 = \beta_T = \frac{1}{T}, \tag{10}$$

and

$$\lambda_1 : \frac{3x_{(1)}}{4} + \frac{x_{(2)}}{4} = \int_{-\infty}^{\frac{x_{(1)}+x_{(2)}}{2}} dx\, x \lambda_1 e^{-\lambda_1 \left(\frac{x_{(1)}+x_{(2)}}{2} - x \right)}, \tag{11}$$

$$\lambda_T : \frac{x_{(T-1)}}{4} + \frac{3x_{(T)}}{4} = \int_{\frac{x_{(T-1)}+x_{(T)}}{2}}^{\infty} dx\, x \lambda_T e^{-\lambda_T \left(x - \frac{x_{(T-1)}+x_{(T)}}{2} \right)} \tag{12}$$

giving $\lambda_1 = \dfrac{4}{x_{(2)} - x_{(1)}}$ and $\lambda_T = \dfrac{4}{x_{(T)} - x_{(T-1)}}$, respectively.

Finally, Steps 2-7 of the MEBB algorithm for a given subset X^* are the same as for the MEB. To draw an observation from I_1 and I_T, we employ the knowledge of the analytic solution to the cdf of f_{I_1} and f_{I_T}, respectively. The advantage is that the mean preserving constraint is by construction satisfied for the tails, as opposed to use the trimmed values.

Such a modified algorithm forms the basis of the MEBB for the non-stationary time series, it preserves the perfect rank correlation locally and uses the proper form of the tails in the maximum entropy distribution function[2].

2.2 *A Numerical Illustration*

In Figure 5, we report a bootstrap sample path based on the new MEBB algorithm. We use the same setup as in Section 2 with $T = 100$ and $x_0 = 0$. The original sample path and the bootstrap one are different and thus the new algorithm does not mimic the sample data too closely.

FIGURE 5

MEBB SAMPLE PATH REPLICATION

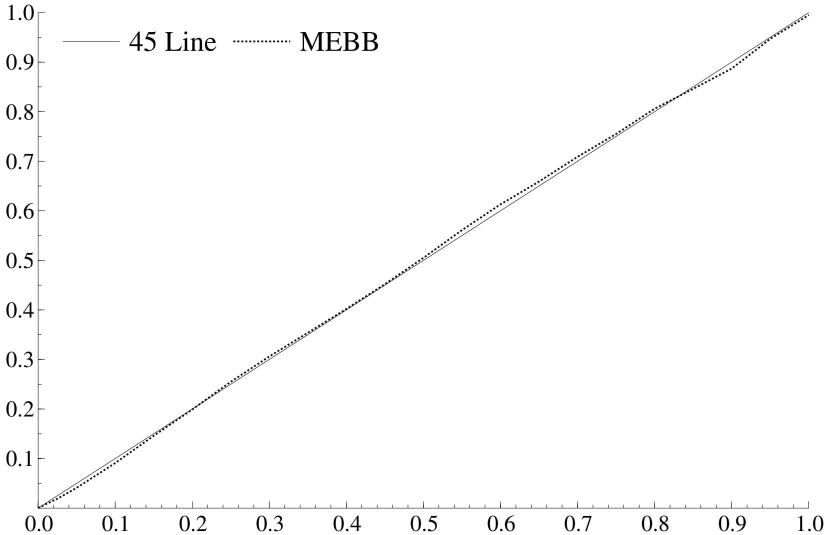
NOTE: The figure reports the sample path x_t and the replicated path x_t^* by the MEBB. The true data generating process is given as $x_t = x_{t-1} + \varepsilon_t$, with $\varepsilon_t \sim N(0,1)$, $x_0 = 0$, and $T = 100$.

2. The tails are of the exponential form, which advocates the use of the trimming algorithm. On the other hand, it is the general properties of tails which allows us to introduce in the next section the notion of the generalized non-extensive entropy.

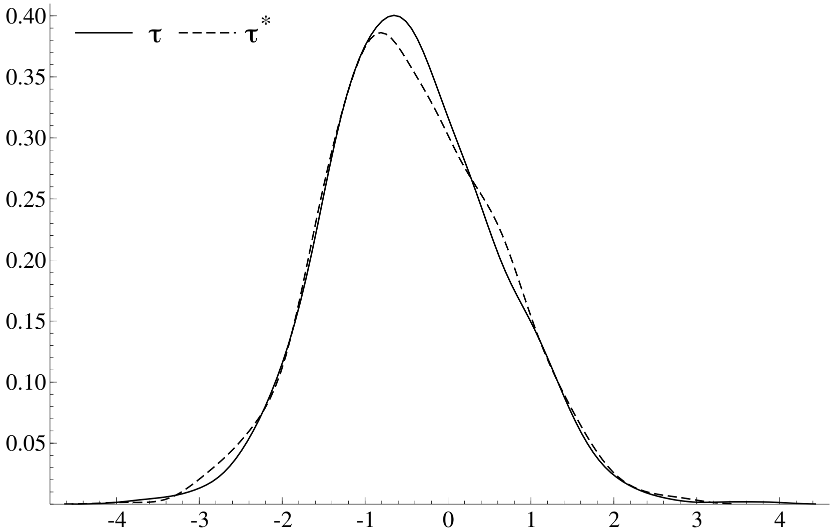
FIGURE 6

THE MEBB

(a) Empirical rejection frequencies



(b) Distribution of the bootstrapped statistic under $H_0: \rho = 1$

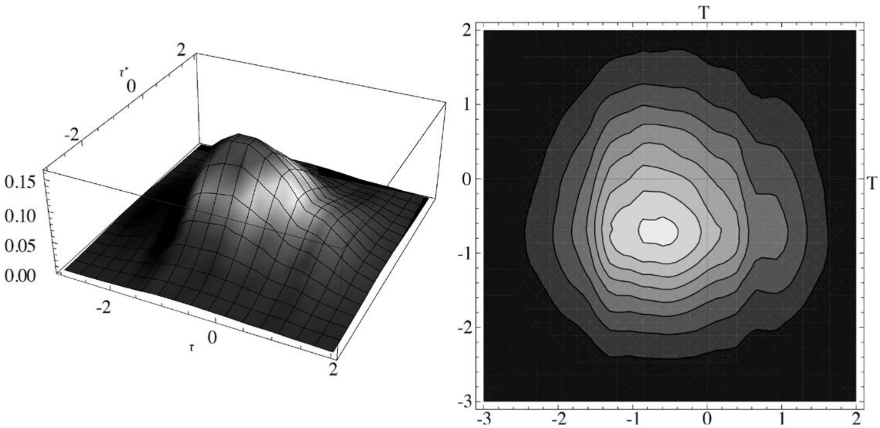


NOTE: The A panel reports the plot of the empirical rejection frequencies against the nominal levels of the test under $H_0: \rho = 1$ against $H_A: |\rho| < 1$ with the distribution approximated by the MEBB. The B panel depicts the distribution of the original t-statistic and the distribution of the t-statistic obtained by bootstrapping under $H_0: \rho = 1$.

FIGURE 7

JOINT KERNEL DENSITY

(a) Gaussian Kernel (b) Contour



In order to provide evidence that our proposed method works, we report in the left panel of Figure 6 the empirical rejection frequency of the *t*-statistic to test for $p = 1$ using the same design as in the previous section. The MEBB rejection frequencies coincide with the nominal values as required. In the right panel of Figure 6, we present the distribution of the *t*-statistic based on the simulated series as well as the distribution of the corresponding bootstrapped *t*-statistic based on the MEBB. The two distributions coincides in the same manner as for the original MEB algorithm, see Figure 3.

By comparing Figure 6 with Figure 2, it is evident that our proposed procedure allows us to restore the correct size. Indeed, a regression similar to that in (7) gives

$$\tau_i^* = \underset{(0.0354)}{-\,0.4541} + \underset{(0.0323)}{0.005663}\,\tau_i \qquad i = 1,\ldots,1000, \tag{13}$$

with the adjusted $R^2 = 0.00097$. As a further evidence that the MEBB works properly, we report in Figure 7 the counterpart of Figure 4. As expected, the absence of correlation between τ and τ^* results in a contour plot that resembles a circle rather than ellipse.

The link between the data generating process and the MEBB draws are weak in this case and therefore our method is able to preserve the perfect rank correlation locally for each block and it provides a truly viable and working entropy-based bootstrap algorithm.

3. Simulation Study of the MEBB

In this section, we carry out a comprehensive Monte Carlo simulation study to evaluate the finite sample properties of the MEBB. We compare the entropy-based resampling schemes with standard bootstrap approaches, i.e. residual-based bootstrap and non-parametric block-bootstrapping techniques. The comparison is based on the bootstrapped empirical rejection frequencies of the t-statistics for testing that the value of the AR(1) coefficient, p, estimated by OLS, equals unity (Dickey-Fuller test). The unit-root set-up is of particular interest as bootstrapping directly a non-stationary series is not common practice in the econometric literature. In particular, for non-stationary time series there exists in the literature only the *Continuous Path Block Bootstrap* (CPBB) developed by Paparoditis and Politis (2001) and studied also by Phillips (2010).

Bootstrapping unit-root tests is thus one of the potentially most interesting application of the entropy-based bootstrap method constituting an alternative to standard residual-based bootstrapped unit-root tests (see Palm *et al.* 2007 for a review) which are not easy to apply when the dependence structure of the residuals is difficult to ascertain.

3.1 *Monte Carlo Design*

We consider approximating by bootstrapping the t-statistic distribution for the AR(1) coefficient. As benchmark, we consider the empirical rejection frequencies based on the standard residuals-based bootstrap (RB) and CPBB. Then, we compute empirical rejection frequencies for the MEBB.

In particular, we generate a time series according to the following data generating process

$$x_{m,t} = \rho_0 x_{m,t-1} + \eta_{m,t}$$
$$x_{m,0} = 0 \quad t = 1,\ldots,T,$$

where $m = 1,\ldots,M$ denotes a realizations of the sample path. Further, we consider different sample sizes $T = \{50,100,300\}$ and we set $\rho_0 = 1$ in order to assess the size, while we consider progressive deviations from the unit root hypothesis, specifically $\rho_0 = \{0.99, 0.95, 0.90, 0.80, 0.70, 0.60, 0.5\}$, in order to assess the power.

Moreover, we generate the $\{\eta_{m,t}\}$ series allowing for "progressive" fat-tails by considering in turn the following distributions

$$\eta_t \overset{iid}{\sim} \begin{cases} N(0,1) \\ t(5) \\ t(3) \end{cases},$$

with $t(k)$ denoting the standard t-distribution with k degrees of freedom. Next, we fit an AR(1) model and compute the t-statistic

$$\tau_m = \frac{\hat{\rho}_m - 1}{\hat{\sigma}_{\eta_m}\left(\sum_{t=1}^{T} x_{m,t-1}^2\right)^{-1/2}}, \text{ for testing } \begin{cases} H_0: & \rho = 1 \\ H_1: & |\rho| < 1 \end{cases}, \tag{14}$$

where $\hat{\rho}_m = \left(\sum_{t=1}^{T} x_{m,t-1}^2\right)^{-1}\left(\sum_{t=1}^{T} x_{m,t-1}x_{m,t}\right)$ is the least squares estimator of the autoregressive coefficient and $\hat{\sigma}_{\eta_m} = (T-1)^{-1/2}\left(\sum_{t=1}^{T}\hat{\eta}_{m,t}^2\right)^{1/2}$ is the residuals standard deviation.

To compute the empirical rejection frequencies, we draw $b = 1,\ldots,B$ bootstrapped samples under H_0 denoted $\{x_{b,m,t}^*\}_{t=1}^T$ either by resampling from the residuals when using the RB approach or by resampling directly from levels $\{x_{m,t}\}$ when using both MEBB and CPBB.

3.1.1 *Residuals Bootstrap*

$$x_{b,m,0}^* = x_{m,0},$$

$$x_{b,m,t}^* = \rho_0 x_{b,m,t-1}^* + \eta_{b,m,t}^*, \qquad t = 1,\ldots,T,$$

where $\eta_{b,m,t}^*$ are drawn from the centred residuals $\left\{\hat{\eta}_{m,t} - \frac{1}{T}\sum_{t=1}^{T}\hat{\eta}_{m,t}\right\}_{t=1}^{T}$ obtained from the residuals of the regression of $x_{b,m,t}$ on its first lag either parametrically or non-parametrically.

3.1.2 *Continuous Path Block Bootstrap*

Paparoditis and Politis (2001), we implement the CPBB procedure as follows:

1. We compute the centred residuals

$$\hat{u}_{m,t} = x_{m,t} - x_{m,t-1} - \frac{1}{T-1}\sum_{t=1}^{T}(x_{m,t} - x_{m,t-1})$$

and define

$$\tilde{x}_{m,t} = \begin{cases} x_{m,1} & t = 1 \\ x_{m,1} + \sum_{j=2}^{t}\hat{u}_{m,j} & t = 2,\ldots,T. \end{cases}$$

The null is imposed by building the intermediate time-series $\{\tilde{x}_{m,t}\}$ generated through $\{\hat{u}_{m,t}\}$.

2. We choose the block length $\ell < T$ and let $i_0, i_1, \ldots, i_{u-1}$ i.i.d. uniform random numbers on the set $[1,2,\ldots,T-\ell]$, where $u = \lfloor T/\ell \rfloor$ (number of blocks).

FIGURE 8

EMPIRICAL REJECTION FREQUENCIES

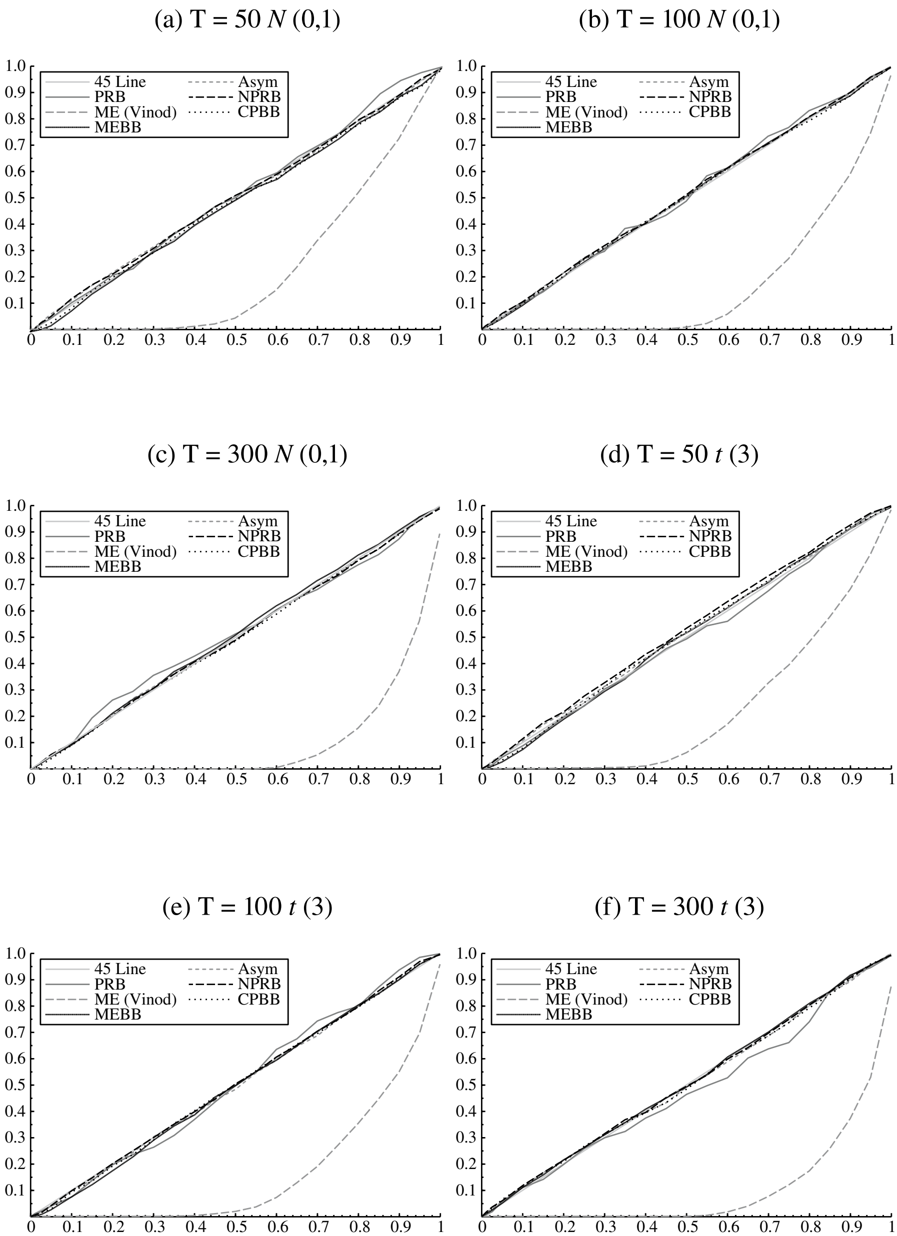
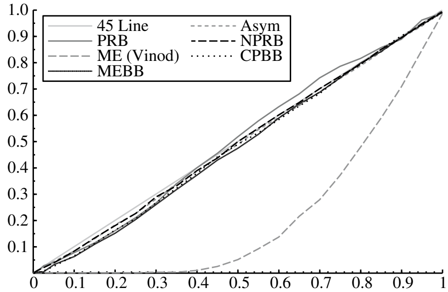
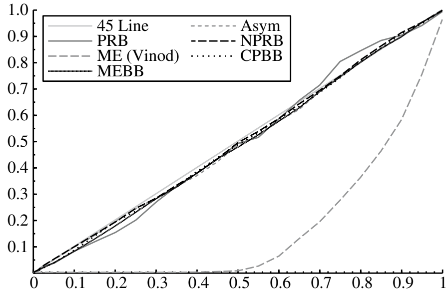


FIGURE 8 (CONTINUED)

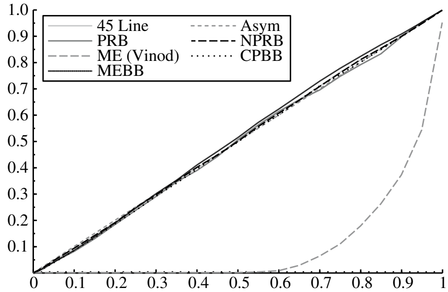
(g) T = 50 *t* (5)



(h) T = 100 *t* (5)



(i) T = 300 *t* (5)

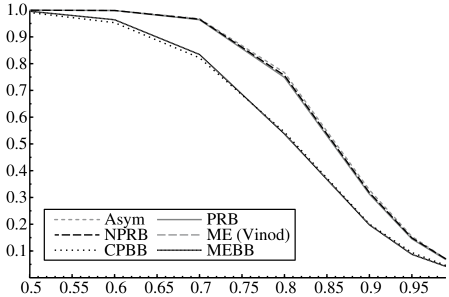


NOTE: The figure reports the empirical rejection frequencies for the MEBB and MnEBB. The data generating process is given as $x_t = x_{t-1} + \varepsilon_t$ with $\varepsilon_t \sim N(0,1)$, $t(3)$ and $t(5)$, respectively, with $x_0 = 0$, and $T = 50$, 100, and 300 respectively.

FIGURE 9

POWER OF THE TESTS

(a) T = 50 *N* (0,1)



(b) T = 100 *N* (0,1)

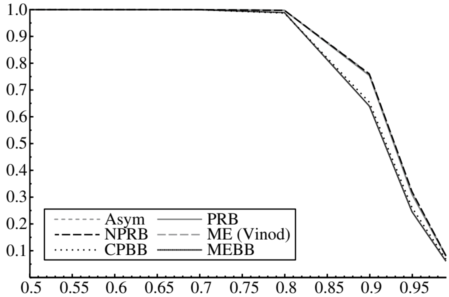
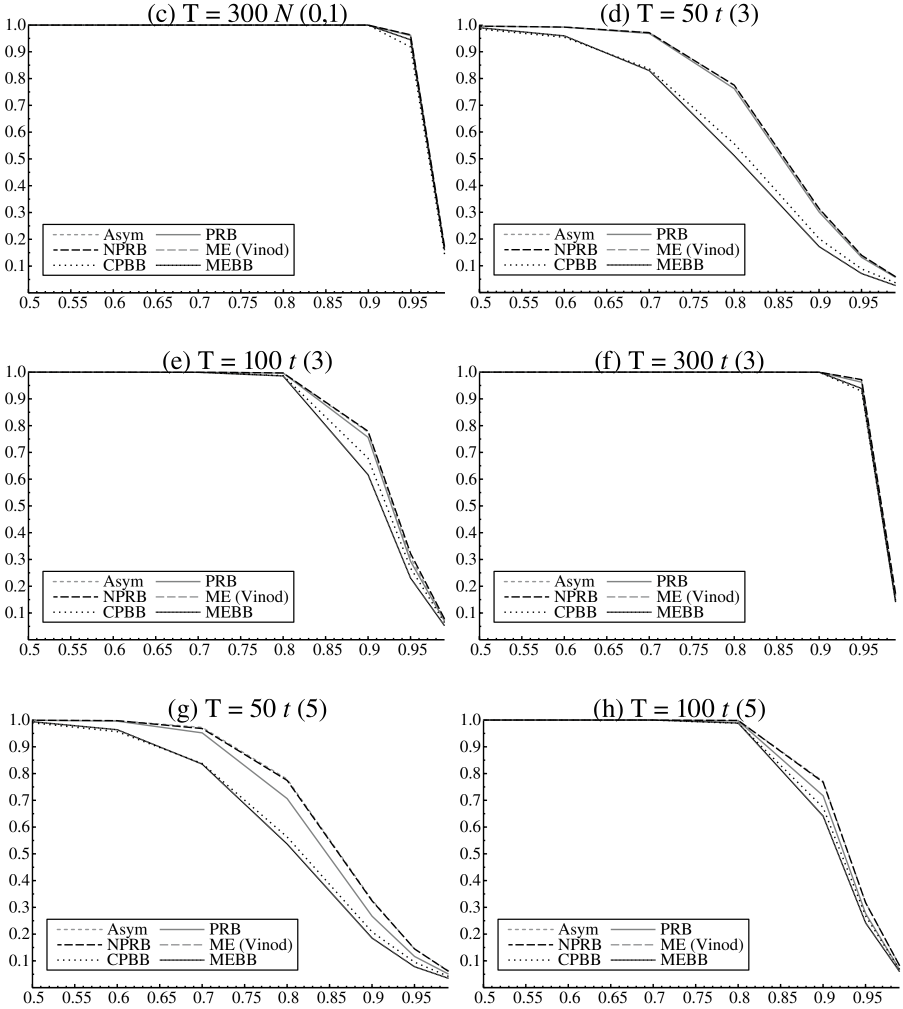
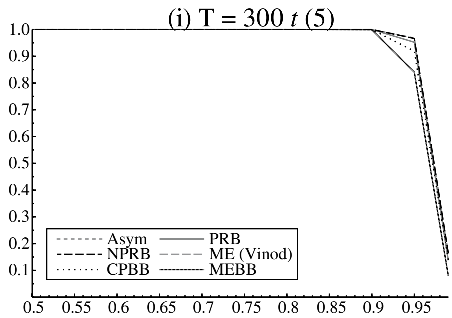


FIGURE 9 (CONTINUED)

POWER OF THE TESTS





NOTE: The figure depicts the power of the test for several bootstrap methods. The data generating process is given as $x_t = x_{t-1} + \varepsilon_t$, with $\varepsilon_t \sim N(0,1)$, $t(3)$ and $t(5)$, respectively, with $x_0 = 0$, and T = 50, 100, and 300, respectively.

3. We build the bootstrapped series of length $l = u \cdot \ell$ as

$$x^*_{b,m,j} = x_{m,1} + [\tilde{x}_{i_0+j-1} - \tilde{x}_{i_0}] \quad \text{first block},$$

$$x^*_{b,m,rs+j} = x^*_{b,m,r\ell} + [\tilde{x}_{i_m+j} - \tilde{x}_{i_r}] \quad (r+1)^{\text{th}}\,\text{block},$$

for $j = 1,\ldots,\ell$ and $r = 1,\ldots,u-1$.

3.2 *Simulation Results*

We then use $\{x^*_{b,m,t}\}_{t=1}^{T}$ to compute the bootstrapped counterpart of (14),

$$\tau^*_{m,b} = \frac{\hat{\rho}^*_{m,b} - 1}{\hat{\sigma}^*_{\eta_{m,b}} (\sum_{t=1}^{T} x^{*}_{m,b,t-1}{}^2)^{-1/2}} \qquad b = 1,\ldots,B, \tag{15}$$

and we select the α-quantile $\tau^*_m(\alpha)$ of the distribution of the bootstrapped statistic (at the *mth* iteration) such that $B^{-1}\sum_{b=1}^{B} I(\tau^*_{m,b} \le \tau^*_m(\alpha)) \approx \alpha$. The empirical rejection frequencies are computed as

$$\frac{1}{M}\sum_{m=1}^{M} I(\tau_m \le \tau^*_m(\alpha)) \tag{16}$$

being a one-sided test with rejection to the left.

To compare the size of the different approaches, we compute (16) for $\alpha \in [0.01, 0.025, 0.05, 0.10, 0.20, \ldots, 1]$ and we plot the calculated values against α. If the bootstrap works well, as we are simulating under H_0, we should observe a graph close to the 45^0 line.

Figure 8 reports the rejection frequencies under H_0 (size) for the bootstrap methods introduced in the previous paragraphs. For the CPBB and MEBB, the block length is set to $u = \lfloor T^{1/3} \rfloor$. Overall, the figures suggest that even for small samples with $T = 50$, our proposed MEBB provides rejection frequencies close to the nominal level, outperforming the original MEB.

Figure 9 reports the power curves for the alternative bootstrap procedures. The closest method to the MEBB is the CPBB, which shows also very similar power to our proposed method. In conclusion, the MEBB provides a valid alternative to the (unique) existing bootstrap method, which allows to replicate the levels for non-stationary time series.

In the next section, we extend the MEBB and introduce the Maximum non-extensive Entropy Block Bootstrap, which is based on the generalization of the non-extensive entropy.

4. Beyond the Shannon Entropy: The Maximum Non-extensive Entropy Block Bootstrap

We extend the Maximum Entropy Block Bootstrap to the case when the underlying principle is driven by maximization of the non-extensive entropy and define the Maximum non-extensive Entropy Block Bootstrap (MnEBB)[3].

The key concept of our framework is the generalized Tsallis (1988) possible entropy defined in the discrete form as

$$H_q = -\frac{1}{1-q}\left(1-\sum_{i=1}^{N}(p_i)^q\right),$$

and in the continuous form as

$$H_q = -\frac{1}{1-q}\left(1-\int dx\left(p(x)\right)^q\right),$$

where the parameter q governs the non-extensiveness of the system. The Tsallis entropy converges to the Shannon entropy in the limit when $q \to 1$.

For a given q, the density function f_q is given as

$$f_q(x) = \frac{\left[1-\beta(1-q)x\right]^{1/(1-q)}}{Z_q}, \tag{17}$$

with normalization constant

$$Z_q = \int dx\left[1-\beta(1-q)x\right]^{1/(1-q)}.$$

The non-extensiveness of the system can be expressed for two systems A and B as

$$H_q(A+B) = H_q(A) + H_q(B) + (1-q)H_q(A)H_q(B).$$

The density function resulting from the optimization of the Tsallis entropy has a finite integral over the semi-definite interval for $q \in (1,2)$. The limit of this distribution function is exponential function for $q \to 1$. For $q \in (1,2)$, we get power law behaviour and thus fatter tails for the distribution. For $q < 1$, we get a non-standard behaviour of the distribution function as it is infinite over the semi-definite interval and thus it requires a normalization by an infinite normalization factor. This still allows for a comparison between the two draws as $\infty/\infty \sim c$, but it provides unnecessary complications. As $q \to 0$, we get uniform distribution, or, $f_{q \to 0}(x) = c/\infty$, where c does not depend on x. For $q > 5/3$, we get a distribution with non-existing

3. The MEBB presented in Section 2 is based on the Shannon entropy. Such an entropy is suitable for systems, which are extensive and thus do not depend on the initial conditions. Furthermore, extensive systems are additive, where two independent pieces of information additively sum up. Tsallis (1988) introduces the concept of the non-extensive entropy, where the information is not simply additive and the system itself depends on the initial conditions. In addition, the non-extensive systems leads to the power law behaviour and fat tails, which is a feature present in time series found in economics and finance.

second moment, or $E\left[x^2\right] = \infty$[4]. At $q = 2$, the first moment cease to exist. In this paper, we explicitly consider $q \in [1,5/3)$, which covers a broad range specifications ranging from the standard Shannon entropy to cases with non-existent second moments and fat tails.

4.1 *The MnEBB Algorithm*

The MnEBB algorithm is defined as follows. First, the mass preserving constraint states that, on average, $1/T$ of the mass of the density function lies in each of the intervals I_t. This is achieved by bootstrapping the set x_t as a T draws, each from the different interval I_t. Second, the mean preserving constraint says that

$$\sum_{t=1}^{T} x_t = \sum_{t=1}^{T} x_{(t)} = \sum_{t=1}^{T} m_t,$$

where m_t is the mean of $f(x)$ over the interval I_t.

The new framework thus leads to the following choice of the density functions

$$f_q(x) = \alpha_q \left(1 - \beta(1-q)x\right)^{\frac{1}{1-q}}, \quad x \in I_1 \; m_1 = \frac{3x_{(1)}}{4} + \frac{x_{(2)}}{4}, \tag{18}$$

$$\alpha_q : \int_{I_1} x dx f_q(x) = m_1 , \tag{19}$$

$$f_q(x) = \frac{1}{z_k - z_{k-1}}, \quad x \in I_k\big|_{k=2,\ldots,T-1} \; m_k = \frac{x_{(k-1)}}{4} + \frac{x_{(k)}}{2} + \frac{x_{(k+1)}}{4}, \tag{20}$$

$$f_q(x) = \omega_q \left(1 - \beta(1-q)x\right)^{\frac{1}{1-q}}, \quad x \in I_T \; m_T = \frac{x_{(T-1)}}{4} + \frac{3x_{(T)}}{4} , \tag{21}$$

$$\omega_q : \int_{I_T} x dx f_q(x) = m_T . \tag{22}$$

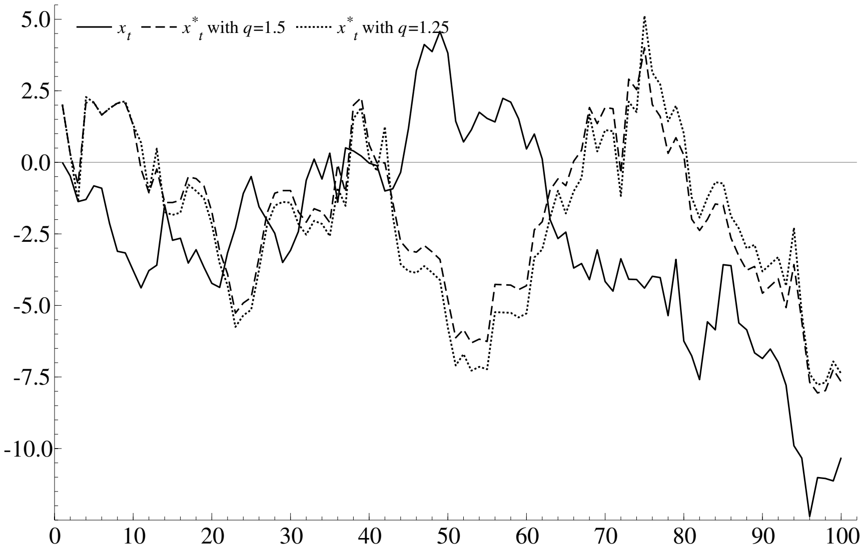
The remaining structure of the MnEBB algorithm is just as defined for the MEBB. The choice of $q > 1$ suggests using the distribution with fatter tails than implied by the standard entropy.

Figure 10 reports a bootstrap sample path based on the new MnEBB algorithm. We use the same setup as in the previous section with $T = 100$, $y_0 = 0$ and two values of non-extensiveness parameter: $q = 1.25$ and $q = 1.5$. The figure supports the intuition that more the non-extensive the system is, the more variation in the replication sample is present.

4. In general, it would be more appropriate to deal with the q-expectations, which remain finite and have a proper meaning. However, this is goes beyond the scope of this paper.

FIGURE 10

MNEBB SAMPLE PATH REPLICATION



NOTE: The figure reports the sample path x_t and the replicated path x_t^* by the MnEBB. The short-dash line corresponds to MnEBB with q = 1.25 and dot line corresponds to MnEBB with q = 1.5. The true data generating process is given as $x_t = x_{t-1} + \varepsilon_t$, with $\varepsilon_t \sim N(0,1)$, $x_0 = 0$, and T = 100.

4.2 *Simulation Results*

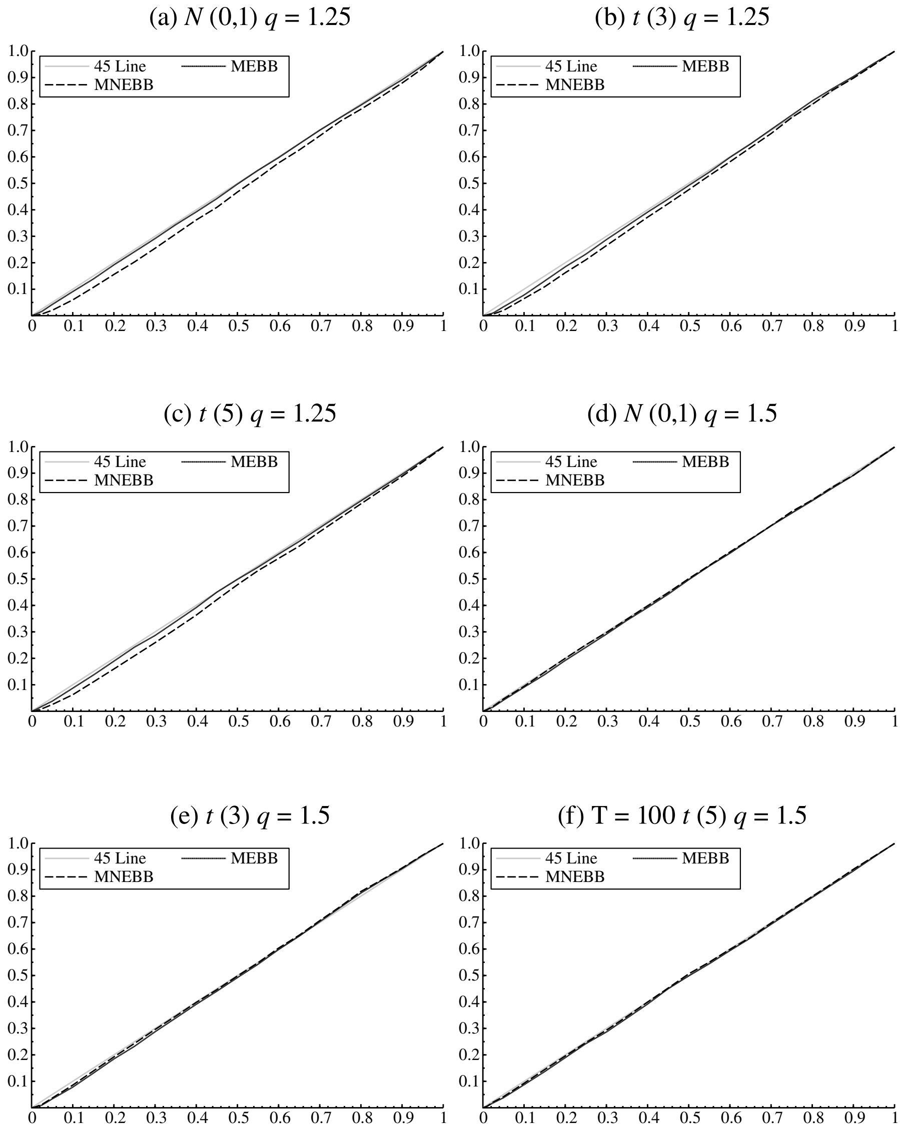
We replicate the extensive simulation analysis as in Section 4 and focus on the comparison between the MEBB and the MnEBB with q = 1.25 and q = 1.5. Note that the MEBB corresponds to the case of the MnEBB with q = 1.

Figure 11 reports the rejection frequencies for three MnEBB methods with q = 1, 1.25, and 1.5, respectively, introduced in the previous paragraphs. The data generating process is given as $x_t = x_{t-1} + \varepsilon_t$, with $\varepsilon_t \sim N(0,1)$, $t(3)$ and $t(5)$, respectively, with $x_0 = 0$, and T = 100. Overall, the figures suggest that for q = 1.25, the MnEBB underperforms in terms of size the MEBB. However, with increasing q, the MEBB and MnEBB are becoming indistinguishable.

Figure 12 reports the power of the alternative bootstrap procedures. The power supports the inferiority of MnEBB with q = 1.25 to the original MEBB, while the MnEBB with q = 1.5 is dominating in terms of the power. Therefore, the MnEBB method with fatter tails provides a significant improvement to the entropy based algorithm even for the data generating process based on the Gaussian distribution.

FIGURE 11

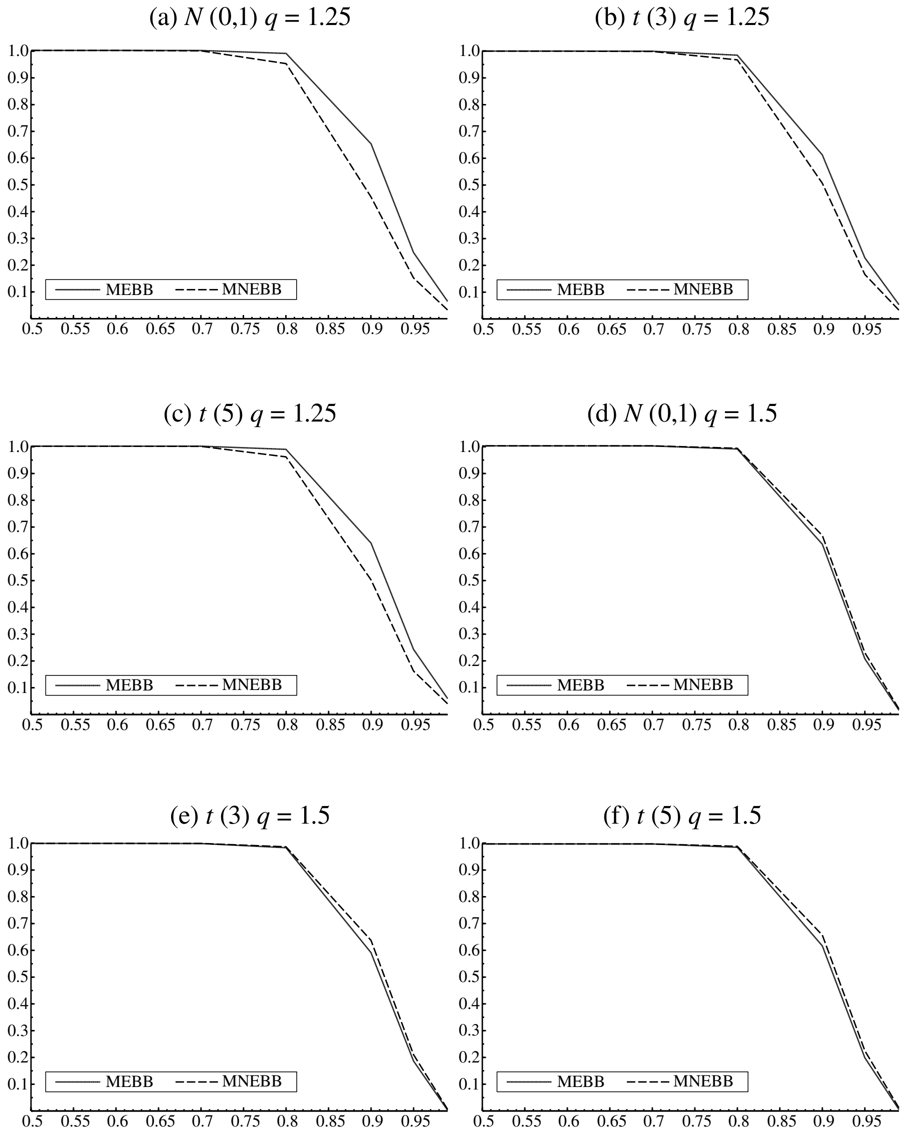
NON-EXTENSIVE SIZE



NOTE: The figure reports the empirical rejection frequencies for the MEBB and MnEBB. The data generating process is given as $x_t = x_{t-1} + \varepsilon_t$, with $\varepsilon_t \sim N(0,1)$, $t(3)$ and $t(5)$, respectively, with $x_0 = 0$, and T = 100.

FIGURE 12

NON-EXTENSIVE POWER



NOTE: The figure depicts the power of the test for the MEBB and MnEBB. The data generating process is given as $x_t = x_{t-1} + \varepsilon_t$ with $\varepsilon_t \sim N(0,1)$, $t(3)$ and $t(5)$, respectively, with $x_0 = 0$, and $T = 100$.

Conclusion

In this paper, we proposed the Maximum Entropy Block Bootstrap, a fully non-parametric bootstrap procedure, to resample directly from time series with a general persistence structure. Our procedure employed the maximum entropy bootstrap while preserving locally the rank correlation between the original sample and the bootstrap draws. We used unit root test to illustrate that our procedure performs well. In addition, we employed the notion of non-extensive entropy and introduce the Maximum non-extensive Entropy Bootstrap, which allows for the inclusion of fat tails and power-law behaviour. This generalized procedure outperforms the Maximum Entropy Bootstrap for large values of the non-extensiveness even when the underlying data generating process involves the normal distribution.

The results in this paper suggest some interesting developments. First, it would be useful to derive the limiting theory of the MEBB and MnEBB methods proposed in this paper. Second, it would be interesting to extend the proposed procedure to a non-stationary framework such as co-integration analysis. This is part of an ongoing research agenda.

REFERENCES

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

DAVIDSON, R. and A. MONTICINI (2014). "Heteroskedasticity-and-autocorrelation-consistent Bootstrapping." Technical report, Università Cattolica del Sacro Cuore, Dipartimenti e Istituti di Scienze Economiche (DISCE).

EFRON, B. (1979). "Bootstrap Methods: Another Look at the Jackknife." *Annals of Statistics*, 7: 1–26.

KÜNSCH, H. (1989). "The Jack-knife and the Bootstrap for General Stationary Observations." *Annals of Statistics*, 17: 1217–1241.

MAMMEN, E. (1993). "Bootstrap and Wild Bootstrap for High Dimensional Linear Models." *Annals of Statistics*, 21: 255–285.

PALM, F. C., S. SMEEKES, and J.-P. URBAIN (2007). "Bootstrap Unit-Root Tests: Comparisons and Extensions." *Journal of Time Series Analysis*, 29: 371–401.

PAPARODITIS, E., and D. POLITIS (2001). "The Continuous Path Block-Bootstrap." In PURI, M. (Ed.), *Asymptotics in Statistics and Probability*, pp. 305–320. VSP Publications.

PHILLIPS, P. C. B. (2010). "Bootstrapping $I(1)$ Data." *Journal of Econometrics*, 158: 280–284.

POLITIS, D., J. ROMANO, and M. WOLF (1999). *Subsampling*. Springer Series in Statistics. Springer New York.

TSALLIS, C. (1988). "Possible Generalization of Boltzmann-Gibbs Statistics." *Journal of Statistical Physics*, 52: 479–487.

VINOD, H. D. and J. LÓPEZ-DE LACALLE (2009). "Maximum Entropy Bootstrap for Time Series: The Meboot R Package." *Journal of Statistical Software*, 29: 1–19.

WU, C.-F. J. (1986). "Jackknife, Bootstrap and Other Resampling Methods in Regression Analysis." *Annals of Statistics*, 14: 1261–1295.