Robust Resampling Methods for Time Series*

Lorenzo Camponovo

University of Lugano

Olivier Scaillet

Université de Genève and Swiss Finance Institute

Fabio Trojani

University of Lugano and Swiss Finance Institute

First Version: November 2007; This Version: January 26, 2010

Abstract

We study the robustness of block resampling procedures for time series. We first derive a set of formulas to quantify their quantile breakdown point. For the block bootstrap and the subsampling, we find a very low quantile breakdown point. A similar robustness problem arises in relation to data-driven methods for selecting the block size in applications, which can render inferences based on standard resampling methods useless already in simple estimation and testing settings. To solve this problem, we introduce a robust fast resampling scheme that is applicable to a wide class of time series settings. Monte Carlo simulations and sensitivity analysis for the simple AR(1) model confirm the dramatic fragility of classical resampling procedures in presence of contaminations by outliers. They also show the better accuracy and efficiency of the robust resampling approach under different types of data constellations. A real data application to testing for stock returns predictability shows that our robust approach can detect predictability structures more consistently than classical methods.

Keywords: Subsampling, bootstrap, breakdown point, robustness, time series.
JEL: C12, C13, C15.
MSC 2000: Primary 62F40; Secondary 62F35.

*The authors acknowledge the financial support of the Swiss National Science Foundation (NCCR FINRISK and grants 101312-103781/1, 100012-105745/1, and PDFM1-114533). We thank participants at the International Conference on Computational Management Science 2009 in Geneva, the International Conference on Robust Statistics 2009 in Parma and the International Conference on Computational and Financial Econometrics 2009 in Limassol for helpful comments. Correspondence address: Fabio Trojani, Faculty of Economics, University of Lugano, Via Buffi 13, CH-6900 Lugano, e-mail: Fabio.Trojani@usi.ch.

1 Introduction

Resampling methods, including the bootstrap (see, e.g., Hall, 1992, Efron and Tibshirani, 1993, and Hall and Horowitz, 1996) and the subsampling (see, e.g., Politis and Romano, 1992, 1994a, Politis, Romano and Wolf, 1999), are useful tools in modern statistics and econometrics. The simpler consistency conditions and the wider applicability in some cases (see, e.g., Andrews, 2000, and Bickel, Gotze and van Zwet, 1997) have made the subsampling a useful and valid alternative to the bootstrap in a number of statistical models. Bootstrap and subsampling procedures for time series typically rely on different block resampling schemes, in which selected sub-blocks of the data, having size strictly less than the sample size, are randomly resampled. This feature is necessary in order to derive consistent resampling schemes under different assumptions on the asymptotically vanishing time series dependence between observations. See, among others, Hall (1985), Carlstein (1986), Künsch (1989), and Politis, Romano and Wolf (1999).

The low robustness of classical bootstrap and subsampling methods is a known feature in the iid setting; see, among others, Singh (1998), Salibian-Barrera and Zamar (2002), Salibian-Barrera, Van Aelst and Willems (2006,2007), and Camponovo, Scaillet and Trojani (2009). These papers study global robustness features and highlight a typically very low breakdown point of classical bootstrap and subsampling quantiles. Essentially, the breakdown point quantifies the smallest fraction of outliers in the data which makes a statistic meaningless. Therefore, standard iid resampling methods produce estimated quantiles that are heavily dependent on a few possible outliers in the original data. Intuitively, this lack of robustness is related to the (typically high) probability of resampling a large number of outliers in a random sample using an iid bootstrap or subsampling scheme. To overcome this problem, robust bootstrap and subsampling approaches with desirable quantile breakdown point properties have been developed in the iid context by Salibian-Barrera and Zamar (2002), Salibian-Barrera, Van Aelst and Willems (2006,2007), and Camponovo, Scaillet and Trojani (2009), among others.

In this paper, we study the robustness of block resampling methods for time series and we develop fast robust resampling approaches that are applicable to a variety of time series models. We first characterize the breakdown properties of block resampling procedures for time series by deriving upper bounds for their quantile breakdown point; these results cover both overlapping and nonoverlapping bootstrap and subsampling procedures. Concrete computations show that block resampling methods for time series suffer of an even larger robustness problem than in the iid context. In the extreme case, a single outlier in the original sample can dramatically affect the accuracy of block resampling methods and make the resulting inference effectively useless. This problem cannot be mitigated simply by applying standard block resampling methods to a more robust statistic, indicating the high need for a more robust resampling scheme applicable in the time series context.

We develop our robust resampling approach for time series following the fast resampling idea putted forward, among others, in Shao and Tu (1995), Davidson and McKinnon (1999), Hu and Kalbfleisch (2000), Andrews (2002), Salibian-Barrera and Zamar (2002), Goncalves and White (2004), Hong and Scaillet (2006), Salibian-Barrera, Van Aelst and Willems (2006,2007), and Camponovo, Scaillet and Trojani (2009). Our resampling method is applicable to a wide class of resampling procedures, including both the block bootstrap and the subsampling, and it provides robust estimation and inference results under weak conditions. Moreover, it inherits the low computational cost of fast resampling approaches. This makes it applicable to nonlinear models when classical methods might become computationally too expensive, or in combination with computationally intensive data-driven procedures for the selection of the optimal block size; see, for instance, Sakata and White (1998), Ronchetti and Trojani (2001), Mancini, Ronchetti and Trojani (2005), Ortelli and Trojani (2005), and Muler and Yohai (2008) for recent examples of robust estimators for nonlinear time series models. By means of explicit breakdown point computations, we also find that the better breakdown properties of our fast robust resampling scheme are inherited by data-driven choices of the block size based on either the minimum confidence index volatility (MCIV) and the calibration method (CM), proposed in Romano and Wolf (2001) for the subsampling, or the data-driven method in Hall, Horowitz and Jing (1995) (HHJ) for the moving block bootstrap.

We investigate by Monte Carlo simulations the performance of our robust resampling approach in the benchmark context of the estimation of the autoregressive parameter in an AR(1) model both in a strictly stationary and near-to-unit root setting. Overall, our Monte Carlo experiments highlight a dramatic fragility of classical resampling methods in presence of contaminations by outliers, and a more reliable and efficient inference produced by our robust resampling method under different types of data constellations. Finally, in an application to real data, we find that our robust resampling approach detects predictability structures in stock returns more consistently than standard methods.

The paper is organized as follows. Section 2 oulines the main setting and introduces the quantile breakdown point formulas of different block resampling procedures. In Section 3 we develop our robust approach and derive the relevant expression for the associated quantile breakdown point formula. We show that, under weak conditions, the resulting quantile breakdown point is maximal. In Section 4, we study the robustness properties of data-driven block size selection procedures based on the MCIV, the CM and the HHJ method. Monte Carlo experiments, sensitivity analysis and the empirical application to stock returns predictability are presented in Section 5. Section 6 concludes.

2 Resampling Distribution Breakdown Point Quantile

We start our analysis by characterizing the robustness of resampling procedures for time series and by deriving formulas for their quantile breakdown point.

2.1 Definition

Let $X_{(n)} = (X_1, \ldots, X_n)$ be a sample from a real valued stationary process $\mathcal{X} = \{X_t, t \in \mathbb{Z}\}$ defined on the probability space $(\Omega, \mathcal{F}, P)$, and consider a real valued statistic $T_n := T(X_{(n)})$.

In the time series setting, block bootstrap procedures split the original sample in overlapping or nonoverlapping blocks of size $m < n$. Then, new random samples of size n are constructed assuming an approximate independence between blocks. Finally, the statistic T is applied to the so generated random samples; see, e.g., Hall (1985), Carlstein (1986), Künsch (1989), and Andrews (2004). The more recent subsampling method (see, e.g., Politis, Romano and Wolf, 1999), instead, directly applies statistic T to overlapping or nonoverlapping blocks of size m strictly less than n.

Let $X^*_{(k)} = (X^*_1, \ldots, X^*_k)$ denote for brevity a bootstrap ($k = n$) or a subsampling ($k = m < n$) random sample and $T^*_{n,k} := T(X^*_{(k)})$ be the bootstrap or subampling statistic, respectively. Then, for $t \in (0, 1)$, the quantile Q^*_t of $T^*_{n,k}$ is defined by

$$Q^*_t = \inf\{x | P^*(T^*_{n,k} \leq x) \geq t\}, \tag{1}$$

where P^* is the corresponding bootstrap or subsampling distribution and, by definition, $\inf(\emptyset) = \infty$.

We characterize the robustness of quantile (1) via its breakdown point, i.e., the smallest fraction of outliers in the original sample such that Q^*_t degenerates, making inference based on (1) meaningless. Different than in the iid case, in time series we can consider different possible models of contamination by outliers, like for instance additive outliers, replacement outliers and innovation outliers; see, e.g., Martin and Yohai (1986). Because of this additional complexity, we first introduce a notation that can better capture the effect of such contaminations, following Genton and Lucas (2003). Denote by $\mathcal{Z}^\zeta_p$ the set of all n-components outlier samples, where p is the number of outliers and index $\zeta \in \bar{R}$ indicates their size. When $p > 1$ we do not necessarily assume outliers $\zeta_1, \ldots, \zeta_p$ to be all equal to ζ, but we rather assume existence of constants $c_1, \ldots, c_p$, such that $\zeta_i = c_i \zeta$.

Let $0 \leq b \leq 0.5$ be the upper breakdown point of statistic T_n, i.e., nb is the smallest number of outliers such that $T(X_{(n)} + Z^{\zeta}_{nb}) = +\infty$ for some $Z^{\zeta}_{nb} \in \mathcal{Z}^{\zeta}_{nb}$. Breakdown point b is an intrinsic characteristic of a statistic. It is explicitly known in some cases and it can be gauged most of the time, for instance by means of simulations and sensitivity analysis. In this section, we focus for brevity on one-dimensional real valued statistics. As discussed for instance by Singh (1998) in the iid context, our quantile breakdown point results for time series can be naturally extended to consider multivariate and scale statistics. Formally, the quantile breakdown point of Q^*_t is defined as follows:

Definition 1 *The upper breakdown point of the t-quantile Q^*_t is given by*

$$b_t = \frac{1}{n} \cdot \left[\inf_{\{1 \leq p \leq \lceil n/2 \rceil\}} \left\{ p \middle| \text{there exists } Z^{\zeta}_p \in \mathcal{Z}^{\zeta}_p \text{ such that } Q^*_t(X_{(n)} + Z^{\zeta}_p) = +\infty \right\} \right], \quad (2)$$

where $\lceil x \rceil = \inf\{n \in \mathbb{N} | x \leq n\}$.

2.2 Quantile Breakdown Point

We derive formulas for the quantile breakdown point of the overlapping subsampling and both nonoverlapping and overlapping moving block bootstrap procedures. Similar results can be obtained for the nonoverlapping subsampling. Since that case is of little practical interest, because unless the sample size is very large the number of blocks is too small to make reliable inference, we do not report results for this case. For brevity, we denote by b^K_t, $K = OS, NB, OB$, the upper t-quantile breakdown point of the overlapping subsampling and the nonoverlapping and overlapping moving block bootstrap, respectively. Results for the overlapping moving block bootstrap can be modified to cover asymptotically equivalent variations such as the stationary bootstrap of Politis and Romano (1994b).

2.2.1 Subsampling

For simplicity, let $n/m = r \in \mathbb{N}$. The overlapping subsampling splits the original sample $X_{(n)} = (X_1, \ldots, X_n)$ into $n - m + 1$ overlapping blocks $(X_i, \ldots, X_{i+m-1})$, $i = 1, \ldots, n - m + 1$. Finally, it applies statistic T to these blocks.

Theorem 2 *Let b be the breakdown point of T_n and $t \in (0, 1)$. The quantile breakdown point of overlapping subsampling procedures satisfies the following property:*

$$b^{OS}_t \leq \inf_{\{p \in \mathbb{N}, p \leq r-1\}} \left\{ p \cdot \frac{\lceil mb \rceil}{n} \middle| p > \frac{(1-t)(n-m+1) + \lceil mb \rceil - 1}{m} \right\}. \quad (3)$$

The term $\frac{(1-t)(n-m+1)}{m}$ represents the number of degenerated statistics necessary in order to cause the breakdown of Q_t^*, while $\frac{\lceil mb \rceil}{n}$ is the fraction of outliers which is sufficient to cause the breakdown of statistic T in a block of size m. In time series, the number of possible subsampling blocks of size m is typically lower than the number of iid subsamples of size m. Therefore, the breakdown of a statistic in one random block tends to have a larger impact on the subsampling quantile than in the iid case. Intuitively, this feature implies a lower breakdown point of subsampling quantiles in time series than in iid settings. Table 1 confirms this basic intuition. Using Theorem 2 we compute the breakdown point of the overlapping subsampling quantile for a sample size $n = 120$, for $b = 0.5$ and for block sizes $m = 5, 10, 15$. We see that even for a maximal breakdown point statistic ($b = 0.5$), the overlapping subsampling implies a very low quantile breakdown point, which is increasing in the block size, but very far from the maximal value $b = 0.5$. Moreover, this breakdown point is clearly lower than in the iid case; see Camponovo, Scaillet and Trojani (2009). For instance, for $m = 10$, the 0.95-quantile breakdown point of the overlapping subsampling is lower than 0.05, which is less than a quarter of the breakdown point of 0.23 for the same block size in the iid setting.

2.2.2 Moving Block Bootstrap

Let $X_{(m),i}^N = (X_{(i-1)\cdot m+1}, \ldots, X_{i\cdot m})$, $i = 1, \ldots, r$, be the r nonoverlapping blocks of size m. The nonoverlapping moving block bootstrap selects randomly with replacement r nonoverlapping blocks $X_{(m),i}^{N*}$, $i = 1, \ldots, r$. Then, it applies statistic T to the n-sample $X_{(n)}^{N*} = (X_{(m),1}^{N*}, \ldots, X_{(m),r}^{N*})$. Similarly, let $X_{(m),i}^O = (X_i, \ldots, X_{i+m-1})$, $i = 1, \ldots, n-m+1$, be the $n-m+1$ overlapping blocks. The overlapping moving block bootstrap selects randomly with replacement r overlapping blocks $X_{(m),i}^{O*}$, $i = 1, \ldots, r$. Then, it applies statistic T to the n-sample $X_{(n)}^{O*} = (X_{(m),1}^{O*}, \ldots, X_{(m),r}^{O*})$.

Theorem 3 *Let b be the breakdown point of T_n and $t \in (0,1)$. The quantile breakdown points b_t^{NB} and b_t^{OB} of the nonoverlapping and overlapping moving block bootstrap respectively satisfy the following properties:*

(i) $b_t^{NB} \leq \frac{1}{n} \cdot \left[\inf_{\{p_1,p_2 \in \mathbb{N}, p_1 \leq m, p_2 \leq r\}} \left\{ p = p_1 \cdot p_2 \middle| P\left(BIN\left(r, \frac{p_2}{r}\right) > \frac{nb}{p_1}\right) > 1-t \right\} \right],$

(ii) $b_t^{OB} \leq \frac{1}{n} \cdot \left[\inf_{\{p_1,p_2 \in \mathbb{N}, p_1 \leq m, p_2 \leq r\}} \left\{ p = p_1 \cdot p_2 \middle| P\left(BIN\left(r, \frac{mp_2-p_1+1}{n-m+1}\right) > \frac{nb}{p_1}\right) > 1-t \right\} \right].$

Similar to the findings for the subsampling, the right part of (i) and (ii) are similar for large $n >> m$. Indeed, (ii) implies $\frac{mp_2-p_1+1}{n-m+1} \approx \frac{mp_2}{n} = \frac{p_2}{r}$, which is the right part of (i). Further the breakdown point formula for the iid bootstrap in Singh (1998) emerges as a special case of the formulas in Theorem 3, for $m = 1$. This is intuitive: a nonoverlapping moving block bootstrap with block size

m is essentially an iid bootstrap based on a sample of size r, in which each block of size m corresponds to a single random realization in the iid bootstrap. As for the subsampling, the reduction in the number of possible blocks when $m \neq 1$ increases the potential impact of a contamination and it implies a lower quantile breakdown point. In Table 1, we compute the breakdown point of the nonoverlapping and overlapping moving block bootstrap quantile for $n = 120$, $b = 0.5$ and block sizes $m = 5, 10, 15$. These breakdown points are decreasing in the block size. Again, they are far from the maximal value $b = 0.5$. For instance, for $m = 15$ the 0.99 quantile breakdown point is less than 0.2126, which is approximatively half the breakdown point of 0.392 in the iid setting.

3 Robust Resampling Procedures

The results in the last section show that, even using statistics with maximal breakdown point, classical block resampling procedures imply a low quantile breakdown point. To overcome this problem it is necessary to introduce a different and more robust resampling approach. We develop such robust resampling methods for M-estimators, starting from the fast resampling approach studied, among others, in Shao and Tu (1995), Davidson and McKinnon (1999), Hu and Kalbfleisch (2000), Andrews (2002), Salibian-Barrera and Zamar (2002), Goncalves and White (2004), Hong and Scaillet (2006), Salibian-Barrera, Van Aelst and Willems (2006,2007), and Camponovo, Scaillet and Trojani (2009).

3.1 Definition

Given the original sample $X_{(n)} = (X_1, \ldots, X_n)$, we consider the class of robust M-estimators $\hat{\theta}_n$ for parameter $\theta \in \mathbb{R}^d$, defined as the solution of the equations:

$$\psi_n(X_{(n)}, \hat{\theta}_n) := \frac{1}{n-q+1} \sum_{i=q}^{n} g(X_{i-q+1}, \ldots, X_i; \hat{\theta}_n) = 0, \tag{4}$$

where $\psi_n(X_{(n)}, \cdot) : \mathbb{R}^d \to \mathbb{R}^d$ depends on parameter θ and a bounded estimating function g. Boundedness of estimating function g is a characterizing feature of robust M-estimators. Standard block resampling approaches need to solve equation $\psi_k(X^*_{(k)}, \hat{\theta}^*_k) = 0$ for each bootstrap ($k = n$) or subsampling ($k = m < n$) random sample $X^*_{(k)}$. Instead, we consider the following Taylor expansion of (4) around the true parameter θ_0:

$$\hat{\theta}_n - \theta_0 = -[\nabla_\theta \psi_n(X_{(n)}, \theta_0)]^{-1} \psi_n(X_{(n)}, \theta_0) + o_p(1), \tag{5}$$

where $\nabla_\theta \psi_n(X_{(n)}, \theta_0)$ denotes the derivative of function ψ_n with respect to θ. Based on this expansion, we use $-[\nabla_\theta \psi_n(X_{(n)}, \hat{\theta}_n)]^{-1}\psi_k(X^*_{(k)}, \hat{\theta}_n)$ as an approximation of $\hat{\theta}^*_k - \hat{\theta}_n$ in the definition of the resampling scheme estimating the sampling distribution of $\hat{\theta}_n - \theta_0$.

Given a normalization constant τ_n, a robust fast resampling distribution for $\tau_n(\hat{\theta}_n - \theta_0)$ is defined by

$$L^{RF*}_{n,m}(x) = \frac{1}{N}\sum_{s=1}^{N} \mathbb{I}\big(\tau_k(-[\nabla_\theta \psi_n(X_{(n)}, \hat{\theta}_n)]^{-1}\psi_k(X^*_{(k),s}, \hat{\theta}_n)) \leq x\big), \tag{6}$$

where $\mathbb{I}(\cdot)$ is the indicator function and s indexes the N possible random samples generated by subsampling and bootstrap procedures, respectively. The main assumptions under which the fast resampling distribution (6) consistently estimates the unknown sampling distribution of $\tau_n(\hat{\theta}_n - \theta_0)$ in a time series context are given, e.g., in Hong and Scaillet (2006) for the subsampling (Assumption 1) and in Goncalves and White (2004) for the bootstrap (Assumption A and Assumptions 2.1 and 2.2).

3.2 Robust Resampling Methods and Quantile Breakdown Point

In the computation of (6) we only need point estimates for θ_0 and $-[\nabla_\theta \psi_n(X_{(n)}, \theta_0)]^{-1}$, based on the whole sample $X_{(n)}$. These estimates are given by $\hat{\theta}_n$ and $-[\nabla_\theta \psi_n(X_{(n)}, \hat{\theta}_n)]^{-1}$, respectively. Thus, a computationally very fast procedure is obtained. This feature is not shared by standard resampling schemes, which can easily become unfeasible when applied to robust statistics.

A close look at $-[\nabla_\theta \psi_n(X_{(n)}, \hat{\theta}_n)]^{-1}\ \psi_k(X^*_{(k),s}, \hat{\theta}_n)$ reveals that this quantity can degenerate to infinity when (i) the matrix $\nabla_\theta \psi_n(X_{(n)}, \hat{\theta}_n)$ is singular or (ii) the estimating function is not bounded. Since we are making use of a robust (bounded) estimating function g situation (ii) cannot arise. From these arguments, we obtain the following corollary.

Corollary 4 *Let b be the breakdown point of the robust M-estimator $\hat{\theta}_n$ defined by (4). The t-quantile breakdown point of resampling distribution (6) is given by $b_t = \min(b, b_{\nabla\psi})$, where*

$$b_{\nabla\psi} = \frac{1}{n} \cdot \inf_{1\leq p\leq \lceil n/2\rceil} \big\{p \big| \textit{there exists } Z^\zeta_p \in \mathcal{Z}^\zeta_p \textit{ such that } det(\nabla_\theta \psi_n(X_{(n)} + Z^\zeta_p, \hat{\theta}_n)) = 0\big\}. \tag{7}$$

The quantile breakdown point of our robust fast resampling distribution is the minimum of the breakdown point of M-estimator $\hat{\theta}_n$ and matrix $\nabla_\theta \psi_n(X_{(n)}, \hat{\theta}_n)$. In particular, if $b_{\nabla\psi} \geq b$, the quantile breakdown point of our robust resampling distribution (6) is maximal, independent of confidence level t.

4 Breakdown Point and Data Driven Choice of the Block Size

A main issue in the application of block resampling procedures is the choice of the block size m, since accuracy of the resampling distribution depends strongly on this parameter. In this section, we study the robustness of data driven block size selection approaches for subsampling and bootstrap procedures. We first consider the MCIV and CM proposed in Romano and Wolf (2001) for the subsampling. In a second step, we analyze the HHJ method for the bootstrap. For these methods, we compute the smallest fraction of outliers in the original sample such that the data driven choice of the block size fails and diverges to infinity. For brevity, we denote by $m_u(X_{(n)})$, u=MCIV, CM, HHJ, the block size choice implied by each of these methods. By definition, the breakdown point of m_u is defined by

$$b_t^u := \frac{1}{n} \cdot \inf_{1 \leq p \leq \lceil n/2 \rceil} \left\{ p \middle| \text{there exists } Z_p^\zeta \in \mathcal{Z}_p^\zeta \text{ such that } \ m_u(X_{(n)} + Z_p^\zeta) = \infty \right\}. \tag{8}$$

4.1 Subsampling

Denote by $b_t^{OS,J}$, $J = MCIV, CM$, the breakdown point of the overlapping subsampling based on the MCIV and CM methods, respectively.

4.1.1 Minimum Confidence Index Volatility

A consistent method for a data driven choice of the block size m is based on the minimization of the confidence interval volatility index across the admissible values of m. For brevity, we present the method for one–sided confidence intervals. Modifications for the case with two–sided intervals are obvious.

Definition 5 *Let $m_{\min} < m_{\max}$ and $k \in \mathbb{N}$ be fixed. For $m \in \{m_{\min} - k, .., m_{\max} + k\}$, denote by $Q_t^*(m)$ the $t-$subsampling quantile for the block size m. Further, let $\overline{Q}_t^{*k}(m)$ be the average quantile $\overline{Q}_t^{*k}(m) := \frac{1}{2k+1} \sum_{i=-k}^{i=k} Q_t^*(m+i)$. The confidence interval volatility (CIV) index is defined for $m \in \{m_{\min}, ..., m_{\max}\}$ by*

$$CIV(m) := \frac{1}{2k+1} \sum_{i=-k}^{i=k} \left(Q_t^*(m+i) - \overline{Q}_t^{*k}(m) \right)^2. \tag{9}$$

Let $\mathcal{M} := \{m_{\min}, \ldots, m_{\max}\}$. The data driven block size that minimizes the confidence interval *volatility index is*

$$m_{MCIV} = \arg\inf_{m\in\mathcal{M}}\{CIV(m) : CIV(m) \in \mathbb{R}^{+}\} \ , \tag{10}$$

where, by definition, $\arg\inf(\emptyset) := \infty$.

The block size m_{MCIV} minimizes the empirical variance of the upper bound in a subsampling confidence interval with nominal confidence level t. Using Theorem 2, the formula for the breakdown point of m_{MCIV} is given in the next corollary.

Corollary 6 *Let b be the breakdown point of estimator $\hat{\theta}_n$. For given $t \in (0,1)$, let $b_t^{OS}(m)$ be the overlapping subsampling upper $t-$quantile breakdown point in Theorem 2, as a function of the block size $m \in \mathcal{M}$. It then follows:*

$$b_t^{OS,MCIV} = \sup_{m\in\mathcal{M}} \inf_{j\in\{-k,..,k\}} b_t^{OS}(m+j). \tag{11}$$

The dependence of the breakdown point formula for the MCIV on the breakdown point of subsampling quantiles is identical to the iid case. However, the much smaller quantile breakdown points in the time series case make the data driven choice m_{MCIV} very unreliable in presence of outliers. For instance, for the block size $n = 120$ and a maximal breakdown point statistic such that $b = 0.5$, the breakdown point of MCIV for $t = 0.95$ is less than 0.05, i.e., just 6 outliers are sufficient to break down the MCIV data driven choice of m. For the same sample size, the breakdown point of the MCIV method is larger than 0.3 in the iid case.

4.1.2 Calibration Method

Another consistent method for a data driven choice of the block size m can be based on a calibration procedure in the spirit of Loh (1987). Again, we present this method for the case of one–sided confidence intervals only. The modifications for two-sided intervals are straightforward.

Definition 7 *Fix $t \in (0,1)$ and let $(X_1^*, \ldots, X_n^*)$ be a nonoverlapping moving block bootstrap sample generated from $X_{(n)}$ with block size m. For each bootstrap sample, denote by $Q_t^{**}(m)$ the $t-$subsampling quantile according to block size m. The data driven block size according to the calibration method is defined by*

$$m_{CM} := \arg\inf_{m\in\mathcal{M}}\{|t - P^*\left[\hat{\theta}_n \leq Q_t^{**}(m)\right]| : P^*\left[Q_t^{**}(m) \in \mathbb{R}\right] > 1-t\}, \tag{12}$$

where, by definition, $\arg\inf(\emptyset) := \infty$*, and* P^* *is the nonoverlapping moving block bootstrap probability distribution.*

In the approximation of the unknown underlying data generating mechanism in Definition 7, we use a nonoverlapping moving block bootstrap for ease of exposition. It is possible to consider also other resampling methods; see, e.g., Romano and Wolf (2001). By definition, m_{CM} is the block size for which the bootstrap probability of the event $[\hat{\theta}_n \leq Q_t^{**}(m)]$ is as near as possible to the nominal level t of the confidence interval, but which at the same time ensures that the resampling quantile breakdown probability of the calibration method is less than t. The last condition is necessary to ensure that the calibrated block size m_{CM} does not imply a degenerate subsampling quantile $Q_t^{**}(m_{CM})$ with a too large probability.

Corollary 8 *Let* b *be the breakdown point of estimator* $\hat{\theta}_n$*,* $t \in (0,1)$*, and define:*

$$b_t^{OS**}(m) = \frac{1}{n} \cdot \left[\inf_{q \in \mathbb{N}, q \leq r} \left\{ p = \lceil mb \rceil \cdot q \middle| P\left(BIN\left(r, \frac{q}{r}\right) < Q^{OS} \right) < 1 - t \right\} \right],$$

where $Q^{OS} = \frac{\lceil (n-m+1)(1-t) \rceil + \lceil mb \rceil - 1}{m}$*. It then follows:*

$$b_t^{OS,CM} \leq \sup_{m \in \mathcal{M}} \{ b_t^{OS**}(m) \}. \tag{13}$$

Because of the use of the moving block bootstrap instead of the standard iid bootstrap in the CM for time series, equation (13) is quite different from the formula for the iid case in Camponovo, Scaillet and Trojani (2009). Similar to the iid case, the theoretical results in Table 2 and the Monte Carlo results in the last section of this paper indicate a higher stability and robustness of the CM relative to the MCIV method. Therefore, from a robustness perspective, the former should be preferred when consistent bootstrap methods are available. As discussed in Romano and Wolf (2001), the application of the calibration method in some settings can be computationally expensive. In contrast to our fast robust resampling approach, a direct application of the subsampling to robust estimators can easily become computationally prohibitive in combination with the CM.

4.2 Moving Block Bootstrap

The data driven method for the block size selection in Hall, Horowitz and Jing (1995) first computes the optimal block size for a subsample of size $m < n$. In a second step it uses Richardson extrapolation in order to determine the optimal block size for the whole sample.

Definition 9 *Let $m < n$ be fixed and split the original sample in $n - m + 1$ overlapping blocks of size m. Fix $l_{\min} < l_{\max} < m$ and for $l \in \{l_{\min}, .., l_{\max}\}$ denote by $Q_t^*(m, l, i)$ the $t-$moving block bootstrap quantile computed with the block size l using the bootstrap m-block $(X_i, \ldots, X_{i+m-1})$, $1 \leq i \leq n - m + 1$. $\overline{Q}_t^*(m, l) := \frac{1}{n-m+1} \sum_{i=1}^{i=n-m+1} Q_t^*(m, l, i)$ is the corresponding average quantile. Finally, denote by $Q_t^*(n, l')$ the $t-$moving block bootstrap quantile computed with block size $l' < n$ based on the original sample $X_{(n)}$. For $l \in \{l_{\min}, .., l_{\max}\}$ define the MSE index is defined as*

$$MSE(l) := \left(\overline{Q}_t^*(m, l) - Q_t^*(n, l')\right)^2 + \frac{1}{n - m + 1} \sum_{i=1}^{n-m+1} \left(Q_t^*(m, l, i) - Q_t^*(n, l')\right)^2 , \qquad (14)$$

and set:

$$l_{HHJ} = \arg\inf_{l \in \{l_{\min}, .., l_{\max}\}} \{MSE(l) : MSE(l) \in \mathbb{R}^+\} \ , \qquad (15)$$

where, by definition, $\arg\inf(\emptyset) := \infty$. *The optimal block size for the whole n-sample is defined by*

$$m_{HHJ} = l_{HHJ} \left(\frac{n}{m}\right)^{1/5} . \qquad (16)$$

As discussed in Bühlmann and Künsch (1999), the HHJ method is not fully data driven, because it is based on some starting parameter values m and l'. However, the algorithm can be iterated. After computing the first value m_{HHJ}, we can set $l' = m_{HHJ}$ and iterate the same procedure. As pointed out in Hall, Horowitz and Jing (1995) this procedure often converges in one step. Also for this data-driven method, the application of the classical bootstrap approach to robust estimators easily becomes computationally unfeasible.

Corollary 10 *Let b be the breakdown point of estimator $\hat{\theta}_n$. For given $t \in (0, 1)$, let $b_t^{NB,m}(l)$ and $b_t^{OB,m}(l)$ be the nonoverlapping and overlapping moving block upper $t-$quantile breakdown point in Theorem 2, as a function of the block size $l \in \{l_{\min}, .., l_{\max}\}$ and a size m of the initial sample. It then follows for $K = NS, OS$:*

$$b_t^{K,MCIV} = \frac{m}{n} \cdot \sup_{l \in \{l_{\min}, .., l_{\max}\}} b_t^{K,m}(l). \qquad (17)$$

The computation of the optimal block size l_{HHJ} based on smaller subsamples of size $l << m < n$, causes a large instability in the computation of m_{HHJ}. Because of this effect, the MSE index in (14) can easily deteriorate even with a small contamination. Indeed, it is enough that the computation of the quantile degenerates just in a single m-block in order to imply a degenerated MSE. Table 2 confirms this intuition. For $n = 120$, $b = 0.5$ and $t = 0.95$, the upper bound on the breakdown point of the HHJ method is half that of CM, even if for small block sizes the quantile breakdown point of subsampling procedures is typically lower than that of bootstrap methods.

5 Monte Carlo Simulations and Empirical Application

We compare through Monte Carlo simulations the accuracy of classical resampling procedures and our fast robust approach in estimating the confidence interval of the autoregressive parameter in a linear AR(1). Moreover, as a final exercise, we consider an application to real data testing the predictability of future stock returns with the classic and our robust fast subsampling.

5.1 AR(1) Model

Consider the linear AR(1) model of the form:

$$X_t = \theta X_{t-1} + \epsilon_t, \qquad X_0 \sim N\Big(0, \frac{1}{1-\theta^2}\Big), \tag{18}$$

where $|\theta| < 1$ and $\{\epsilon_t\}$ is a sequence of iid standard normal innovations. We denote by $\hat{\theta}_n^{OLS}$ the (nonrobust) OLS estimator of θ_0, which is the solution of equation:

$$\psi_n^{OLS}(X_{(n)}, \hat{\theta}_n^{OLS}) := \frac{1}{n-1}\sum_{t=2}^{n} X_{t-1}(X_t - \hat{\theta}_n^{OLS} X_{t-1}) = 0. \tag{19}$$

To apply our robust fast resampling approach, we consider a robust estimator $\hat{\theta}_n^{ROB}$ defined by

$$\psi_n^{ROB}(X_{(n)}, \hat{\theta}_n^{ROB}) := \frac{1}{n-1}\sum_{t=2}^{n} h_c(X_{t-1}(X_t - \hat{\theta}_n^{ROB} X_{t-1})) = 0, \tag{20}$$

where $h_c(x) := x \cdot min(1, c/|x|)$, $c > 1$, is the Huber function; see Künsch (1984).

To study the robustness of the different resampling methods under investigation, we consider for brevity replacement outliers random samples $(\tilde{X}_1, \ldots, \tilde{X}_n)$ generated according to

$$\tilde{X}_t = (1 - p_t)X_t + p_t \cdot X_{1.5max}, \tag{21}$$

where $X_{1.5max} = 1.5 \cdot \max(X_1, \ldots, X_n)$ and p_t is an iid $0-1$ random sequence, independent of process (18) and such that $P[p_t = 1] = \eta$. The probability of contamination is set to $\eta = 1.5\%$, which is a very small contamination of the original sample.

5.1.1 The Standard Strictly Stationary Case

We construct symmetric resampling confidence intervals for the true parameter θ_0. Hall (1988) and more recent contributions, as for instance Politis, Romano and Wolf (1999), highlight a better accuracy of symmetric confidence intervals, which even in asymmetric settings can be shorter than asymmetric confidence intervals. Andrews and Guggenberger (2009, 2010a) and Mikusheva (2007) also show that because of a lack of uniformity in pointwise asymptotics, nonsymmetric subsampling confidence intervals for autoregressive models can imply a distorted asymptotic size, which is instead correct for symmetric confidence intervals.

Using OLS estimator (19), we compute both overlapping subsampling and moving block bootstrap distributions for the distribution of $\sqrt{n}|\hat{\theta}_n^{OLS} - \theta_0|$. Using robust estimator (20), we compute overlapping robust fast subsampling and moving block bootstrap distributions for the distribution of $\sqrt{n}|\hat{\theta}_n^{ROB} - \theta_0|$. Standard resampling methods combined with data driven block size selection methods for robust estimator (20) are computationally too expensive.

We generate N=1000 samples of size $n = 180$ according to model (18) for the parameter choices $\theta_0 = 0.5, 0.6, 0.7, 0.8$. We select the subsampling block size using MCIV and CM for $\mathcal{M} = \{9, 10, 12, 15, 18\}$. For the bootstrap, we apply HHJ method with $l' = 12$, $m = 30$, $l_{min} = 6$, and $l_{max} = 10$. The degree of robustness is $c = 5$, the significance level is $1 - \alpha = 0.95$.

We first analyze the finite sample coverage and the power of resampling procedures in a test of the null hypothesis $\mathcal{H}_0 : \theta_0 = 0.5$. Figure 1 plots the empirical frequencies of rejection of the null hypothesis $\mathcal{H}_0 : \theta_0 = 0.5$ for different values of the alternative hypothesis: $\theta_0 = 0.5, 0.6, 0.7, 0.8$.

Without contamination (left column, $\eta = 0\%$), we find that our robust fast approach and the classical procedures provide accurate and comparable results. In particular, when $\theta_0 = 0.5$, the size values for the classical moving block bootstrap and subsampling with CM are 0.045 and 0.056, respectively. With our robust approach, for the robust fast bootstrap and robust fast subsampling with CM we obtain 0.055 and 0.061, which both imply size values very close to the nominal level $\alpha = 0.05$. For the robust fast subsampling and the classical subsampling with MCIV the size is larger than 0.067, which suggests a lower accuracy of the MCIV relative to the CM. When $\theta_0 \neq 0.5$, the proportion of rejections of our robust fast approach remains larger than that of the classical methods. For instance, when $\theta_0 = 0.7$, this difference in power between robust fast subsampling and subsampling with CM is close to 10%. It is even larger than 10% in a comparison between the moving block robust fast bootstrap and the classical bootstrap.

If we consider the contaminated Monte Carlo simulations (right column, $\eta = 1.5\%$), the size dramatically increases for $\theta_0 = 0.5$ for nonrobust methods, which are found to be dramatically oversized. In the case of nonrobust subsampling methods the size is even larger than 0.3. In contrast, the size of our robust fast approach remains closer to the nominal level $\alpha = 0.05$. In particular, the size is 0.082 for the robust fast subsampling with CM. A contamination tremendously deteriorates also the power of nonrobust methods. As θ_0 increases, we find that the power curve of nonrobust methods is not monotonically increasing, with low frequencies of rejection even when θ_0 is far from 0.5. For instance, for $\theta_0 = 0.8$, the power of nonrobust methods is close to 50%, but that of our robust approach is larger than 90%.

In a second exercise, we examine the sensitivity of the different resampling procedures with respect to a single point contamination of the original sample. For each Monte Carlo sample, let:

$$X_{\max} = \arg \max_{X_1,\ldots,X_n} \{u(X_i) | u(X_i) = X_i - \theta X_i, \text{under}\, \mathcal{H}_0\}. \tag{22}$$

We modify $X_{\max}$ over a grid within the interval $[X_{\max}+1, X_{\max}+4]$. Then, we analyze the sensitivity of the resulting empirical averages of p-values for testing the null hypothesis $\mathcal{H}_0 : \theta_0 = 0.5$. In Figure 2, we plot the resulting empirical p-values. As expected, our robust fast approach shows a desirable stability for both subsampling and bootstrap methods.

5.1.2 The Near-to-Unit-Root Case

As a second application, we consider the near-to-unit-root case. Moving block bootstrap procedures are then inconsistent, but the studentized subsampling based on symmetric confidence intervals is consistent; see Andrews and Guggenberger (2009, 2010b) and Mikusheva (2007). Therefore, we focus exclusively on the latter method.

Consider the OLS estimator (19) and denote by $\hat{\sigma}_n^{OLS}$ the estimated standard deviation of $\hat{\theta}_n^{OLS}$. The studentized subsampling approximates the distribution of $|\hat{\theta}_n^{OLS} - \theta_0|/\hat{\sigma}_n^{OLS}$ by the empirical distribution of $|\hat{\theta}_m^{*OLS} - \hat{\theta}_n^{OLS}|/\hat{\sigma}_m^{*OLS}$, where $\hat{\sigma}_m^{*OLS}$ denotes the estimated standard deviation of $\hat{\theta}_m^{*OLS}$ based on the subsampling block. Let $\hat{\sigma}_n^{ROB}$ be the estimated standard deviation of $\hat{\theta}_n^{ROB}$. Using our robust approach, the robust fast subsampling approximates the distribution of $|\hat{\theta}_n^{ROB} - \theta_0|/\hat{\sigma}_n^{ROB}$ by the empirical distribution of $|(-[\nabla_\theta \psi_m^{ROB}(X^*_{(m)}, \hat{\theta}_n^{ROB})]^{-1} \psi_m^{ROB}(X^*_{(m)}, \hat{\theta}_n^{ROB}))|/\hat{\sigma}_n^{*ROB}$, where $\hat{\sigma}_n^{*ROB}$ denotes the estimated standard deviation of $\hat{\theta}_m^{*ROB}$; see also Hong and Scaillet (2006).

We generate N=1000 samples of size $n = 180$ according to model (18) for the parameter choices $\theta_0 = 0.8, 0.85, 0.9, 0.95$, and simulate contaminated samples $(\tilde{X}_1, \ldots, \tilde{X}_n)$ according to (26) as before. Since bootstrap methods are inconsistent, for the selection of the block size only MCIV is recommended. However, in this setting Romano and Wolf (2001) obtain accurate results even with the subsampling based on CM. Consequently, for comparison purposes in our experiments we consider both MCIV and CM with $\mathcal{M} = \{9, 10, 12, 15, 18\}$. The degree of robustness is $c = 9$.

We analyze the finite sample size and the power of resampling procedures in a test of the null hypothesis $\mathcal{H}_0 : \theta_0 = 0.8$. The significance level is $1 - \alpha = 0.95$. Figure 3 plots the empirical frequencies of rejection of the null hypothesis $\mathcal{H}_0 : \theta_0 = 0.8$ for different values $\theta_0 = 0.8, 0.85, 0.9, 0.95$ of the alternative. As in the previous Monte Carlo setting, we find that without contamination (left column, $\eta = 0\%$) our robust fast approach and the classical procedures yield accurate and comparable results. When $\theta_0 = 0.8$, the difference between the nominal level $\alpha = 0.05$ and the size of all methods under investigation is less than 1.3%. For large θ_0, we find that the power of the robust fast subsampling is higher. The difference in power between robust and nonrobust methods is near to 10% for $\theta_0 = 0.9$ and for both data driven choice of the block size. When we consider the contaminated Monte Carlo simulations (right column, $\eta = 1.5\%$) the size of the robust fast subsampling with MCIV and CM (0.086 and 0.057, respectively) is slightly closer to the nominal level than that of the subsampling (0.103 with MCIV and 0.096 with CM) for $\theta_0 = 0.8$. More strikingly, we also find that a contamination by outliers tremendously deteriorates the power of the subsampling approach. As θ_0 increases towards the boundary value 1, the power curve of the subsampling is nonmonotonic, with frequencies of rejection less than 20% even when $\theta_0 = 0.95$ for both MCIV and CM. In contrast, the power of the robust fast subsampling is substantial and larger than 80% for $\theta_0 = 0.95$.

5.2 Stock Returns Predictability

Consider the predictive regressions model:

$$y_t = \alpha + \beta x_{t-1} + \epsilon_t, \tag{23}$$

where, for $t = 1, \ldots, n$, $\{y_t\}$ denotes the stock return, $\{x_t\}$ denotes the explanatory variable and $\{\epsilon_t\}$ is the error term. We use the subscript 0 to indicate the true value β_0 of the parameter β.

Recently, several testing procedures have been proposed in order to test the non predictability hypothesis $\mathcal{H}_0 : \beta_0 = 0$; see among others Campbell and Yogo (2006), Jansson and Moreira (2006) and Amihud, Hurvich and Wang (2008). Indeed, because of the endogeneity of the explanatory variables in this setting, classic asymptotic theory based on OLS estimator becomes inaccurate. Moreover, as emphasized in Torous, Valkanov and Yan (2004), various state variables considered as predictors follows a nearly integrated process, which complicates inference on parameter β. As advocated e.g. in Wolf (2000), the subsampling approach can be applied for testing the hypothesis of non predictability.

In this study, we analyze the predictive power of dividend yields for stock returns with the classic studentized subsampling and our robust approach. We define the one-period real total return as

$$R_t = (P_t + d_t)/P_{t-1}, \tag{24}$$

where P_t is the end of month real stock price and d_t is the real dividends paid during month t. Furthermore, we define the annualized dividend series D_t as

$$D_t = d_t + (1 + r_t)d_{t-1} + (1 + r_t)(1 + r_{t-1}d_{t-2} + \cdots + (1 + r_t)(1 + r_{t-1} \ldots (1 + r_{t-10})d_{t-11}, \tag{25}$$

where r_t is the one-month treasury-bill rate. Finally, we set $y_t = \ln(R_t)$ and $x_t = D_t/P_t$.

We implement the studentized classic subsampling and our robust approach based on MCIV and CM on US equity data. We consider monthly S&P 500 index data (1871-2008) from Shiller (2000). We consider the period 1964-2008, consisting of 540 observations, and the subperiods 1964-1993 and 1994-2008, consisting of 360 and 180 observations, respectively. In order to test $\mathcal{H}_0 : \beta_0 = 0$, for each method and period, we construct 95% confidence intervals for parameter β. Table 3 reports our empirical results.

In the whole period 1964-2008 and in the subperiod 1964-1993, classic and robust subsampling with MCIV and CM provide very similar confidence intervals. In both these periods, the procedures under investigations provide also evidence in favor of predictability. However, it is interesting to note that our robust approach implies smaller confidence intervals than the classic method, leading to a stronger rejection of the null of no predictability. Finally, in the subperiod 1994-2008, only our robust approach produces significant evidence of predictability. The classic subsampling with MCIV and CM implies extremely large confidence intervals which lead to a nonrejection of $\mathcal{H}_0$. This finding seems to confirm the robustness problem of the classic approach in our Monte Carlo simulations. Indeed, the subperiod 1994-2008 and in particular the year 2008 is characterized by a large proportion of unusual observations. For instance, we find that observation October 2008 is the most influential data point for the whole period 1964-2008. As shown in the previous section, the presence of anomalous observations may dramatically deteriorate the performance of nonrobust resampling methods. Consequently, the nonrejection of $\mathcal{H}_0$ caused by the large confidence intervals provided by the classic subsampling suggests a low power of this approach in these cases.

6 Conclusions

Theoretical breakdown point formulas and Monte Carlo evidence highlight a dramatic unexpected lack of robustness of classical block resampling methods for time series. This problem affects block bootstrap and subsampling procedures as well, and it is much worse than a related problem analyzed recently by the literature in the iid context. To overcome the problem, we propose a general robust fast resampling approach, which is applicable to a wide class of block resampling methods, and show that is implies good theoretical quantile breakdown point properties. In the context of a simple linear AR(1) model, our Monte Carlo simulations show that the robust resampling delivers more accurate and efficient results, in some cases to a dramatic degree, than other standard block resampling schemes in presence and absence of outliers in the original data. A real data application to testing for stock returns predictability provides more consistent evidence in favor of the predictability hypothesis using our robust resampling approach.

Appendix: Proofs

Proof of Theorem 2. Denote by $X^N_{(m),i} = (X_{(i-1)m+1}, \ldots, X_{im})$, $i = 1, \ldots, r$ and $X^O_{(m),i} = (X_i, \ldots, X_{i+m-1})$, $i = 1, \ldots, n-m+1$ the nonoverlapping and overlapping blocks of size m respectively. Given the original sample $X_{(n)}$, for the first nonoverlapping block $X^N_{(m),1}$, consider following type of contamination:

$$X^N_{(m),1} = (X_1, \ldots, X_{m-\lceil mb \rceil}, Z_{m-\lceil mb \rceil+1}, \ldots, Z_m), \tag{26}$$

where X_i, $i = 1, \ldots, m - \lceil mb \rceil$ and Z_j, $j = m - \lceil mb \rceil + 1, \ldots, m$, denote the non contaminated and contaminated points respectively. By construction, the first $m - \lceil mb \rceil + 1$ overlapping blocks $X^O_{(m),i}$, $i = 1, \ldots, m - \lceil mb \rceil + 1$, contain $\lceil mb \rceil$ outliers. Consequently, $T(X^O_{(m),i}) = +\infty$, $i = 1, \ldots, m - \lceil mb \rceil + 1$. Assume that the first $p < r-1$ nonoverlapping blocks $X^N_{(m),i}$, $i = 1, \ldots, p$ have the same contamination as in (26). Because of this contamination, the number of statistics $T^*_{n,m}$ which diverge to infinity is $mp - \lceil mb \rceil + 1$.

$Q^*_t = +\infty$ when the proportion of statistics $T^*_{n,m}$ with $T^*_{n,m} = +\infty$ is larger than $(1-t)$. Therefore, $b^{OS}_t \leq \inf_{\{p \in \mathbb{N}, p \leq r-1\}} \left\{ p \cdot \frac{\lceil mb \rceil}{n} \,\middle|\, \frac{mp - \lceil mb \rceil + 1}{n-m+1} > 1-t \right\}$ ∎

Proof of Theorem 3. *Case (i): Nonoverlapping Moving Block Bootstrap.* Consider $X^N_{(m),i}$, $i = 1, \ldots, r$. Assume that p_2 of these nonoverlapping blocks are contaminated with exactly p_1 outliers for each block, while the remaining $(r - p_2)$ are non contaminated (0 outliers), where $p_1, p_2 \in \mathbb{N}$ and $p_1 \leq m$, $p_2 \leq r-1$. The nonoverlapping moving block bootstrap constructs a n-sample randomly selecting with replacement r nonoverlapping blocks. Let X be the random variable which denotes the number of contaminated blocks in the random bootstrap sample. It follows that $X \sim BIN(r, \frac{p_2}{r})$.

By Definition 1, $Q^*_t = +\infty$ when the proportion of statistics $T^*_{n,n}$ with $T^*_{n,n} = +\infty$ is larger than $(1-t)$. The smallest number of outliers such that $T^*_{n,n} = +\infty$ is by definition nb. Consequently, $b^{NB}_t \leq \frac{1}{n} \cdot \left[\inf_{\{p_1, p_2 \in \mathbb{N}, p_1 \leq m, p_2 \leq r-1\}} \left\{ p = p_1 \cdot p_2 \middle| P\left(BIN\left(r, \frac{p_2}{r}\right) > \frac{nb}{p_1} \right) > 1-t \right\} \right]$.

Case (ii): Overlapping Moving Block Bootstrap. Given the original sample $X_{(n)}$, consider the same nonoverlapping blocks as in (i), where the contamination of the p_2 contaminated blocks has the structure defined in (26). The overlapping moving block bootstrap constructs a n-sample randomly selecting with replacement r overlapping blocks of size m. Let X be the random variable which denotes the number of contaminated blocks in the random bootstrap sample. It follows that $X \sim$

$BIN(r, \frac{mp_2-p_1+1}{n-m+1})$.

By Definition 1, $Q_t^* = +\infty$ when the proportion of statistics $T_{n,n}^*$ with $T_{n,n}^* = +\infty$ is larger than $(1-t)$. The smallest number of outliers such that $T_{n,n}^* = +\infty$ is by definition nb. Consequently, $b_t^{OB} \leq \frac{1}{n} \cdot \left[\inf_{\{p_1,p_2 \in \mathbb{N}, p_1 \leq m, p_2 \leq r-1\}} \left\{ p = p_1 \cdot p_2 \middle| P\left(BIN\left(r, \frac{mp_2-p_1+1}{n-m+1} \right) > \frac{nb}{p_1} \right) > 1-t \right\} \right]$ ■

Proof of Corollary 4. Consider the robust fast approximation of $(\hat{\theta}_k^* - \hat{\theta}_n)$ given by

$$-[\nabla_\theta \psi_n(X_{(n)}, \hat{\theta}_n)]^{-1} \psi_k(X_{(k),s}^*, \hat{\theta}_n), \tag{27}$$

where $k = n$ or $k = m$. Assuming a bounded estimating function, expression (27) may degenerate only when, (i) $\hat{\theta}_n \notin \mathbb{R}$ or (ii) the matrix $[\nabla_\theta \psi_n(X_{(n)}, \hat{\theta}_n)]$ is singular, i.e. $det([\nabla_\theta \psi_n(X_{(n)}, \hat{\theta}_n)]) = 0$. If (i) and (ii) are not satisfied, it turns out that the quantile Q_t^* is bounded, $\forall t \in (0,1)$. Let b be the breakdown point of $\hat{\theta}_n$ and $b_{\nabla\psi}$ be the smallest fraction of outliers in the original sample such that condition (ii) is satisfied, the breakdown point of Q_t^* is given by $b_t = \min(b, b_{\nabla\psi})$. ■

Proof of Corollary 6. Denote $b_t^{OS}(m)$, the overlapping subsampling quantile breakdown point based on blocks of size m. By definition, in order to get $m_{MCIV} = \infty$ we must have $CIV(m) = \infty$ for all $m \in \mathcal{M}$. Given $m \in \mathcal{M}$, $CIV(m) = \infty$ if and only if the fraction of outliers p in the sample $\{X_1, \ldots, X_n\}$ satisfies $p \geq \min\{b_t^{OS}(m-k), b_t^{OS}(m-k+1), .., b_t^{OS}(m+k-1), b_t^{OS}(m+k)\}$. This concludes the proof. ■

Proof of Corollary 8. By definition, in order to get $m_{CM} = \infty$ we must have $P\left[Q_t^{**}(m) = \infty\right] \geq t$ for all $m \in \mathcal{M}$. Given the original sample, Assume that q nonoverlapping blocks are contaminated with exactly $\lceil mb \rceil$ outliers for each block, while the remaining $(r-q)$ are non contaminated (0 outliers), where $q \in \mathbb{N}$ and $q \leq r$. Moreover, assume that the contamination of the contaminate blocks has the structure defined in (26). Let X be the random variable which denotes the number of contaminated blocks in the nonoverlapping moving block bootstrap sample. As in (i), $X \sim BIN(r, q/r)$. For the construction of the nonoverlapping moving block bootstrap sample, the selection of $p \leq r-1$ contaminated blocks implies the break of $mp - \lceil mb \rceil + 1$ overlapping subsampling statistics.

$Q_t^{**}(m) = \infty$ when the proportion of contaminated blocks is larger than $1-t$, i.e. $\frac{mp-\lceil mb \rceil+1}{n-m+1} > 1-t \Leftrightarrow p > \frac{\lceil (n-m+1)(1-t) \rceil + \lceil mb \rceil - 1}{m}$. This concludes the proof of the second statement. ■

Proof of Corollary 10. By definition, in order to get $m_{HHJ} = \infty$ we must have $l_{HHJ} = \infty$, i.e. $MSE(l) = \infty$, for all $l \in \{l_{min} \ldots, l_{max}\}$. For l fixed, $MSE(l) = \infty$ if just a single $Q_t^*(m, l, i)$,

$i = 1, \ldots, n - m + 1$ diverges to infinity. This concludes the proof. ■

References

[1] Amihud, Y., Hurvich C.M., and Y. Wang, 2008. Multiple-predictor regressions: hypothesis testing. The Review of Financial Studies, 22, 413–434.

[2] Andrews, D., 2000. Inconsistency of the bootstrap when a parameter is on the boundary of the parameter space. Econometrica, 68, 399–405.

[3] Andrews, D., 2002. Higher-order improvements of a computationally attractive k-step bootstrap for extremum estimators. Econometrica, 70, 119–162.

[4] Andrews, D., 2004. The block-block bootstrap: improved asymptotic refinements. Econometrica, 72, 673–700.

[5] Andrews, D., and P. Guggenberger, 2009. Hybrid and size-corrected subsample methods. Econometrica, 77, forthcoming.

[6] Andrews, D., and P. Guggenberger, 2010a. Asymptotic size and a problem with subsampling and with the m out of n bootstrap. Econometric Theory, 26, forthcoming.

[7] Andrews, D., and P. Guggenberger, 2010b. Application of subsampling, hybrid and size-correction methods. Journal of Econometrics, forthcoming.

[8] Bickel, P. J., Gotze, F., and W. R. van Zwet, 1997. Resampling fewer than n observations: Gains, losses, and remedies for losses. Statistica Sinica, 7, 1–31.

[9] Bühlmann, P., and H. Künsch, 1999. Block length selection in the bootstrap for time series. Computational Statistics and Data Analysis, 31, 295-310.

[10] Campbell, J.Y., and M. Yogo, 2006. Efficient tests of stock return predictability. Journal of Financial Economics, 81, 27–60.

[11] Camponovo, L., Scaillet, O., and F. Trojani, 2009. Robust subsampling, Working Paper.

[12] Carlstein, E., 1986. The use of subseries methods for estimating the variance of a general statistic from a stationary time series. Annals of Statistics, 14, 1171–1179.

[13] Davidson, R., and J. McKinnon, 1999. Bootstrap testing in nonlinear models. International Economic Review, 40, 487–508.

[14] Efron, B., and R. Tibshirani, 1993. An Introduction to the Bootstrap. New York: Chapman and Hall.

[15] Genton, M., and A. Lucas, 2003. Comprehensive definitions of breakdown point for independent and dependent observations. Journal of the Royal Statistical Society, Series B, 65, 81–94.

[16] Goncalves, S., and H. White, 2004. Maximum likelihood and the bootstrap for nonlinear dynamic models. Journal of Econometrics, 119, 199–220.

[17] Hall, P., 1985. Resampling a coverage process. Stochastic Processes and their Applications, 19, 259–269.

[18] Hall, P., 1988. On symmetric bootstrap confidence intervals. Journal of the Royal Statistical Society, Ser.B, 50, 35–45.

[19] Hall, P., 1992. The Bootstrap and Edgeworth Expansion. New York: Springer-Verlag.

[20] Hall, P., and J. Horowitz, 1996. Bootstrap critical values for tests based on Generalized-Method-of-Moment estimators. Econometrica, 64, 891–916.

[21] Hall, P., Horowitz J., and B.-Y. Jing, 1995. On blocking rules for the bootstrap with dependent data. Biometrika, 82, 561–574.

[22] Hong, H., and O. Scaillet, 2006. A fast subsampling method for nonlinear dynamic models. Journal of Econometrics, 133, 557–578..

[23] Hu, F., and J. Kalbfleisch, 2000. The estimating function bootstrap. Canadian Journal of Statistics, 28, 449–499

[24] Künsch, H., 1984. Infinitesimal robustness for autoregressive processes. Annals of Statistics, 12, 843–863.

[25] Künsch, H., 1989. The jacknife and the bootstrap for general stationary observations. Annals of Statistics, 17, 1217–1241.

[26] Jansson, M., and M.J. Moreira, 2006. Optimal inference in regression models with nearly integrated regessors. Econometrica, 74, 681–714.

[27] Loh, W. Y., 1987. Calibrating confidence coefficients. Journal of the American Statistical Association, 82, 155–162.

[28] Mancini, L., Ronchetti, E., and F. Trojani, 2005. Optimal conditionally unbiased bounded-influence inference in dynamic location and scale models. Journal of the American Statistical Association, 100, 628–641.

[29] Martin, R.D., and V. Yohai, 1986. Influence functionals for time series. Annals of Statistics, 14, 781–818.

[30] Mikusheva, A., 2007. Uniform inference in autoregressive models. Econometrica, 75, 1411–1452.

[31] Muler, N., and V. Yohai, 2008. Robust estimates for GARCH models. Journal of Statistical Planning and Inference, 138, 2918–2940.

[32] Ortelli, C., and F. Trojani, 2005. Robust efficient method of moments. Journal of Econometrics, 128, 69–97.

[33] Politis, D. N., and J. P. Romano, 1992. A general theory for large sample confidence regions based on subsamples under minimal assumptions. Technical Report 399, Dept of Statistics, Stanford University.

[34] Politis, D. N., and J. P. Romano, 1994a. Large sample confidence regions based on subsamples under minimal assumptions. Annals of Statistics, 22, 203–2050.

[35] Politis, D. N., and J. P. Romano, 1994b. The stationary bootstrap. Journal of the American Statistical Association, 89, 1303–1313.

[36] Politis, D. N., Romano J. P., and M. Wolf, 1999. Subsampling. Springer, New York.

[37] Romano, J. P., and M. Wolf, 2001. Subsampling intervals in autoregressive models with linear time trend. Econometrica, 69, 1283-1314.

[38] Ronchetti, E., and F. Trojani, 2001. Robust inference with GMM estimators, Journal of econometrics, 101, 37–69.

[39] Sakata, S., and H. White, 1998. High breakdown point conditional dispersion estimation with application to S&P500 daily returns volatility. Econometrica, 66, 529–567.

[40] Salibian-Barrera, M., Van Aelst, S., and G. Willems, 2006. Principal components analysis based on multivariate MM estimators with fast and robust bootstrap. Journal of the American Statistical Association, 101, 1198–1211.

[41] Salibian-Barrera, M., Van Aelst, S., and G. Willems, 2007. Fast and robust bootstrap. Statistical Methods and Applications, 17, 41–71

[42] Salibian-Barrera, M., and R. Zamar, 2002. Boostrapping robust estimates of regression. Annals of Statistics, 30, No. 2, 556–582.

[43] Shao, J., and D. Tu, 1995. The jackknife and bootstrap. Springer, New York.

[44] Singh, K., 1998. Breakdown theory for bootstrap quantiles. Annals of Statistics, 26, No. 5, 1719–1732.

[45] Shiller, R.J., 2000. Irrational Exuberance. Princeton University Press, Princeton, New York.

[46] Torous, W., R. Valkanov, and S. Yan, 2004. On predicting stock returns with nearly integrated explanatory variables. Journal of Business, 77, 937-966.

[47] Wolf, M., 2000. Stock returns and dividend yields revisited: a new way to look at an old problem. Journal of Business and Economic Statistics, 18, 18-30.

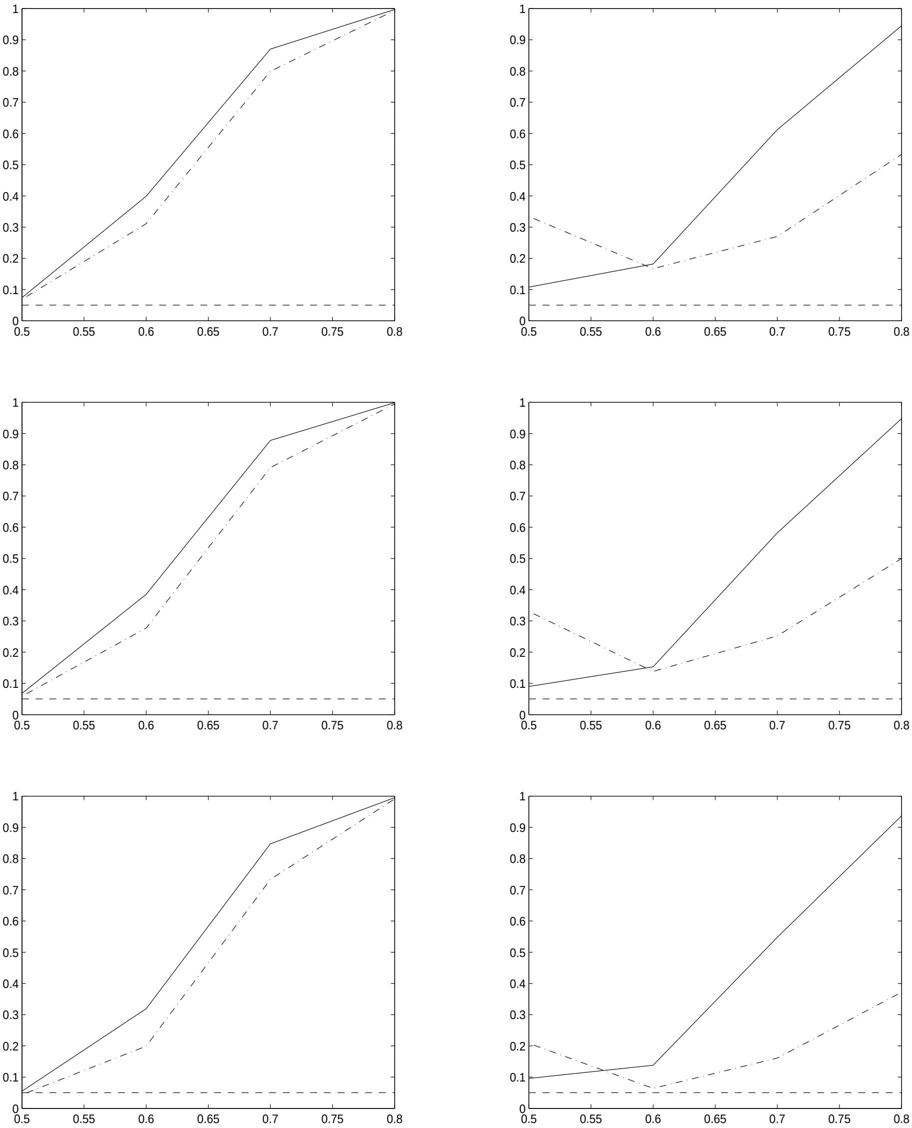

Figure 1: **Power curves in the standard strictly stationary case.** We plot the proportion of rejections of the null hypothesis $\mathcal{H}_0 : \theta_0 = 0.5$, when the true parameter value is $\theta_0 \in [0.5, 0.8]$. From the top to the bottom, we present the overlapping subsampling with MCIV, the subsampling with CM and the moving block bootstrap with HHJ. We consider our robust fast approach (straight line) and the classic approach (dash-dotted line). In the left column, we consider a non contaminated sample ($\eta = 0\%$). In the right column, the proportion of outliers is $\eta = 1.5\%$.

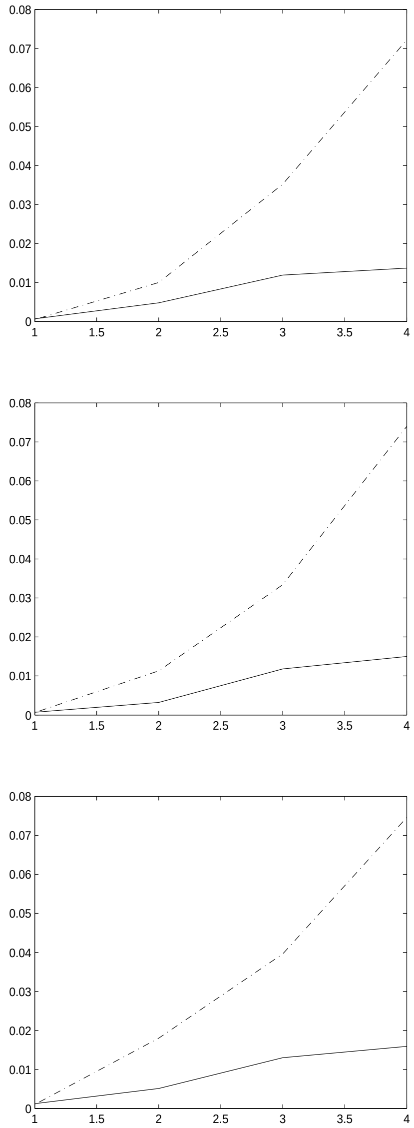

Figure 2: **Sensitivity analysis.** Sensitivity plots of the variation of the empirical $p-$value average, for a test of the null hypothesis $\mathcal{H}_0 : \theta_0 = 0.5$, with respect to variations of $X_{\max}$, in each Monte Carlo sample, within the interval $[1, 4]$. The random samples were generated under $\mathcal{H}_0$ and, from the top to the bottom, we present the overlapping subsampling with MCIV, the subsampling with CM and the moving block bootstrap with HHJ. We consider the robust fast approach (straight line) and the classic nonrobust approach (dash-dotted line).

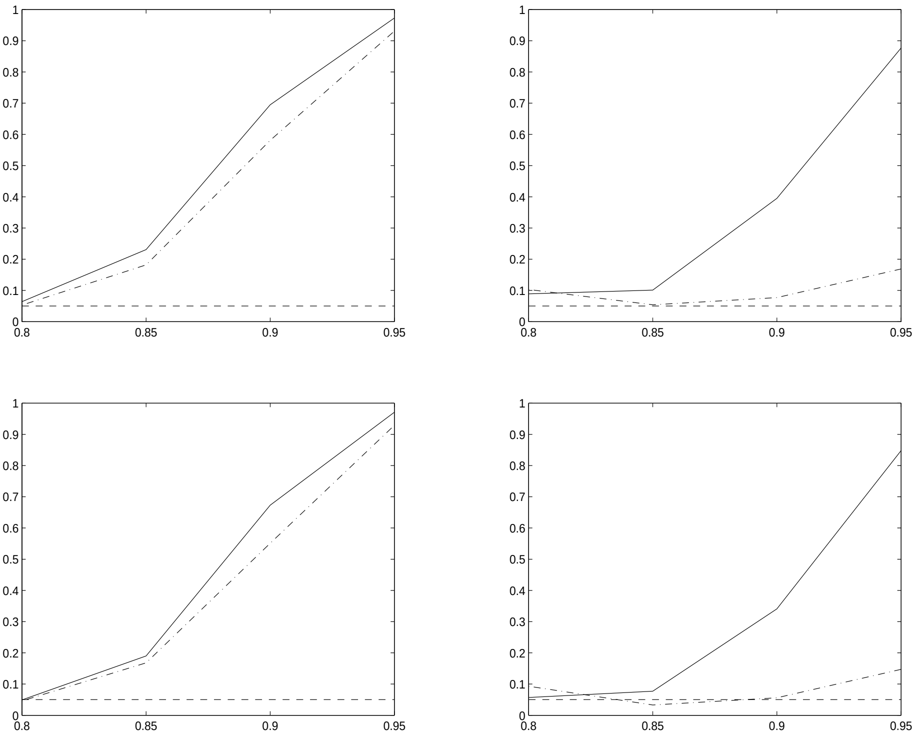

Figure 3: **Power curves in the near-to-unit-root case.** We plot the proportion of rejections of the null hypothesis $\mathcal{H}_0 : \theta_0 = 0.8$, when the true parameter value is $\theta_0 \in [0.8, 0.95]$. From the top to the bottom, we present the overlapping subsampling with MCIV and CM. We consider our robust fast approach (straight line) and the classic approach (dash-dotted line). In the left column, we consider a non contaminated sample ($\eta = 0\%$). In the right column, the proportion of outliers is $\eta = 1.5\%$.

$n = 120$, $b = 0.5$	0.95	0.99
O. Subsampling ($m = 5$)	≤ 0.0500	≤ 0.0250
O. Subsampling ($m = 10$)	≤ 0.0500	≤ 0.0500
O. Subsampling ($m = 15$)	≤ 0.0667	≤ 0.0667
N. Bootstrap ($m = 5$)	≤ 0.3750	≤ 0.3333
N. Bootstrap ($m = 10$)	≤ 0.3333	≤ 0.2667
N. Bootstrap ($m = 15$)	≤ 0.3250	≤ 0.2167
O. Bootstrap ($m = 5$)	≤ 0.3750	≤ 0.3333
O. Bootstrap ($m = 10$)	≤ 0.3333	≤ 0.2667
O. Bootstrap ($m = 15$)	≤ 0.3250	≤ 0.2167

Table 1: **Subsampling and Moving Block Bootstrap Quantile Breakdown Point.** Breakdown point of the overlapping (O.) subsampling and nonoverlapping (N.) and overlapping (O.) moving block bootstrap quantile. The sample size is $n = 120$, the block size $m = 5, 10, 15$. We assume a statistic with breakdown point $b = 0.5$ and confidence level $t = 0.95, 0.99$. Quantile breakdown points are computed using Theorem 2 and 3.

$n = 120$	$t = 0.95$	$t = 0.99$
O. Subsampling MCIV	≤ 0.0500	≤ 0.0500
O. Subsampling CM	≤ 0.2000	≤ 0.2667
N. Bootstrap HHJ	≤ 0.1000	≤ 0.0667
O. Bootstrap HHJ	≤ 0.1000	≤ 0.0667

Table 2: **Breakdown point of Block Size Selection Procedures**. We compute the breakdown point of the minimum confidence index volatility (MCIV), the calibration method (CM) and the data driven method in Hall, Horowitz and Jing (1995) (HHJ) for the nonoverlapping (N.) and overlapping (O.) cases. For (MCIV) and (CM) we use Corollary 6, 8 with $\mathcal{M} = \{6, 8, 10, 12, 15\}$. For (HHJ) we use Corollary 10 with $m = 30$, $l_{min} = 3$, and $l_{\max} = 10$. The breakdown point of the statistic is $b = 0.5$ and the confidence levels are $t = 0.95, 0.99$. The sample size is $n = 120$.

	(1964 − 2008) (540)	(1964 − 1993) (360)	(1994 − 2008) (180)
Subsampling (MCIV)	$[0.0247, 0.0520]$	$[0.0098, 0.0591]$	$[-0.0290, 0.1995]$
Subsampling (CM)	$[0.0220, 0.0547]$	$[0.0141, 0.0548]$	$[-0.0290, 0.1995]$
R. F. Subsampling (MCIV)	$[0.0277, 0.0491]$	$[0.0177, 0.0509]$	$[0.0176, 0.1583]$
R. F. Subsampling (CM)	$[0.0268, 0.0501]$	$[0.0200, 0.0486]$	$[0.0201, 0.1559]$

Table 3: **Stock Returns Predictability.** We report 95% confidence intervals for the parameter β in model (23). We consider the classic subsampling and our robust fast subsampling with MCIV and CM, for the period 1964-2008 (540 observations) and the subperiods 1964-1993 and 1993-2008 (360 and 180 observations, respectively).