

Robust Resampling Methods and Stock Returns Predictability

Submitted for the degree of Ph.D. in Economics at

Faculty of Economics

University of Lugano

Lugano, Switzerland

Presented by

Lorenzo Camponovo

Dipl.Math. ETH Zürich

Accepted on the recommendation of

Prof. F. Trojani, advisor, University of Lugano

Prof. A. Mira, University of Lugano

Prof. E. Ronchetti, University of Geneva and University of Lugano

Prof. O. Scaillet, HEC University of Geneva

2009

To my Family.

Acknowledgments

First of all I wish to thank my advisor Professor Fabio Trojani for his encouragement and illuminating support during my Ph.D. studies. Furthermore, I would like also to thank Professor Elvezio Ronchetti and Professor Olivier Scaillet for their precious advices, valuable suggestions and brilliant ideas on several research topics.

I wish to thank Professor Antonietta Mira for accepting to be a member of the thesis committee and for many useful suggestions on my thesis.

Finally, I would like to thank Professor Patrick Gagliardini, Professor Claudio Ortelli, Davide La Vecchia and Andrea Vedolin for many stimulating discussions on different financial topics and for the beautiful time spent in Lugano.

Contents

Introduction	1
1 Robust Subsampling	5
1.1 Abstract	5
1.2 Introduction	6
1.3 Subsampling Breakdown Point and Robust Subsampling	7
1.3.1 Explicit Breakdown Point Formula for Subsampling Quantiles	9
1.3.2 Breakdown Point and Data Driven Choice of the Block Size	11
1.3.3 Robust Subsampling	14
1.3.4 Robust Subsampling in the Linear Regression Model	17
1.4 Monte Carlo Study and Sensitivity Analysis	23
1.4.1 Square of the Sample Average	24
1.4.2 Linear Regression	26
1.5 Conclusions	31
A.1 Proofs	32
2 Robust Resampling Methods for Time Series	44
2.1 Abstract	44

2.2	Introduction	45
2.3	Resampling Distribution Breakdown Point Quantile	47
2.3.1	Definition	47
2.3.2	Quantile Breakdown Point	49
2.4	Robust Resampling Procedures	51
2.4.1	Definition	52
2.4.2	Robust Resampling Methods and Quantile Breakdown Point	53
2.5	Breakdown Point and Data Driven Choice of the Block Size	54
2.5.1	Subsampling	54
2.5.2	Moving Block Bootstrap	58
2.6	Monte Carlo Simulations	59
2.6.1	The Standard Strictly Stationary Case	60
2.6.2	The Near-to-Unit-Root Case	62
2.7	Conclusions	64
A.2	Proofs	65
3	Robust Predictive Regression and Stock Returns Predictability	73
3.1	Abstract	73
3.2	Introduction	74
3.3	Predictive Regression Model	77
3.3.1	Bonferroni Method	78
3.3.2	Bias Correction Methods	79
3.3.3	Subsampling Method	80
3.3.4	The Robustness Problem of Standard Tests of Predictability	81
3.4	Robust Predictive Regression	82

3.5	Predictability of Stock Returns	85
3.6	Conclusions	87
A.3	Robust Fast Subsampling and Predictive Regression Model	88
A.4	Block Size Selection	90
	Bibliography	93

Introduction

Resampling methods are powerful tools in modern statistics and econometrics; bootstrap procedures (see, e.g., Hall, 1992, Efron and Tibshirani, 1993, and Hall and Horowitz, 1996) and subsampling procedures (Politis and Romano, 1992, 1994a) have widespread applicability, and are useful for a wide variety of inference problems in many fields. Indeed, besides the opportunity of asymptotic refinements provided by bootstrap procedures (see, e.g., Hall 1992), resampling methods may represent a better alternative to the classic asymptotic theory. Some examples where resampling methods outperform classic asymptotic theory include settings where the latter poorly performs (see, e.g., Salibian-Barrera and Zamar, 2002) and settings where the latter is not applicable (see, e.g., Andrews and Guggenberger, 2010a). The bootstrap has been the object of a huge research in statistics and econometrics, since its introduction by Efron (1979). More recently, subsampling procedures have been also investigated as a valid alternative to the bootstrap for settings in which the bootstrap fails; see for instance Gonzalo and Wolf (2005), Hong and Scaillet (2006) and Andrews and Guggenberger (2009a,b, 2010a,b) for recent applications.

Singh (1998), Salibian-Barrera and Zamar (2002) and Salibian-Barrera, Van Aelst and Willems (2006, 2007) study the robustness of the bootstrap in iid settings, by characterizing its quantile breakdown point, i.e., the smallest proportion of contamination in the original sample such that the bootstrap quantile diverges to infinity. They find a small bootstrap quantile breakdown point, which

implies a large instability of bootstrap inference in presence of model contamination, and develop different robust bootstrap methods to overcome this problem.

This thesis studies the robustness properties of general block resampling methods and it develops robust block resampling procedures both for iid and time series settings. An application of these robust resampling methods to tests for stock return predictability shows their usefulness in a concrete inference setting.

In the first chapter, we develop robust subsampling procedures for the iid case. Under this last assumption, Singh (1998) and Salibian-Barrera et al. (2006, 2007) have proposed different robust bootstrap procedures. In some econometrics models, the weaker consistency conditions of the subsampling are satisfied, but those of the bootstrap are not; see for instance Andrews (2000) and Bickel et al. (1997) for some famous examples. Consequently, in order to obtain a robust resampling procedure for cases in which the bootstrap fails a robust version of the subsampling is necessary. We extend the work of Salibian-Barrera et al. (2006, 2007) to the subsampling framework and develop robust subsampling procedures for a general class of fixed-point estimators.

In the second chapter, we study robust block resampling methods for time series. In the time series setting, the dependence between observations requires a different resampling scheme than in the iid context. The general approach splits the data in (overlapping or non-overlapping) blocks. In a second step, either it applies the statistic directly to these blocks, as in subsampling procedures, or it generates new random samples, assuming an approximate independence between the blocks, and finally applies the statistic to the so generated random sample, as in the block-bootstrap; see for instance Carlstein (1986), Hall (1985), Künsch (1989) and Andrews (2004)).

After having characterized the robustness of resampling methods for time series, we develop robust resampling procedures for general M-estimators in the time series context. Our explicit breakdown formulas show that in the time series case the robustness problem of resampling methods is even larger

than in the iid setting. To develop our robust resampling approach, we extend the fast resampling procedures introduced in Salibián-Barrera and Zamar (2002), Salibián-Barrera et al. (2006, 2007) and Hong and Scaillet (2006). In our robust resampling method, we avoid the estimation of the parameter of interest in each random block by using a linear approximation based on the estimating function of the given M-estimator. In this way, the robustness properties of the resampling distribution of the estimator depends directly on the robustness of the given estimating function, which can be controlled if this function is bounded.

The third chapter of the thesis considers an application to finance in the context of testing for returns predictability. The forecasting ability of some explanatory variables for future stock returns is largely debated in the financial literature. A possible source of these contrasting discussions can derive from the inappropriate statistical tools used in some case to test for predictability. For instance, Nelson and Kim (1993) and Goetzmann and Jorion (1993) emphasize that in the predictive regression model with an endogenous predictor and correlated innovations classical asymptotic theory causes small sample biases that overreject the hypothesis of no predictability.

Recent work has proposed different statistical approaches to this problem, especially for the case where the endogenous predictor is nearly integrated. Torous, Valkanov and Yan (2004) show that many variables used in the literature to predict stock returns follow an autoregressive model with local-to-unit root. Torous et al. (2004), Campbell and Yogo (2006), Jansson and Moreira (2006) and Polk, Thompson and Vuolteenaho (2006), propose testing procedures for predictive regression settings with a persistent predictor and correlated innovations. However, extensions of these tests to more general models, e.g. in the multivariate setting, are not easily feasible: The Bonferroni-type of approach in Torous et al. (2004) and Campbell and Yogo (2006) is intrinsically univariate; The procedures in Jansson and Moreira (2006) and Polk et al. (2006) are computationally too expensive in the multivariate context. As an alternative, subsampling methods are applicable in this setting.

For instance, Wolf (2000) and Choi and Chue (2007) propose a non robust studentized subsampling approach for testing the forecasting ability of an explanatory variable.

We find that in testing for predictability the robustness of the resampling method used is a key issue. Monte Carlo simulation and sensitivity analysis confirm the larger accuracy of our robust procedure with respect to non robust subsampling methods. In the application to real data, we find that price-earnings ratios have the most significant forecast ability for predicting stock returns, e.g., in comparison to the dividend yield. We obtain smaller confidence intervals and a higher power of the test using our robust approach. In addition, we find different conclusions in some subsamples using classical and robust subsampling methods. This indicates the nonrobustness of test results for predictability in some subsamples when classical resampling methods are used.

Chapter 1

Robust Subsampling

1.1 Abstract

We characterize the robustness of subsampling procedures by deriving a general formula for the breakdown point of subsampling quantiles. This breakdown point can be very low for moderate subsampling block sizes, which implies the fragility of subsampling procedures, even if they are applied to robust statistics. This instability arises also for data driven block size selection procedures minimizing the minimum confidence interval volatility index, but can be mitigated if a more robust calibration method is applied instead. To overcome these robustness problems, we propose a consistent robust subsampling procedure for M-estimators and derive explicit subsampling quantile breakdown point characterizations for MM-estimators in the linear regression model. Monte Carlo simulations in two settings where the bootstrap fails show the accuracy and robustness of the robust subsampling relative to the classical subsampling.

1.2 Introduction

Resampling methods are widely applied statistical tools in modern econometric and statistical analysis. Among the different resampling methods, the bootstrap, since its introduction by Efron (1979), has been the object of important research; see for instance Hall (1992), Efron and Tibshirani (1993) and Hall and Horowitz (1996). Subsampling procedures (Politis and Romano, 1992, 1994a) are more recent, but have gained rapidly considerable attention. The simpler consistency conditions and the wider applicability in some cases (see, e.g., Andrews, 2000, and Bickel et al., 1997, for some famous examples) make subsampling a useful and valid alternative to the bootstrap. Some examples of recent applications of subsampling procedures include: Chernozhukov and Fernandez-Val (2005), who analyze subsampling inference of quantile regression processes; Gonzalo and Wolf (2005), who study subsampling inference in threshold autoregressive models; Linton, Maasoumi and Whang (2005), who develop a subsampling testing procedure for stochastic dominance; Hong and Scaillet (2006), who propose a fast subsampling method for nonlinear dynamic models; Lee and Pun (2006), who investigate subsampling in nonstandard M-Estimation with nuisance parameters.

As emphasized, for instance, by Bickel et al. (1997), a key issue in the application of subsampling methods is the selection of an adequate subsampling block size m among the n data points, because subsampling accuracy can highly depend on this parameter. Hall and Yao (2003) highlight this problem for GARCH settings with asymmetric heavy-tailed errors. Cowell and Flachaire (2007) and Davidson and Flachaire (2007) observe a similar problem when resampling inequality and poverty measures.

Our goal is to study the robustness of subsampling methods in relation to the choice of the subsampling block size. The need for robust statistical procedures has been stressed by many authors and is now widely recognized; see, e.g., Huber (1981), Hampel, Ronchetti, Rousseeuw and Stahel (1986), Heritier and Ronchetti (1994), Sakata and White (1998), Ronchetti and Trojani (2001), Ortelli

and Trojani (2005) and Mancini, Ronchetti and Trojani (2005). We focus on global subsampling instability and derive a formula for the breakdown point of subsampling quantiles in Section 1.3.1. This breakdown point is increasing in the subsampling block size, the sample size and the breakdown point of the statistic used. Concrete computations show that moderate block sizes typically chosen in applications can imply very unstable subsampling quantiles also when exploiting robust statistics. This instability is larger than the one observed for standard bootstrap quantiles; see, e.g., Singh (1998), and Salibian-Barrera and Zamar (2002). As shown in Section 1.3.2, it also arises for data driven block size selection procedures based on the minimum confidence interval volatility (MCIV) index, but can be mitigated by a more robust calibration approach (Romano and Wolf, 2001). To overcome these robustness problems, we introduce a robust subsampling method for M-estimators in Section 1.3.3. We further analyze in detail the properties of the robust subsampling for MM-estimates in the linear regression setting, by computing its breakdown point and by proving its consistency in Section 1.3.4. Monte Carlo simulations and sensitivity analysis are presented in Section 1.4, for two settings where the bootstrap fails. In the second example, we study a model with a parameter of interest possibly near a boundary. Andrews and Guggenberger (2009a, 2010a,b) show that subsampling methods may imply a distorted asymptotic size, when applied to statistics with a discontinuous asymptotic distribution in some model parameter, and they propose hybrid subsampling methods to overcome the problem. We borrow from their approach to compute confidence intervals for the relevant parameter using hybrid robust subsampling procedures in our second Monte Carlo example. Section 1.5 gathers concluding remarks.

1.3 Subsampling Breakdown Point and Robust Subsampling

Let (X_1, \dots, X_n) be an iid random sample from a probability distribution H and $T_n := T(X_1, \dots, X_n)$ a one-dimensional real valued statistic. Let $0 < b \leq 0.5$ be the upper breakdown point of T_n , that

is, nb is the smallest number of observations that need to go to $\pm\infty$ in order to force T_n to go to ∞ (symmetrically, for one-dimensional real valued statistics, the lower breakdown point of T_n is the smallest number of observations that need to go to $\pm\infty$ in order to force T_n to go to $-\infty$). The breakdown point b is an intrinsic characteristic of the chosen statistic. It is explicitly known in some cases, and can be gauged most of the time, for instance by means of simulation or sensitivity analysis. Many nonrobust statistics have a breakdown point $b = 1/n$. Given a subsampling block size $m < n$, a random subsample (X_1^*, \dots, X_m^*) is drawn without replacement from the original sample (X_1, \dots, X_n) . $T_{n,m}^* := T(X_1^*, \dots, X_m^*)$ denotes the subsampling statistic. Given $t \in (0, 1)$, the t -quantile of $T_{n,m}^*$ is $Q_t^* := \inf\{x | P[T_{n,m}^* \leq x] \geq t\}$, where, by definition, $\inf(\emptyset) := \infty$.

Definition 1 *The subsampling upper t -quantile breakdown point b_t of statistic T_n is defined by*

$$b_t := \inf\{p \in [1/n, 1] : np \in \mathbb{N} \text{ and } Q_t^* = \infty\}, \quad (1.1)$$

where p is the fraction of observations $X_{i_1}, \dots, X_{i_{np}}$ in original sample (X_1, \dots, X_n) such that $X_{i_1} \rightarrow \pm\infty, X_{i_2} \rightarrow \pm\infty, \dots, X_{i_{np}} \rightarrow \pm\infty$.

By definition, b_t is the smallest fraction of outliers in original sample (X_1, \dots, X_n) such that the t -quantile of $T_{n,m}^*$ diverges to infinity. Intuitively, b_t is a measure of the stability of quantile estimates provided by subsampling procedures, with respect to data contaminations of the original sample. In this section, we focus for brevity on one-dimensional real valued statistics, even if, as discussed for instance by Singh (1998) in relation to the bootstrap, our subsampling breakdown point results extend naturally to multivariate and scale statistics. The extension of our theory to the m out of n bootstrap is also straightforward. Asymptotic confidence intervals built by subsampling and the m out of n bootstrap are equivalent for iid observations when $m^2/n \rightarrow 0$; see Politis, Romano and Wolf (1999), Section 2.3, and Andrews and Guggenberger (2009a, 2010a,b). Therefore, for brevity, we focus in the

sequel on subsampling procedures only.

1.3.1 Explicit Breakdown Point Formula for Subsampling Quantiles

The formula for the breakdown point of subsampling quantiles is given in the next theorem.

Theorem 2 *The subsampling upper t -quantile breakdown point is*

$$b_t = \inf\{p \in [1/n, b] : np \in \mathbb{N} \text{ and } P[X(n, m, p) < mb] < t\}, \quad (1.2)$$

where $X(n, m, p)$ is a hypergeometrically distributed random variable with parameters n , np , and m .

From formula (1.2), b_t depends on the quantile probability t , the breakdown point b of T_n , the block size m and the sample size n . It is decreasing in t , and increasing in b , m , for $mb \in$ and n . Moreover, $b_t = b$ for $m = n$. The formula for the subsampling lower t -quantile breakdown point is analogous.

The main implication of Theorem 2 is that it pays to start with a robust statistic T_n having nontrivial breakdown point, to stay away from extreme quantiles, and to avoid small block sizes. Table 1.1 emphasizes this point by computing the subsampling quantile breakdown points when $n = 40, 80, 120$, and for $b = 0.25, 0.5$. The bootstrap quantile breakdown points based on Singh (1998) formula are often close to the ones given by medium subsample sizes.

Theorem 2 implies that we can always obtain a target upper quantile breakdown point $\hat{b}_t \in (1/n, b]$ by selecting a suitable block size $\hat{m}_t = m(\hat{b}_t)$. The formula for the smallest block size ensuring a given upper breakdown point of subsampling quantiles is given below.

Corollary 3 *For given $t \in (0, 1)$, let $\hat{b}_t \in (1/n, b]$ be such that $n\hat{b}_t \in \mathbb{N}$. The smallest block size \hat{m}_t*

such that $b_t \geq \hat{b}_t$ is given by

$$\hat{m}_t = \inf \left\{ m : P \left[\hat{X}(n, m, \hat{b}_t - 1/n) < mb \right] \geq t \right\},$$

where $\hat{X}(n, m, \hat{b}_t - 1/n)$ is a hypergeometrically distributed random variable with parameters n , $n\hat{b}_t - 1$, and m .

Corollary 3 implies that for $\hat{b}_t = b$ it is possible to obtain a breakdown point b_t as large as the one of the statistic T_n . As highlighted by Table 1.1, in order to achieve this goal it is not in general necessary to select a trivial block size $m = n$.

According to Theorem 2, the block size m has to be sufficiently high, in order to avoid undesired subsampling breakdown properties. However to get consistency in a general setting a condition like $m/n \rightarrow 0$ should hold as $n, m \rightarrow \infty$ (see, for instance, Politis, Romano and Wolf, 1999). This means that the application of Corollary 3 is essentially relevant to particular settings for which the consistency of the subsampling holds with $m = O(n)$ (see Wu, 1990, and Remark 2.2.2 in Politis, Romano and Wolf, 1999).

The asymptotic subsampling breakdown behavior is characterized as follows.

Corollary 4 *Let subsampling block size m satisfy $m/n \rightarrow r \in [0, 1)$, $m, n \rightarrow \infty$. Then, $b_t = b - z_t \sqrt{b(1-b)(1-r)}/\sqrt{m} + O(1/m)$, for n large enough, where z_t is the t -quantile of the standard normal distribution.*

From Corollary 4, the subsampling breakdown point b_t converges to the breakdown point of statistic T as $n, m \rightarrow \infty$. Therefore, similar to the asymptotic bootstrap breakdown point formula in Singh (1998), Corollary 4 rules out the breakdown problem of subsampling quantiles for large samples and subsampling block sizes.

1.3.2 Breakdown Point and Data Driven Choice of the Block Size

A main issue in the application of subsampling procedures is the choice of block size m , because the subsampling accuracy heavily depends on this parameter. In this section, we study the robustness of data driven block size selection procedures and derive the breakdown behavior of procedures based on either a minimization of the confidence interval volatility index (MCIV) or the calibration method (CM); see Romano and Wolf (2001). In particular, we are interested in computing the minimal proportion of contamination in the original sample such that the data driven choice of the block size fails and diverges to infinity. Let m_u be the block size selected using MCIV or CM. We consider the following definition for the breakdown point:

Definition 5 *The breakdown point of m_u is defined as*

$$b_t^u := \inf\{p \in [1/n, p] : np \in \mathbb{N} \text{ and } m_u = \infty\}, \quad (1.3)$$

where p is the fraction of observations $X_{i_1}, \dots, X_{i_{np}}$ in original sample (X_1, \dots, X_n) such that $X_{i_1} \rightarrow \pm\infty$, $X_{i_2} \rightarrow \pm\infty, \dots, X_{i_{np}} \rightarrow \pm\infty$.

In the next sections we briefly describe the MCIV and CM approaches, and compute their breakdown points.

Minimum Confidence Interval Volatility Method

A consistent method for a data driven choice of m determines the block size by minimizing the confidence interval volatility index across the admissible values of m . For brevity, we present the method for one-sided intervals. Modifications for the case with two-sided intervals are obvious.

Definition 6 *Let $m_{\min} < m_{\max}$ and $k \in \mathbb{N}$ be fixed. For $m \in \{m_{\min} - k, \dots, m_{\min}, \dots, m_{\max}, \dots, m_{\max} + k\}$ denote by $Q_t^*(m)$ the (lower) t -subsampling quantile for block size m . Further, define $\overline{Q}_t^{*k}(m)$ as the*

average

quantile

$\bar{Q}_t^{*k}(m) := \frac{1}{2k+1} \sum_{j=-k}^{j=k} Q_t^*(m+j)$. The confidence interval volatility (CIV) index is defined for $m \in \{m_{\min}, m_{\min} + 1, \dots, m_{\max} - 1, m_{\max}\}$ by

$$CIV(m) := \frac{1}{2k+1} \sum_{j=-k}^{j=k} \left(Q_t^*(m+j) - \bar{Q}_t^{*k}(m) \right)^2. \quad (1.4)$$

Let $\mathcal{M} := \{m_{\min}, m_{\min} + 1, \dots, m_{\max}\}$. The data driven block size that minimizes the confidence interval volatility index is

$$m_v = \arg \inf_{m \in \mathcal{M}} \{CIV(m) : CIV(m) \in \mathbb{R}^+\}, \quad (1.5)$$

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

The block size m_v minimizes the empirical variance of the upper bound in a subsampling confidence interval with nominal confidence level t . Typical recommended choices for k , m_{\min} and m_{\max} are $k = 2, 3$, $m_{\min} = c_1 n^\zeta$ and $m_{\max} = c_2 n^\zeta$, respectively, where $c_1 \in [0.5, 1]$, $c_2 \in [2, 3]$ and $\zeta = 0.5$; see Romano and Wolf (2001). Moreover, according to Theorem 2, in order to ensure a minimal breakdown point for the quantile of the subsampling distribution, we can select the value of m_{\min} as

$$m_{\min} = \max(c_1 n^\zeta, \hat{m}_t), \quad (1.6)$$

where \hat{m}_t is the minimal subsampling block size in Corollary 3, which ensures a breakdown point larger than \hat{b}_t .

Using Theorem 2, the formula for the breakdown point of m_v follows from Definition 5.

Corollary 7 For given $t \in (0, 1)$, let $b_t(m)$ be the subsampling upper t -quantile breakdown point in

Theorem 2, as a function of the block size $m \in \mathcal{M}$. Then we have: $b_t^v = \sup_{m \in \mathcal{M}} \inf_{j \in \{-k, \dots, k\}} b_t(m+j)$.

Since m_v is a crucial parameter for the accuracy of the resulting subsampling inference, it is convenient to quantify b_t^v for realistic applications. To this end, we can use Corollary 20. For instance, for a sample size $n = 100$ and for $t = 0.99$, we obtain $m_{\min} = 8$ and $m_{\max} = 25$, using the average recommended choice in Romano and Wolf (2001), i.e., $c_1 = 0.75$ and $c_2 = 2.5$. For a statistic with breakdown point $b = 0.1$ and for $k = 3$, this parameter setting implies $b_t^v = 0.03$. In other words, three outliers out of a hundred data points are sufficient to break down the data driven choice of m based on the MCIV index.

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). As above, we present this method for the case of a one-sided confidence interval only. The modifications for two-sided intervals are obvious.

Definition 8 Fix $t \in (0, 1)$ and let (X_1^*, \dots, X_n^*) be a bootstrap sample from (X_1, \dots, X_n) . For each bootstrap sample, denote by $Q_t^{**}(m)$ the t -sub-sampling quantile according to block size m . The data driven block size according to the calibration method is defined by

$$m_c := \arg \inf_{m \in \mathcal{M}} \{ |t - P^*[T_n \leq Q_t^{**}(m)]| : P^*[Q_t^{**}(m) \in \mathbb{R}] > 1 - t \}, \quad (1.7)$$

where, by definition, $\arg \inf(\emptyset) := \infty$, and P^* denote the probability with respect to the bootstrap distribution.

By definition, m_c is the block size for which the bootstrap probability of the event $\{T_n \leq Q_t^{**}(m)\}$ is as near as possible to the nominal level t of the confidence interval, but which at the same time ensures that the subsampling quantile breakdown probability of the calibration method is less than t . The last condition is necessary to ensure that the calibrated block size m_c does not imply a degenerate

subsampling quantile $Q_t^{**}(m_c)$ with a too large probability. By definition, the breakdown point of m_c is the smallest fraction of outliers such that equation (2.12) is degenerate, similar to the MCIV index method. The formula for the breakdown point of m_c is given next.

Corollary 9 *Let $t \in (0, 1)$. The breakdown point of m_c is given by $b_t^c = \max_{m \in \mathcal{M}} \{b_t^{**}(m)\}$, with*

$$b_t^{**}(m) = \inf\{p \in [1/n, b] : np \in \mathbb{N} \text{ and } P[BIN(n, p) < nb_t(m)] < 1 - t\},$$

where $b_t(m)$ is for given $m \in \mathcal{M}$ the quantile subsampling breakdown point in Theorem 2 and $BIN(n, p)$ is a binomial random variable with parameters n and p .

Table 1.2 compares the breakdown point of m_v and m_c for some concrete parameter choices, given a statistic with breakdown point $b = 0.5$.

These theoretical results corroborated by unreported Monte Carlo results in linear regression models indicate a higher robustness of the calibration method relative to the MCIV index method. Therefore, from a robustness perspective, the former should be preferred when consistent bootstrap methods are available. However, as discussed in Romano and Wolf (2001), the application of the calibration method in some settings can be computationally too expensive. In these cases, it is necessary to select an appropriate subset of \mathcal{M} for the admissible block size (see Romano and Wolf (2001), Remark 5.4).

1.3.3 Robust Subsampling

To overcome the problem of the low breakdown point of subsampling quantiles, it is necessary first to apply subsampling methods to robust statistics, in order to avoid a trivial breakdown point from the beginning, and, second, to robustify the subsampling procedure itself. We first show how this goal can be achieved for the class of robust M-estimators, by applying the fast subsampling approach in Hong and Scaillet (2006). This approach, putted forward among others in Davidson and McKinnon

(1999) and Andrews (2002) in relation to the bootstrap, can be used to extend in a convenient way the robust bootstrap procedure for fixed point estimators in Salibian-Barrera, Van Aelst and Willems (2006, 2007) to the robust subsampling setting with M-estimators. In a second step, we study in more detail the linear regression setting, where explicit breakdown point characterizations are possible. We develop robust subsampling procedures for robust MM-estimators and derive a formula for the implied subsampling quantile breakdown point. These results are a natural complement to the theoretical findings obtained in Salibian-Barrera and Zamar (2002) for the robust bootstrap.

Let (X_1, \dots, X_n) be an iid sample governed by the probability law H . We consider the class of robust M-estimators $\hat{\theta}_n$ for parameter $\theta \in \mathbb{R}^d$, defined by the solution of

$$\psi_n(\hat{\theta}_n) = \sum_{i=1}^n f(X_i, \hat{\theta}_n) = 0, \quad (1.8)$$

for some function $\psi_n : \mathbb{R}^d \rightarrow \mathbb{R}^d$ depending on the parameter θ and on the sample (X_1, \dots, X_n) .

Robust M -estimators typically have a bounded estimating function f . This feature is key for developing our robust subsampling approach in the M -estimation setting. As shown previously, a high breakdown point of $\hat{\theta}_n$ does not have to imply a high breakdown point for the corresponding subsampling quantiles. For instance, in Table 1.1, we obtain very low breakdown points of subsampling quantiles, especially for small subsample sizes, even using robust estimators. A second issue is the fact that the application of robust estimators in resampling schemes can rapidly become prohibitive from a computational point of view.

To obtain a robust and computationally feasible subsampling method, we consider the following Taylor expansion of (1.8) around the true parameter value θ_* : $\psi_n(\hat{\theta}_n) = \psi_n(\theta_*) + \nabla \psi_n(\theta_*)(\hat{\theta}_n - \theta_*) + o_P(1)$, where $\nabla \psi_n \in \mathbb{R}^{d \times d}$ is the matrix of partial derivatives with respect to parameter θ . This implies: $(\hat{\theta}_n - \theta_*) = (-\nabla \psi_n(\theta_*))^{-1}(\psi_n(\theta_*)) + o_P(1)$. Thus, we can consider: $(-\nabla \psi_n(\hat{\theta}_n))^{-1}(\psi_{n,m}^*(\hat{\theta}_n))$ as an approximation of $\hat{\theta}_{n,m}^* - \hat{\theta}_n$, where $\psi_{n,m}^*$ is computed from the subsampling block (X_1^*, \dots, X_m^*) .

Given the normalization constant τ_n , the robust subsampling distribution approximating the sampling distribution of $\tau_n(\hat{\theta}_n - \theta_*)$ is defined by

$$L_{n,m}^R(x) = \frac{1}{N_{n,m}} \sum_{s=1}^{N_{n,m}} \mathbb{I}\left\{\tau_m(-\nabla\psi_n(\hat{\theta}_n))^{-1}(\psi_{n,m,s}^*(\hat{\theta}_n)) \leq x\right\}, \quad (1.9)$$

where s indexes the set of possible subsamples, $N_{n,m} = \binom{n}{m}$, and $\mathbb{I}\{\cdot\}$ is the indicator function. The following standard high-level assumptions ensures consistency of the robust subsampling for the class of robust M-estimators; see also Politis, Romano, and Wolf (1999).

$$(A1) \quad \hat{\theta}_n = \theta_* + O_P(1/\tau_n).$$

$$(A2) \quad (-\nabla\psi_n(\hat{\theta}_n))^{-1} = (-\nabla\psi_n(\theta_*))^{-1} + o_P(1).$$

$$(A3) \quad \tau_n(\hat{\theta}_n - \theta_*) = (-\nabla\psi_n(\theta_*))^{-1}\tau_n(\psi_n(\theta_*)) + o_P(1).$$

$$(A4) \quad \text{There exists a limit law } J(H) \text{ such that the distribution of } \tau_n(\hat{\theta}_n - \theta_*) \text{ converges weakly to } J(H).$$

Given Assumptions (A1)-(A4), consistency of the robust subsampling scheme follows in the next theorem.

Theorem 10 *Let Assumptions (A1)-(A4) be satisfied. Assume further that $\tau_m/\tau_n \rightarrow 0$ and $m/n \rightarrow 0$ as $m, n \rightarrow \infty$. Then we get:*

$$(1) \quad \text{If } x \text{ is a continuity point of } J(\cdot, H), \text{ then } L_{n,m}^R(x) \rightarrow J(x, H) \text{ as } n \rightarrow \infty.$$

$$(2) \quad \text{If } J(\cdot, H) \text{ is continuous, then } \sup_x |L_{n,m}^R(x) - J(x, H)| \rightarrow 0 \text{ in probability as } n \rightarrow \infty.$$

$$(3) \quad \text{Given } \alpha \in (0, 1) \text{ define } c_{n,m}(1 - \alpha) = \inf\{x : L_{n,m}^R(x) \geq 1 - \alpha\} \text{ and } c(1 - \alpha, H) = \inf\{x : J(x, H) \geq 1 - \alpha\}. \text{ If } J(\cdot, H) \text{ is continuous at } c(1 - \alpha, H), \text{ it then follows:}$$

$$P\left[\tau_n(\hat{\theta}_n - \theta_*) \leq c_{n,m}(1 - \alpha)\right] \rightarrow 1 - \alpha, \quad \text{as } n \rightarrow \infty.$$

Statements 1–3 in Theorem 10 are standard statements on the weak convergence of the robust subsampling approximation to the true asymptotic distribution $J(H)$ of $\sqrt{n}(\hat{\theta}_n - \theta_*)$. Statement 3 implies that the $(1 - \alpha)$ -quantile of $L_{n,m}$ converges to the corresponding $(1 - \alpha)$ -quantile of $J(H)$. Therefore, the quantities $c_{n,m}(1 - \alpha)$, $\alpha \in (0, 1)$, can be used to construct finite sample tests and confidence intervals for θ_* .

Remark. The fast subsampling approach can be applied also with estimators $\tilde{\theta}_n$ defined by the solution of a set of smooth fixed-point equations $g_n(\tilde{\theta}_n) = \tilde{\theta}_n$, for some function $g_n : \mathbb{R}^d \rightarrow \mathbb{R}^d$ depending on the parameter θ and the sample (X_1, \dots, X_n) . This follows from writing the fixed-point equations in the form $g_n(\tilde{\theta}_n) - \tilde{\theta}_n = 0$. This corresponds to equation (1.8) with $\psi_n = (g_n - Id)$, where Id is the identity function $Id(x) = x$. Consequently, in these cases the robust subsampling is equivalent to the extension of the robust bootstrap approach in Salibián-Barrera, Val Aelst and Willems (2007) to the subsampling setting.

In Section 1.3.4, we characterize explicitly the breakdown point of robust subsampling quantiles in the linear regression setting based on MM-estimates. More generally, we note from the robust subsampling definition (1.9) that the subsampling quantile breakdown point is maximal if: (i) $(-\nabla \psi_n(\hat{\theta}_n))^{-1}$ does not break down as long as $\hat{\theta}_n$ does not break down and (ii) given a subsampling block size m , function $\psi_{n,m}^*(\hat{\theta}_n)$ is bounded with a bound that depends only on the original data set. The last condition is typically satisfied by the estimating functions of robust M-estimators. The first one is often verifiable in concrete model settings.

1.3.4 Robust Subsampling in the Linear Regression Model

We consider the iid linear regression model:

$$Y_i = X_i' \beta + \sigma U_i, \quad i = 1, \dots, n, \quad (1.10)$$

where Y_i is a scalar random variable, X_i an \mathbb{R}^d -valued random variable, $\beta \in \mathbb{R}^d$, and $\sigma \in \mathbb{R}^+$. The joint probability distribution of $(Y_i, X_i)'$ is denoted by H . Several robust estimators of β and σ are available in the literature; see, e.g., Hampel et al. (1986) for a review. We focus on a high-breakdown MM-estimator of β (Yohai, 1987).

Let $\{(y_i, x_i') : i = 1, \dots, n\}$ be a sample of observations of model (1.10). The MM-estimate $\hat{\beta}_n$ of β is defined by the implicit equation:

$$\frac{1}{n} \sum_{i=1}^n \nabla \rho_1 \left(\frac{y_i - x_i' \hat{\beta}_n}{\hat{\sigma}_n} \right) x_i = 0. \quad (1.11)$$

In equation (1.11), $\nabla \rho_1$ is the derivative of a continuously differentiable, bounded and symmetric function ρ_1 , satisfying the assumption (A1)-(A4) below. $\hat{\sigma}_n$ is a scale S -estimate that minimizes with respect to β the M-estimate $\hat{\sigma}_n(\beta)$, defined implicitly by $\frac{1}{n} \sum_{i=1}^n \rho_0 \left(\frac{y_i - x_i' \beta}{\hat{\sigma}_n(\beta)} \right) = B$, where function ρ_0 satisfies the same assumptions as ρ_1 and B is a positive constant. We denote by $\tilde{\beta}_n$ the S -regression estimate, i.e., $\hat{\sigma}_n = \hat{\sigma}_n(\tilde{\beta}_n)$. The choice of B determines the breakdown point of the estimators, which is maximal for $B = 0.5$ (see, e.g., Huber, 1981).

To define the robust subsampling for the linear regression setting we introduce the following notation.

Notation 11 (i) For $i = 1, \dots, n$, define the residuals: $\hat{r}_i = y_i - x_i' \hat{\beta}_n$ and $\tilde{r}_i = y_i - x_i' \tilde{\beta}_n$, and compute the weights: $\hat{w}_i = \nabla \rho_1(\hat{r}_i / \hat{\sigma}_n) / \hat{r}_i$, $\tilde{v}_i = \frac{\hat{\sigma}_n}{nB} \rho_0(\tilde{r}_i / \hat{\sigma}_n) / \tilde{r}_i$. (ii) Given $m < n$, define for every subsampling block $\{(y_i^*, x_i^{*'}) : i = 1, \dots, m\}$ the residuals $\hat{r}_i^* = y_i^* - x_i^{*'} \hat{\beta}_n$ and $\tilde{r}_i^* = y_i^* - x_i^{*'} \tilde{\beta}_n$, and compute the weights:

$$\hat{w}_i^* = \nabla \rho_1(\hat{r}_i^* / \hat{\sigma}_n) / \hat{r}_i^*, \quad \tilde{v}_i^* = \frac{\hat{\sigma}_n}{nB} \rho_0(\tilde{r}_i^* / \hat{\sigma}_n) / \tilde{r}_i^*. \quad (1.12)$$

With these weights, define:

$$\hat{\beta}_{n,m}^* = \left(\sum_{i=1}^m \hat{\omega}_i^* x_i^* x_i^{*'} \right)^{-1} \sum_{i=1}^m \hat{\omega}_i^* x_i^* y_i^*, \quad \hat{\sigma}_{n,m}^* = \sum_{i=1}^m \tilde{v}_i^* (y_i^* - x_i^{*'} \tilde{\beta}_n). \quad (1.13)$$

In equation (1.12), the weights $\hat{\omega}_i^*$ and \tilde{v}_i^* are computed without recalculating the estimators $\hat{\beta}_n$, $\tilde{\beta}_n$ and $\hat{\sigma}_n$ in each subsampling block. The same applies to the quantities $\hat{\beta}_{n,m}^*$ and $\hat{\sigma}_{n,m}^*$ in (1.13), which are therefore only an approximation of the “true” point estimates $\hat{\beta}_m^*$ and $\hat{\sigma}_m^*$ implied by the subsampling block $\{(y_i^*, x_i^{*'}) : i = 1, \dots, m\}$. Following the insight of the previous section, the basic idea is to correct for the asymptotic bias between $(\hat{\beta}_{n,m}^{*'}, \hat{\sigma}_{n,m}^*)$ and $(\hat{\beta}_m^{*'}, \hat{\sigma}_m^*)$ using a first-order linear correction that depends only on $\hat{\beta}_n$, $\tilde{\beta}_n$ and $\hat{\sigma}_n$. In this way, the large breakdown point of these estimators will be inherited by the implied subsampling quantiles. Moreover, since it is not necessary to compute in each subsampling block the implied robust point estimate, the robust subsampling in Definition 12 yields a computationally feasible resampling scheme, which allows us to compute robust confidence intervals for regression parameter β in presence of nuisance scale parameter σ .

Definition 12 Let β_\star be the true parameter value in the regression model (1.10) and $J_n(H)$ be the sampling distribution of $\sqrt{n}(\hat{\beta}_n - \beta_\star)$, i.e., for any $x \in \mathbb{R}^n$: $J_n(x, H) = P \left[\sqrt{n}(\hat{\beta}_n - \beta_\star) \leq x \right]$. The robust subsampling approximation of $J_n(x, H)$ is given by

$$L_{n,m}^R(x) = \frac{1}{N_{n,m}} \sum_{s=1}^{N_{n,m}} \mathbb{I} \left\{ M_n \sqrt{m} (\hat{\beta}_{n,m,s}^* - \hat{\beta}_n) + d_n \sqrt{m} (\hat{\sigma}_{n,m,s}^* - \hat{\sigma}_n) \leq x \right\}, \quad (1.14)$$

where the linear corrections M_n and d_n are defined as follows:

$$M_n = \hat{\sigma}_n \left(\sum_{i=1}^n \nabla^2 \rho_1(\hat{r}_i/\hat{\sigma}_n) x_i x_i' \right)^{-1} \sum_{i=1}^n \hat{\omega}_i x_i x_i',$$

$$d_n = \frac{nB}{\hat{\sigma}_n^2 \sum_{i=1}^n \nabla \rho_0(\tilde{r}_i/\hat{\sigma}_n) \tilde{r}_i/\hat{\sigma}_n} \left(\sum_{i=1}^n \nabla^2 \rho_1(\hat{r}_i/\hat{\sigma}_n) x_i x_i' \right)^{-1} \sum_{i=1}^n \nabla^2 \rho_1(\hat{r}_i/\hat{\sigma}_n) \hat{r}_i x_i.$$

The following are detailed assumptions on the robust linear regression setting based on the above MM-estimator, which ensure consistency of the robust subsampling approximation in Definition 12.

(A5) The sampling distribution $J_n(H)$ converges weakly to a limit distribution $J(H)$ as $n \rightarrow \infty$.

(A6) The following limits in probability hold as $n \rightarrow \infty$: $\hat{\beta}_n \rightarrow \beta_*$, $\tilde{\beta}_n \rightarrow \tilde{\beta}_*$, $\hat{\sigma}_n \rightarrow \sigma_*$, where parameters

β_* , $\tilde{\beta}_*$ and σ_* are the unique solution of the set of moment conditions: $E[\nabla \rho_1((Y_1 - X_1' \beta)/\sigma)] = 0$,

$$E[\rho_0((Y_1 - X_1' \tilde{\beta})/\sigma)] = B, E[\nabla \rho_0((Y_1 - X_1' \tilde{\beta})/\sigma)] = 0.$$

(A7) For $j = 0, 1$, the function ρ_j is three times continuously differentiable and such that: (R1)

$\rho_j(-u) = \rho_j(u)$ for all $u \in \mathbb{R}$; (R2) $\rho_j(0) = 0$; (R3) $\sup_u |\rho_j(u)| = 1$; (R4) If $\rho_j(u) < 1$ and

$0 < v < u$ then $\rho_j(v) < \rho_j(u)$.

(A8) Let $r = Y_1 - X_1' \beta_*$. The following expectations exist:

$$E \left[\frac{\nabla \rho_1(r)}{r} X_1 X_1' \right], E[\nabla \rho_1(r) X_1 X_1'], E[\nabla^2 \rho_1(r) X_1 X_1'], E[\nabla \rho_1(r) r X_1 X_1'], \quad (1.15)$$

$$E \left[\frac{\nabla \rho_0(r)}{r} X_1 X_1' \right], E[\nabla \rho_0(r) r], E[\nabla^2 \rho_0(r) X_1 X_1'], \quad (1.16)$$

$$E[\nabla^2 \rho_0(r) r X_1], E[\nabla^2 \rho_1(r) r X_1].$$

In addition, the first and the third matrices in (1.15) and in (1.16) are invertible, and the second expectation in (1.16) is not zero.

(A9) The following functions are continuous:

$$u \mapsto \frac{\nabla \rho_0(u)}{u}, \quad u \mapsto \frac{\nabla \rho_0(u) - \nabla^2 \rho_0(u)u}{u^2}, \quad u \mapsto \frac{\nabla \rho_1(u) - \nabla^2 \rho_1(u)u}{u^2}.$$

Consistency of the robust subsampling in Definition 12 is stated in the next theorem.

Theorem 13 *Let Assumption (A5)-(A9) be satisfied. Then we get:*

1. *If x is a continuity point of $J(\cdot, H)$, then the following limit in probability holds as $n, m \rightarrow \infty$ and $m/n \rightarrow 0$: $L_{n,m}^R(x) \rightarrow J(x, H)$.*
2. *If $J(\cdot, H)$ is continuous, then the following limit in probability holds as $n, m \rightarrow \infty$ and $m/n \rightarrow 0$:*

$$\sup_x |L_{n,m}^R(x) - J(x, H)| \rightarrow 0.$$
3. *For $\alpha \in (0, 1)$, define $c_{n,m}(1 - \alpha) = \inf\{x : L_{n,m}^R(x) \geq 1 - \alpha\}$, $c(1 - \alpha, H) = \inf\{x : J(x, H) \geq 1 - \alpha\}$. If $J(\cdot, H)$ is continuous at $c(1 - \alpha, H)$, then the following limit holds as $n, m \rightarrow \infty$ and $m/n \rightarrow 0$: $P \left[\sqrt{n}(\hat{\beta}_n - \beta_\star) \leq c_{n,m}(1 - \alpha) \right] \rightarrow 1 - \alpha$.*

In contrast to the general M-estimator case, we can exploit the additional structure of the linear regression setting to explicitly characterize the breakdown point of robust subsampling quantiles. The breakdown point formula for the robust subsampling in Definition 12 is given in the next theorem.

Theorem 14 *Let $\sqrt{\widehat{\omega}_1}x_1, \dots, \sqrt{\widehat{\omega}_n}x_n$ be in general position, i.e., any d row vectors of the $n \times d$ design matrix $X = [\sqrt{\widehat{\omega}_i}x'_i]_{i=1,\dots,n}$ are linearly independent, and fix $t \in (0, 1)$.*

1. *The breakdown point b_t^R of the t -quantile of the robust subsampling in Definition 12 is given by*

$$b_t^R = \inf\{p \in [1/n, b] : np \in \mathbb{N} \text{ and } P[X(n, m, p) \leq m - d] < t\}, \quad (1.17)$$

where $X(n, m, p)$ is a hypergeometrically distributed random variable with parameters n , np , and m , and b is the breakdown point of the robust MM -regression estimator $\hat{\beta}_n$.

2. Let $\hat{b}_t^R \in (1/n, b]$ be such that $n\hat{b}_t^R \in \mathbb{N}$. The smallest block size \hat{m}_t^R such that $b_t^R \geq \hat{b}_t^R$ is given by

$$\hat{m}_t^R = \inf\{m : P\left[\hat{X}(n, m, \hat{b}_t^R - 1/n) \leq m - d\right] \geq t\},$$

where $\hat{X}(n, m, \hat{b}_t^R - 1/n)$ is a hypergeometrically distributed random variable with parameters n , $n\hat{b}_t^R - 1$, and m .

The assumption on the general position of $\sqrt{\hat{\omega}_1}x_1, \dots, \sqrt{\hat{\omega}_n}x_n$ is also used in Salibián-Barrera and Zamar (2002), and is needed here to ensure that the approximation $\hat{\beta}_{n,m}^*$ of the subsampling estimate $\hat{\beta}_m^*$ is well-defined in every subsampling block. By comparing (1.17) with the breakdown formula (1.2) of the standard subsampling in Theorem 2, we note that for reasonable parameter choices $mb \ll m - d = m(1 - d/m)$. Therefore, $P[X(n, m, p) < mb] \ll P[X(n, m, p) \leq m - d]$ and $b_t^R \gg \hat{b}_t^R$. The numerical difference between the two breakdown points can be large. Table 1.3 computes the robust subsampling breakdown point for a setting with $d = 3$ and for sample sizes $n = 40, 80, 120$, in dependence of the breakdown point b of the MM -regression estimator $\hat{\beta}_n$. We find that statement 2 of Theorem 14 can be more relevant for applications than the one of Corollary 3. This is so because a large breakdown point of the robust subsampling can arise also for small subsampling block sizes, which asymptotically can more easily ensure the subsampling consistency conditions.

For $b = 0.225$ and $n = 40$, the robust subsampling breakdown point is $b_t^R = 0.225$ for all $m \geq 6$. For $t = 0.9$, the maximal breakdown point is obtained already for $m = 8$. For $t = 0.95$ and $t = 0.99$, it is obtained for $m = 10$ and $m = 12$, respectively. In general, the maximal breakdown point is obtained for all samples sizes and confidence levels in Table 1.3, independently of b , for $m = 14$. When $b < 0.5$,

the value of m ensuring the maximal breakdown point is even lower. These are large differences with respect to the subsampling breakdown points in Table 1.1.

These results have implications also for the breakdown point of m_v in Corollary 20. For instance, with a sample size $n = 100$, the average recommended choice in Romano and Wolf (2001) yields $m_{\min} = 8$ and $m_{\max} = 25$ (using $c_1 = 0.75$ and $c_2 = 2.5$). For $b = 0.1$ and $k = 3$, the breakdown point of m_v when using the robust subsampling is maximal for all confidence levels, but the one when using the standard subsampling is $b_t^v = 0.03$ for $t = 0.99$. On the other side, the breakdown point of m_c is much higher, as was shown by our previous numerical computations.

Remark. Our results on the robust subsampling extend directly to linear regression models with fixed designs. Assume that the covariates $X_i \in \mathbb{R}^d$ are fixed and part of an infinite sequence $(X_1, \dots, X_n, X_{n+1}, \dots)$. For $Z_i = (Y_i, X_i')'$, let H be the joint probability law governing the infinite sequence $(Z_1, \dots, Z_n, Z_{n+1}, \dots)$. Salibian-Barrera (2006a) proves consistency and asymptotic normality of MM-estimators in this setting. Moreover, Salibian-Barrera (2006b) shows the validity of the first order Taylor expansion for the corresponding fixed-point estimating equation. Under the weak assumptions of Theorem 4.3.1 in Politis, Romano and Wolf (1999), the consistency of the robust subsampling in Definition 1.14 then follows even for settings with fixed designs.

1.4 Monte Carlo Study and Sensitivity Analysis

We study through Monte Carlo simulations the statistical properties of the subsampling and the robust subsampling in estimating (i) the distribution of the square of the sample average for an iid normal sample and (ii) the confidence interval of a parameter of interest in an iid linear regression model (1.10) when this parameter is possibly near a boundary. In both settings, the bootstrap is inconsistent, but subsampling procedures are applicable.

1.4.1 Square of the Sample Average

The first example considers the sampling distribution of the square of the sample average based in an iid normal sample. This is an informative, albeit simple, design to measure the accuracy of the subsampling in presence of model contaminations. As discussed, e.g., in Datta (1995), the bootstrap fails in this setting and only a modified bootstrap procedure is applicable. Instead, the subsampling is consistent without modifications of the standard procedure.

Model and Estimation

Let (X_1, \dots, X_n) be an iid sample with $X_i \sim N(\mu, \sigma^2)$, where $\mu = 0$ and $\sigma^2 = 1$, and $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ be the sample average. Using subsampling, we estimate the distribution of $n((\bar{X}_n)^2 - \mu^2)$. Let (X_1^*, \dots, X_m^*) be a random subsample and $\bar{X}_{n,m}^* = \frac{1}{m} \sum_{i=1}^m X_i^*$ be the subsample average. Then, the subsampling distribution approximation reads:

$$L_{n,m}^{NR}(x) = \frac{1}{N_{n,m}} \sum_{s=1}^{N_{n,m}} \mathbb{I} \{ m((\bar{X}_{n,m,s}^*)^2 - (\bar{X}_n)^2) \leq x \}. \quad (1.18)$$

For the robust subsampling, we consider the robust location estimate \bar{X}_n^R given as solution of the equation $\psi_n(\bar{X}_n^R) = 0$, where function ψ_n is defined by

$$\psi_n(\theta) = \frac{1}{n} \sum_{i=1}^n h_c(X_i - \theta), \quad (1.19)$$

and $h_c(x) = x \cdot \min(1, c/|x|)$ is the Huber function. To estimate the distribution of $n((\bar{X}_n^R)^2 - \mu^2)$, we start by considering the subsample statistic $m((\bar{X}_n^{R*})^2 - (\bar{X}_n^R)^2) = m((\bar{X}_n^{R*} - \bar{X}_n^R)^2 + 2\bar{X}_n^R(\bar{X}_n^{R*} - \bar{X}_n^R))$, where \bar{X}_n^{R*} is the subsample robust estimate of μ . This approximation does not directly generate a robust subsampling distribution. Therefore, in the last expression we use our robust subsampling approach to estimate the distribution of $(\bar{X}_n^{R*} - \bar{X}_n^R)$. Finally, the robust subsampling approximation

of the distribution of $n((\bar{X}_n^R)^2 - \mu^2)$ is defined by

$$L_{n,m}^R(x) = \frac{1}{N_{n,m}} \sum_{s=1}^{N_{n,m}} \mathbb{I} \{ m((A_{n,m,s})^2 + 2(\bar{X}_n^R)(A_{n,m,s})) \leq x \}, \quad (1.20)$$

where $A_{n,m,s} = (-\nabla \psi_n(\bar{X}_n^R))^{-1} \psi_{n,m,s}^*(\bar{X}_n^R)$.

Numerical Results

We consider sample sizes $n = 40, 80, 120$. Since the calibration method is not applicable here, we use a data driven block size obtained by minimizing the CIV index for $k = 2$. For sample sizes $n = 40, 80, 120$, the average recommended choice in Romano and Wolf (2001) implies a lower and upper bound $m_{\min} = 5, 7, 8$ and $m_{\max} = 16, 22, 27$ for $c_1 = 0.75$ and $c_2 = 2.5$, respectively.

We study the finite sample coverage implied by the classical and robust subsampling methods. To this end, we test the null hypothesis $\mathcal{H}_0 : (\mu_\star)^2 = 0$ against $\mathcal{H}_1 : (\mu_\star)^2 > 0$, under a contaminated normal distribution for X_i :

$$X_i \sim (1 - \gamma)N(0, 1) + \frac{\gamma}{2}(N(5, 100) + N(-5, 100)),$$

for $\gamma = 0$ (no contamination), $\gamma = 0.05$ (5% of contaminated data) and $\gamma = 0.10$ (10% of contaminated data). For each of the 2000 Monte Carlo replications, subsampling distributions are computed based on 500 draws. Tables 1.4 summarizes the empirical frequencies of non rejection of null hypothesis \mathcal{H}_0 for a confidence level $1 - \alpha = .95$. In all Monte Carlo simulation settings, we find that the empirical frequencies for the robust subsampling are quite accurate and closer to the nominal frequencies than those of the classical subsampling. Under a model contamination, the underrejection of the null hypothesis using classical subsampling methods can be severe. For all sample sizes we get an empirical rejection frequency of .99 instead of the true .95 value when $\gamma = 0.10$. At the same time, the one-

sided confidence interval implied by the subsampling virtually explodes in presence of contamination, leading to a virtually non informative inference. For instance, when $\gamma = 0.10$ and $n = 120$, the median .95-quantile implied by the subsampling approximation is 48.42 while the one implied by the robust subsampling approximation is 5.56.

1.4.2 Linear Regression

In this section we consider the iid linear regression model (1.10) when a parameter of interest is possibly near a boundary (see, e.g., Kim, Stone and White (2005) for an application in finance). As discussed in detail by Andrews (2000), the bootstrap is inconsistent in this context and the subsampling is a potentially natural alternative to it. Moreover, Andrews and Guggenberger (2009a, 2010a,b) (see also Mikusheva (2007) for a similar problem in autoregressive models with unit roots) show that pure subsampling methods have a lack of uniform asymptotic approximation within a class of models including our Monte Carlo setting. They also develop hybrid and size-correction procedures to fix the arising asymptotic size distortion. Analogous remarks hold in the linear regression model when making inference on a parameter of a given regressor and the parameter of another regressor, a nuisance parameter, may be near a boundary. We follow their hybrid approach in our Monte Carlo study of the classic and robust subsampling.

Model and Estimation

We consider the regression parameter $\beta \in \mathbb{R}$, which is known to satisfy the constraint $\beta \geq 0$ in the iid linear regression model:

$$\begin{aligned} Y_i &= X_i' \theta + W_i \beta + \sigma U_i, \\ &= Z_i' \eta + \sigma U_i, \quad i = 1, \dots, n, \end{aligned} \tag{1.21}$$

where Y_i , W_i are scalar, X_i is an \mathbb{R}^{d-1} -valued random variable, $\theta \in \mathbb{R}^{d-1}$, and $\sigma \in \mathbb{R}^+$. Moreover, $Z_{i(1)} = X_{i(1)} = 1$, $\eta_{(1)} = \theta_{(1)}$, $Z_{i(2)} = W_i$, $\eta_{(2)} = \beta$ and for $3 \leq j \leq d$, $Z_{i(j)} = X_{i(j-1)}$, $\eta_{(j)} = \theta_{(j-1)}$, where $h_{(j)}$ denotes the j -th coordinate of vector h . The common joint distribution of $(Y_i, Z_i)'$ is denoted by H . Let $\{(y_i, z_i)' : i = 1 \dots, n\}$ be a sample of observations of model (1.21). In order to construct confidence intervals for parameter β using the classic subsampling, we consider the constrained estimator $\hat{\beta}_n^{NR} = \max(0, \hat{\eta}_n^{ols})$, where $\hat{\eta}_n^{ols}$ is the (unrestricted) OLS estimator of η . For the robust subsampling, we consider the constrained estimator $\hat{\beta}_n^R = \max(0, \hat{\eta}_n^{rob})$, where $\hat{\eta}_n^{rob}$ is a MM-estimator of η . The S-estimate $\hat{\sigma}_n = \hat{\sigma}_n(\tilde{\eta}_n^{rob})$ is computed from a constrained robust estimator $\tilde{\eta}_n^{rob}$ of η under the constraint $\eta_{(2)} \geq 0$.

Given subsampling blocks $\{(y_i^*, z_i^{*'})' : i = 1, \dots, m\}$, we construct consistent subsampling and robust subsampling methods as follows. For the subsampling, we compute in each block the constrained estimator $\hat{\beta}_{n,m}^{NR} = \max(0, \hat{\eta}_{n,m(2)}^{ols})$. The subsampling distribution function estimating the distribution function of $\sqrt{n}(\hat{\beta}_n^{NR} - \beta_*)$ is then given by

$$L_{n,m}^{NR}(x) = \frac{1}{N_{n,m}} \sum_{s=1}^{N_{n,m}} \mathbb{I} \left\{ \sqrt{m}(\hat{\beta}_{n,m,s}^{NR} - \hat{\beta}_n^{NR}) \leq x \right\}. \quad (1.22)$$

For the robust subsampling, we follow (1.13) and (1.14) and additionally account for the parameter constraint. Thus, we consider the robust subsampling statistic:

$\hat{sub}\beta_{n,m}^* = \max \left((M_n(\hat{\eta}_{n,m}^* - \hat{\eta}_n^{rob}) + d_n(\hat{\sigma}_{n,m}^* - \hat{\sigma}_n))_{(2)} + \hat{\beta}_n^R, 0 \right)$. The robust subsampling distribution function which approximates the distribution function of $\sqrt{n}(\hat{\beta}_n^R - \beta_*)$ is then given by

$$L_{n,m}^R(x) = \frac{1}{N_{n,m}} \sum_{s=1}^{N_{n,m}} \mathbb{I} \left\{ \sqrt{m}(\hat{sub}\beta_{n,m,s}^* - \hat{\beta}_n^R) \leq x \right\}. \quad (1.23)$$

By construction, the theoretical results in Section 1.3.4 for the upper quantile breakdown of the robust subsampling distribution in Definition 12 hold also for (1.23).

Hybrid Procedures

Using (1.22) and (1.23), we construct hybrid, classical and robust, equal-tailed confidence intervals for parameter β as follows. Let $c_{n,m}(1 - \alpha)$ be the $(1 - \alpha)$ -quantile implied by either (1.22) or (1.23). The corresponding hybrid quantile is:

$$c_{n,m}^H(1 - \alpha) = \max(c_{n,m}(1 - \alpha), c_\infty(1 - \alpha)), \quad (1.24)$$

where $c_\infty(1 - \alpha)$ is the quantile of the asymptotic distribution of $\sqrt{n}(\hat{\beta}_n^i - \beta_\star)$, $i = CL, R$, for the unconstrained, either classical or robust, estimator $\hat{\beta}_n^i$ of β . To compute $c_\infty(1 - \alpha)$, one can use standard asymptotic normality results for OLS estimators and the asymptotic normality results for MM-estimates in Yohai (1987). However, Salibian-Barrera and Zamar (2002) show that these asymptotic approximations behave poorly in presence of contamination. Therefore, we use the bootstrap and the robust bootstrap, for the subsampling and the robust subsampling, respectively, to estimate the distribution of the unconstrained estimators in the computation of hybrid quantiles. Unreported numerical results confirm the superiority of this approach. In this way, the construction of hybrid quantiles for our robust subsampling approach can profit also from the robustness properties of the robust bootstrap developed in Salibian-Barrera and Zamar (2002).

Numerical Results

We consider the iid linear regression model (1.21) for $d = 3, 5$. The true parameter vector is $\eta_\star = (0, \beta_\star, 0)'$ and $\eta_\star = (0, \beta_\star, 0, 0, 0)'$, respectively, with $\beta_\star = 0.25$, which is a parameter value near to the boundary 0. Since the calibration method is not applicable here, we use a data driven block size obtained by minimizing the CIV index with $k = 2$. For our sample sizes $n = 40, 60$, the average recommended choice in Romano and Wolf (2001) implies an upper bound $m_{\max} = 16$, $m_{\max} =$

19 for $c_2 = 2.5$ respectively. For the lower bound, we apply equation (1.6) in order to obtain a breakdown point of at least 40% for the 0.975 quantile. This implies $m_{\min} \geq 9, 13$ for $d = 3, 5$, respectively, restricting the standard choice of m_{\min} , especially for $d = 5$. In order to allow for a non trivial data-driven block size selection when $d = 5$, we then set $m_{\max} = 18, 23$ in this case for $n = 40, 60$, respectively. We make use of functions ρ_0 and ρ_1 in Tukey's family. The constant for the MM -regression estimator in our simulations is $B = 0.5$. For this choice, we obtain a breakdown point of $\hat{\eta}_n$ satisfying $b \geq 0.47$; see Yohai (1987, Theorem 2.1).

We first study the finite sample coverage implied by classical and robust subsampling methods. To this end, we test the null hypothesis $\mathcal{H}_0 : \beta_\star = 0.25$ under a contaminated normal distribution for U :

$$U \sim (1 - \gamma)N(0, 1) + \frac{\gamma}{2}(N(C, (0.1)^2) + N(-C, (0.1)^2)), \quad (1.25)$$

where $C = 5$, $\gamma = 0$ (no contamination), $\gamma = 0.15$ (15% of contaminated data) and $\gamma = 0.25$ (25% of contaminated data), as in Salibián-Barrera and Zamar (2002). For each of the 2000 Monte Carlo replications, subsampling distributions are based on 200 draws. Tables 1.5 summarizes the empirical frequencies of non rejection of null hypothesis \mathcal{H}_0 and the median confidence interval lengths for the confidence level $1 - \alpha = .95$.

In all Monte Carlo simulation settings, the empirical frequencies for the robust subsampling with data driven choice of the block size are quite accurate and closer to the nominal frequencies than those of the classical subsampling. Unreported results for the inconsistent bootstrap and robust bootstrap yield empirical rejection frequencies between 67.5% to 72.3%. Similarly, the subsampling and robust subsampling without hybrid correction yield empirical frequencies between 60.1% and 65.4%, which are not too far away from the theoretical distorted asymptotic size $(1 - \alpha)/2$ of equal-tailed confidence intervals; see Andrews and Guggenberger, 2010b. Unreported results for the parameter choice

$\beta_\star = 0.1$ show that the undercoverage of inconsistent procedures is, as expected, larger, while hybrid robust methods maintain an accurate coverage. Results of the robust bootstrap and hybrid robust subsampling for the parameter choice $\beta_\star = 0.5$ are more similar, as expected, but still in favour of the latter. Unreported results with different contamination sizes, model dimensions and sample sizes (e.g., $C = 4$, $d = 20$) produced similar results.

The median length of the robust subsampling confidence intervals is moderately higher in the setting with no contamination ($\gamma = 0\%$). For instance, for the case $n = 40$, $d = 3$, the median confidence interval of the robust subsampling is approximately 14% higher than the median length of the subsampling. However, in presence of contamination the robust subsampling produces clearly more efficient inferences with dramatically smaller median confidence interval lengths. For instance, for the case $n = 40$, $d = 3$, the median confidence interval of the robust subsampling is approximately 38% (28%) lower than the median length of the subsampling when $\gamma = 15\%$ ($\gamma = 25\%$). These are large differences having obvious implications for the power of tests based on subsampling and robust subsampling methods.

We have also studied the sensitivity of the subsampling and robust subsampling inference with respect to empirical contaminations of the data. For each Monte Carlo sample, let:

$$Y_{\max} = \arg \max_{Y_1, \dots, Y_n} \{u(Y_i) | u(Y_i) = Y_i - Z_i' \eta, \text{ under } \mathcal{H}_0\}, \quad (1.26)$$

We modify Y_{\max} over a grid within the interval $[Y_{\max} + 1, Y_{\max} + 4]$. Then, we analyze the sensitivity of the resulting empirical averages of p -values for testing the null hypothesis $\mathcal{H}_0 : \beta_\star = 0.25$. Figure 1.1 summarizes the results.

As expected, we obtain quite large absolute variations in average p -values for the subsampling and an almost flat sensitivity curve for the robust subsampling.

Finally, as a last exercise we have computed the average p -value for Monte Carlo samples generated

under $\mathcal{H}_0 : \beta_\star = 0.25$, with increasing contamination sizes $\gamma \in [0, 0.25]$ in (1.25), and have analyzed the average p -value variation with respect to the setting with no contamination ($\gamma = 0$). Figure 1.2 summarizes the results.

Also in this case, the subsampling clearly implies larger variations in average p -values as a function of the size of contamination in the data, indicating the fragility of the implied inference results.

1.5 Conclusions

We derive a formula for the breakdown point of subsampling quantiles, which is shown to imply fragile subsampling procedures for moderate block sizes, even when subsampling is applied to robust statistics. This instability is inherited by data driven block size selection procedures. We propose consistent robust subsampling methods for the class of M-estimators and derive detailed breakdown point formulas for MM-estimators in the linear regression setting. Monte Carlo simulations in two settings where the bootstrap is known to fail show the usefulness of robust subsampling relative to the classical subsampling for producing accurate inferences in presence of model deviations.

A.1 Proofs

Proof of Theorem 2. The quantile Q_t^* breaks down if and only if the proportion of bounded realizations of the statistic $T_{n,m}^*$ is less than t , i.e., when the proportion of subsamples with less than mb outliers is less than t . Let $X(n, m, p)$ be the number of outliers in subsample (X_1^*, \dots, X_m^*) , when np is the number of outliers in the original sample (X_1, \dots, X_n) . The random variable $X(n, m, p)$ follows a hypergeometric distribution with parameters n , np , and m . Consequently, b_t is the smallest proportion p such that $np \in$ and $P[X(n, m, p) < mb] < t$, which is the stated result. ■

Proof of Corollary 3. Existence of \hat{m}_t is ensured by Theorem 2. For a hypergeometrically distributed variable $X(n, m, p)$ such that $np \in$, the probability $P[X(n, m, p) < mb]$ is decreasing in p . Therefore, $b_t(\hat{m}_t) \geq \hat{b}_t$. By definition, for every integer $m < \hat{m}_t$, $P[\hat{X}(n, m, \hat{b}_t - 1/n) < mb] < t$, and $b_t(m) \leq \hat{b}_t - 1/n$. This concludes the proof. ■

Proof of Corollary 4. Let us take $p = b - z_t \sqrt{b(1-b)(1-r)}/\sqrt{m} + c/m$, for $p \in [0, b]$, and compute a Berry-Esseen type bound for the normal approximation of the hypergeometric distribution, where c is in a fixed compact set. For n and c large enough, $P[X(n, m, p) < mb] < t$, where $X(n, m, p)$ is a hypergeometric random variable with parameters n , np , and m . For n large enough and c small enough, $P[X(n, m, p) < mb] > t$. Therefore, $b_t = b - z_t \sqrt{b(1-b)(1-r)}/\sqrt{m} + O(1/m)$, as stated. ■

Proof of Corollary 20. By definition, in order to get $m_v = \infty$ 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 $\{X_1, \dots, X_n\}$ satisfies $p \geq \min\{b_t(m-k), b_t(m-k+1), \dots, b_t(m+k-1), b_t(m+k)\}$. This concludes the proof. ■

Proof of Corollary 22. By definition, in order to get $m_c = \infty$ we must have $P[Q_t^{**}(m) = \infty] \geq t$ for all $m \in \mathcal{M}$. $Q_t^{**}(m) = \infty$ if the number of outliers in bootstrap sample (X_1^*, \dots, X_n^*) is at least as large as $nb_t(m)$. The number of outliers in the bootstrap sample is distributed as $B(n, p)$. This concludes the proof. ■

Proof of Theorem 10. Under Assumptions (A1)-(A4) the statements of the theorem follow from Theorem 1 in Hong and Scaillet (2006). ■

Proof of Theorem 13. We first rewrite the estimator $\tau_n = (\hat{\beta}'_n, \hat{\sigma}_n, \tilde{\beta}'_n)'$ as the fixed point of the following system of equations:

$$\begin{aligned}\hat{\beta}_n &= A_n(\hat{\beta}_n, \hat{\sigma}_n)^{-1} V_n(\hat{\beta}_n, \hat{\sigma}_n), \\ \hat{\sigma}_n &= \hat{\sigma}_n U_n(\tilde{\beta}_n, \hat{\sigma}_n), \\ \tilde{\beta}_n &= B_n(\tilde{\beta}_n, \hat{\sigma}_n)^{-1} W_n(\tilde{\beta}_n, \hat{\sigma}_n),\end{aligned}\tag{1.27}$$

where

$$\begin{aligned}A_n(\beta, \sigma) &= \frac{1}{n} \sum_{i=1}^n \frac{\nabla \rho_1((y_i - x'_i \beta)/\sigma)}{y_i - x'_i \beta} x_i x'_i, \\ V_n(\beta, \sigma) &= \frac{1}{n} \sum_{i=1}^n \frac{\nabla \rho_1((y_i - x'_i \beta)/\sigma)}{y_i - x'_i \beta} y_i x_i, \\ U_n(\tilde{\beta}, \sigma) &= \frac{1}{n} \sum_{i=1}^n \frac{\rho_0((y_i - \tilde{\beta}' x_i)/\sigma)}{B(y_i - \tilde{\beta}' x_i)} (y_i - \tilde{\beta}' x_i),\end{aligned}$$

and

$$\begin{aligned} B_n(\tilde{\beta}, \sigma) &= \frac{1}{n} \sum_{i=1}^n \frac{\nabla \rho_0((y_i - x_i' \tilde{\beta})/\sigma)}{y_i - x_i' \tilde{\beta}} x_i x_i', \\ W_n(\tilde{\beta}, \sigma) &= \frac{1}{n} \sum_{i=1}^n \frac{\nabla \rho_0((y_i - x_i' \tilde{\beta})/\sigma)}{y_i - x_i' \tilde{\beta}} y_i x_i. \end{aligned}$$

More compactly, the system (1.27) can be written as $\tau_n = F_n(\tau_n)$ for an appropriate function $F_n : \mathbb{R}^{2d+1} \rightarrow \mathbb{R}^{2d+1}$. A first order expansion of (1.27) gives

$$\sqrt{n}(\tau_n - \tau_\star) = [I - \nabla F_n(\tau_\star)]^{-1} \sqrt{n}(F_n(\tau_\star) - \tau_\star) + o_P(1), \quad (1.28)$$

where $\tau_\star = (\beta_\star', \sigma_\star, \tilde{\beta}_\star')'$. The explicit computation of ∇F_n shows that $\tilde{\beta}_n$ does not enter (1.28) in the approximation of the first $d+1$ components of $\sqrt{n}(\tau_n - \tau_\star)$, i.e., the approximation of $\sqrt{n}(\hat{\beta}_n - \beta_\star)$ and $\sqrt{n}(\hat{\sigma}_n - \sigma_\star)$. The $d \times d$ matrix M_n in (1.15) is the left upper diagonal block of $[I - \nabla F_n(\tau_n)]^{-1}$, and the vector d_n in (1.15) is the $d+1$ -th upper d -dimensional column of this matrix. Summarizing, we obtain the approximation:

$$\begin{aligned} \sqrt{n}(\hat{\beta}_n - \beta_\star) &= M_{n,\star} \sqrt{n}(A_n(\beta_\star, \sigma_\star)^{-1} V_n(\beta_\star, \sigma_\star) - \beta_\star) \\ &\quad + d_{n,\star} \sqrt{n}(\sigma_\star U_n(\tilde{\beta}_\star, \sigma_\star) - \sigma_\star) + o_P(1) \\ &=: \xi_n(\tau_\star) + o_P(1), \end{aligned}$$

where $M_{n,\star}$ and $d_{n,\star}$ are the same matrix and the same vector as in (1.15) and (1.15), respectively, but evaluated at τ_\star instead of τ_n . Therefore, we have to show that the limit distribution of $\xi_n(\tau_\star)$ is

the same as the limit distribution of

$$\begin{aligned}
\xi_{n,m}^* &= M_n \sqrt{m} (A_{n,m}^*(\hat{\beta}_n, \hat{\sigma}_n)^{-1} V_{n,m}^*(\hat{\beta}_n, \hat{\sigma}_n) - \hat{\beta}_n) \\
&\quad + d_n \sqrt{m} (\hat{\sigma}_n U_{n,m}^*(\tilde{\beta}_n, \hat{\sigma}_n) - \hat{\sigma}_n) \\
&= M_n \sqrt{m} (\hat{\beta}_{n,m}^* - \hat{\beta}_n) + d_n \sqrt{m} (\hat{\sigma}_{n,m}^* - \hat{\sigma}_n).
\end{aligned}$$

To this end, it is sufficient to prove that the limit distribution of $\zeta_{n,m}^*(\tau_n) := \sqrt{m}(F_{n,m}^*(\tau_n) - \tau_n)$ is the same as the limit distribution of $\zeta_n(\tau_*) := \sqrt{n}(F_n(\tau_*) - \tau_*)$. In order to obtain this, we only need to show that the U -statistic defined by $U_{n,m}(x) = \frac{1}{N_{n,m}} \sum_{s=1}^{N_{n,m}} \mathbb{I} \{ \sqrt{m}(F_{n,m,s}^*(\tau_n) - \tau_*) \leq x \}$ converges to the limit cumulative distribution of $\zeta_n(\tau_*)$, evaluated at any continuity point x . This implication follows, however, with standard arguments; see, e.g., the proof of Theorem 2.2.1 in Politis, Romano and Wolf (1999). ■

Proof of Theorem 14. Under the assumptions of the theorem, we can use the same arguments as in the proof of Theorem 2 in Salibián-Barrera and Zamar (2002) to show that, given a subsampling block of size m , the approximation $\hat{\beta}_{n,m}^*$ is bounded, with a bound that depends only on the original data set, if at least d observations in the block are not outliers. Moreover, $\sigma_{n,m}^*$ remains bounded for every subsampling block. Therefore, the robust subsampling approximation in Definition 12 breaks down if and only if in the subsampling block the number $X(n, m, p)$ of outliers is larger than $m - d$. The proportion p of outliers in the original sample that is needed to drive the t -th subsampling quantile estimate above any bound should then satisfy:

$$P[X(n, m, p) > m - d] \geq 1 - t. \quad (1.29)$$

This proves statement (i) of Theorem 14, after taking complements of the event in (1.29). Statement (ii) now follows with the same arguments used to prove Corollary 3. ■

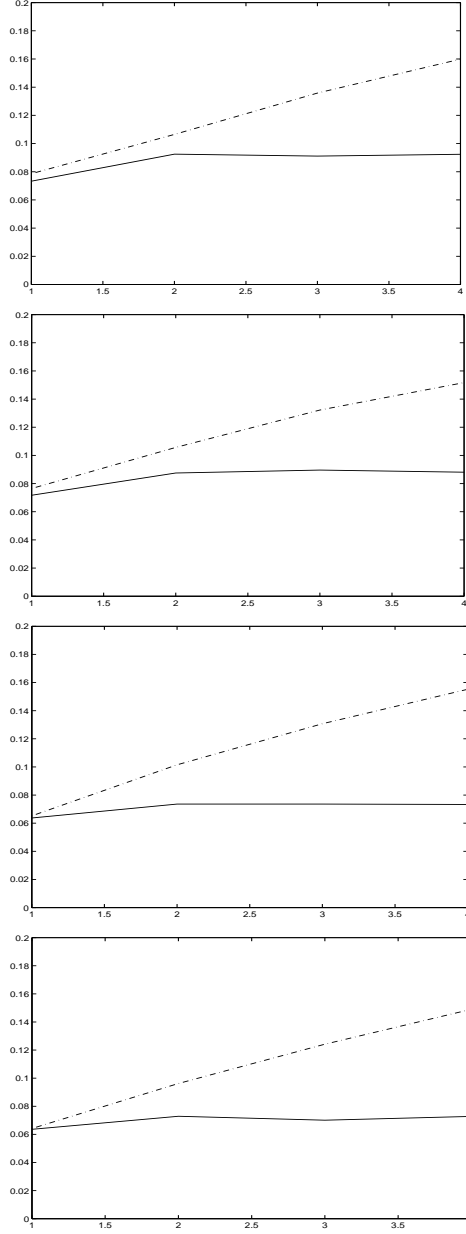


Figure 1.1: **Sensitivity analysis.** Sensitivity plots of the absolute variation of the empirical p -value average, for a test of the null hypothesis $\mathcal{H}_0 : \beta_\star = 0.25$, with respect to variations of Y_{\max} , in each Monte Carlo sample, within the interval $[1, 4]$. The random samples were generated under \mathcal{H}_0 and, from the top to the bottom, with $n = 40$ and $d = 3$, $n = 40$ and $d = 5$, $n = 60$ and $d = 3$, and $n = 60$ and $d = 5$, respectively. We consider the classical subsampling (dash-dotted line) and the robust subsampling (straight line), with the MCIV method for the block size selection.

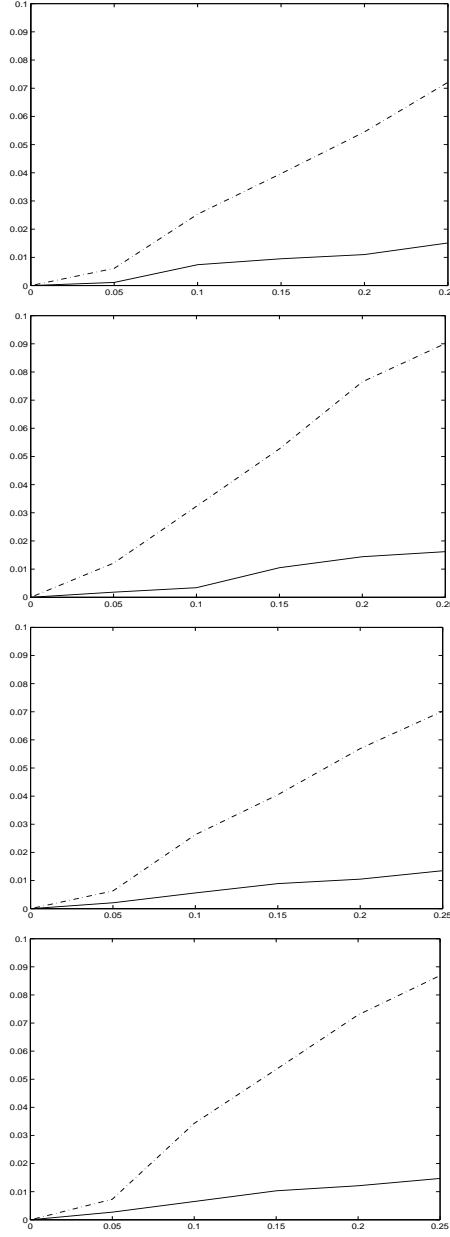


Figure 1.2: **Breakdown point analysis.** Sensitivity plots of the empirical p -value average for a test of the null hypothesis $\mathcal{H}_0 : \beta_\star = 0.25$. Each p -value average is computed using Monte Carlo samples generated with contamination probabilities $\gamma \in [0, 0.25]$. The graphs plot the difference in average p -value relative to the case with no contamination ($\gamma = 0$). The random samples are generated under \mathcal{H}_0 and, from the top to the bottom, with $n = 40$ and $d = 3$, $n = 40$ and $d = 5$, $n = 60$ and $d = 3$, and $n = 60$ and $d = 5$. We consider the classical subsampling (dash-dotted line) and the robust subsampling (straight line), using the MCIV method for the block size selection.

$n = 40$	$t = .9$	$t = .95$	$t = .99$	$n = 40$	$t = .9$	$t = .95$	$t = .99$
B. Point	.2500	.2500	.2500	B. Point	.5000	.5000	.5000
Boots.	.1750	.1500	.1250	Boots.	.4000	.3750	.3250
Subs.				Subs.			
$m = 5$.1250	.1000	.0500	$m = 5$.2750	.2250	.1500
$m = 10$.1500	.1250	.0750	$m = 10$.3000	.2750	.2000
$m = 20$.1750	.1500	.1250	$m = 20$.4000	.3750	.3250
$m = 30$.2250	.2000	.2000	$m = 30$.4500	.4250	.4000
$m = 37$.2500	.2500	.2500	$m = 39$.5000	.5000	.5000
$n = 80$	$t = .9$	$t = .95$	$t = .99$	$n = 80$	$t = .9$	$t = .95$	$t = .99$
B. Point	.2500	.2500	.2500	B. Point	.5000	.5000	.5000
Boots.	.1875	.1750	.1500	Boots.	.4250	.4125	.3750
Subs.				Subs.			
$m = 10$.1250	.1000	.0625	$m = 10$.2875	.2375	.1750
$m = 20$.1500	.1250	.1000	$m = 20$.3625	.3250	.2750
$m = 30$.1875	.1625	.1375	$m = 30$.4000	.3750	.3375
$m = 40$.1875	.1750	.1500	$m = 40$.4250	.4000	.3750
$m = 50$.2125	.2000	.1875	$m = 50$.4375	.4250	.4000
$m = 60$.2125	.2125	.2000	$m = 60$.4625	.4500	.4250
$m = 70$.2375	.2250	.2250	$m = 70$.4750	.4625	.4500
$m = 77$.2500	.2500	.2500	$m = 79$.5000	.5000	.5000
$n = 120$	$t = .9$	$t = .95$	$t = .99$	$n = 120$	$t = .9$	$t = .95$	$t = .99$
B. Point	.2500	.2500	.2500	B. Point	.5000	.5000	.5000
Boots.	.2000	.1917	.1667	Boots.	.4417	.4250	.3917
Subs.				Subs.			
$m = 10$.1250	.1000	.0583	$m = 10$.2750	.2333	.1667
$m = 20$.1417	.1167	.0833	$m = 20$.3500	.3167	.2667
$m = 40$.1750	.1667	.1333	$m = 40$.4083	.3917	.3500
$m = 60$.2000	.1917	.1667	$m = 60$.4417	.4250	.3917
$m = 80$.2167	.2083	.1917	$m = 80$.4583	.4417	.4250
$m = 100$.2250	.2250	.2167	$m = 100$.4750	.4667	.4500
$m = 117$.2500	.2500	.2500	$m = 119$.5000	.5000	.5000

Table 1.1: **Breakdown point of subsampling and bootstrap quantiles.** t -quantile upper breakdown point of the bootstrap (Boots.) and the subsampling (Subs.) for different block sizes, sample sizes $n = 40, 80, 120$, and confidence levels $t = 0.9, 0.95, 0.99$ when the breakdown point (B. Point) is $b = 0.25, 0.5$. Bootstrap breakdown points are computed using Singh (1998) result. Subsampling breakdown points are computed using Theorem 2. The smallest integer such that the t -quantile subsampling breakdown point equals the breakdown point of statistic T for all given confidence levels is equal to 37, 77, and 117, 39, 79 and 119 for sample sizes 40, 80, and 120, $b=0.25, 0.5$, respectively.

$n = 40$	\mathcal{M}	$t = .9$	$t = .95$	$t = .99$
MCIV	[7; 12]	.3000	.2750	.2000
MCIV	[4; 18]	.3750	.3250	.2750
CM	[7; 12]	.4500	.4500	.4250
CM	[4; 18]	.5000	.5000	.5000
$n = 80$	\mathcal{M}	$t = .9$	$t = .95$	$t = .99$
MCIV	[9; 17]	.3250	.2875	.2250
MCIV	[5; 26]	.3750	.3500	.3000
CM	[9; 17]	.4500	.4250	.4000
CM	[5; 26]	.4750	.4625	.4500
$n = 120$	\mathcal{M}	$t = .9$	$t = .95$	$t = .99$
MCIV	[11; 21]	.3417	.3083	.2500
MCIV	[6; 32]	.3917	.3583	.3167
CM	[11; 21]	.4417	.4250	.3917
CM	[6; 32]	.4667	.4583	.4333

Table 1.2: **Breakdown point of Minimum Confidence Index Volatility (MCIV) and Calibration Method (CM)**. We consider a statistic with breakdown point $b = 0.5$ and confidence levels $t = 0.9, 0.95, 0.99$. The set \mathcal{M} of admissible block sizes is implied by the smallest and largest block size according to the suggested choice in Romano and Wolf (2001).

$n = 40, d = 3$	$t = .9$	$t = .95$	$t = .99$
R. Bootstrap	$\min(b, .5000)$	$\min(b, .5000)$	$\min(b, .5000)$
R. Subsampling			
$m = 6$	$\min(b, .3750)$	$\min(b, .3000)$	$\min(b, .2250)$
$m = 8$	$\min(b, .5000)$	$\min(b, .4500)$	$\min(b, .3500)$
$m = 10$	$\min(b, .5000)$	$\min(b, .5000)$	$\min(b, .4500)$
$m = 12$	$\min(b, .5000)$	$\min(b, .5000)$	$\min(b, .5000)$
$n = 80, d = 3$	$t = .9$	$t = .95$	$t = .99$
R. Bootstrap	$\min(b, .5000)$	$\min(b, .5000)$	$\min(b, .5000)$
R. Subsampling			
$m = 6$	$\min(b, .3500)$	$\min(b, .2875)$	$\min(b, .1875)$
$m = 8$	$\min(b, .4750)$	$\min(b, .4250)$	$\min(b, .3125)$
$m = 10$	$\min(b, .5000)$	$\min(b, .5000)$	$\min(b, .4125)$
$m = 12$	$\min(b, .5000)$	$\min(b, .5000)$	$\min(b, .5000)$
$n = 120, d = 3$	$t = .9$	$t = .95$	$t = .99$
R. Bootstrap	$\min(b, .5000)$	$\min(b, .5000)$	$\min(b, .5000)$
R. Subsampling			
$m = 8$	$\min(b, .4750)$	$\min(b, .4167)$	$\min(b, .3083)$
$m = 10$	$\min(b, .5000)$	$\min(b, .5000)$	$\min(b, .4083)$
$m = 12$	$\min(b, .5000)$	$\min(b, .5000)$	$\min(b, .4833)$
$m = 14$	$\min(b, .5000)$	$\min(b, .5000)$	$\min(b, .5000)$

Table 1.3: **Breakdown point of robust subsampling and robust bootstrap quantiles.** t -quantile breakdown point of the robust bootstrap and the robust subsampling in a linear regression model with $d = 3$, for different block sizes, sample sizes $n = 40, 80, 120$, and confidence levels $t = 0.9, 0.95, 0.99$. Robust bootstrap breakdown points are computed using the results in Salibian-Barrera and Zamar (2002). Robust subsampling breakdown points are computed using Theorem 14.

$n = 40$	$\gamma = 0\%$	$\gamma = 5\%$	$\gamma = 10\%$
Subs. $m = 5$.9595(3.5362)	.9870(27.9123)	.9900(54.0288)
Subs. $m = 16$.9700(2.8720)	.9890(14.7520)	.9905(32.0801)
Subs. MCIV	.9695(3.1921)	.9890(14.8329)	.9905(35.4644)
R. Subs. $m = 5$.9575(3.5285)	.9620(4.3960)	.9670(5.4761)
R. Subs. $m = 16$.9690(2.8569)	.9715(3.5609)	.9725(4.4743)
R. Subs. MCIV	.9690(3.1684)	.9710(3.9803)	.9725(5.0520)
$n = 80$	$\gamma = 0\%$	$\gamma = 5\%$	$\gamma = 10\%$
Subs. $m = 7$.9620(3.6404)	.9885(33.1606)	.9900(57.1361)
Subs. $m = 22$.9755(3.1927)	.9915(20.5762)	.9925(40.8722)
Subs. MCIV	.9720(3.4242)	.9915(23.2245)	.9920(44.1523)
R. Subs. $m = 7$.9595(3.5924)	.9615(4.4963)	.9620(5.6244)
R. Subs. $m = 22$.9700(3.1922)	.9720(3.9843)	.9730(5.0482)
R. Subs. MCIV	.9700(3.3689)	.9720(4.3567)	.9720(5.3602)
$n = 120$	$\gamma = 0\%$	$\gamma = 5\%$	$\gamma = 10\%$
Subs. $m = 8$.9580(3.6849)	.9865(36.3722)	.9900(59.0049)
Subs. $m = 27$.9730(3.3121)	.9885(23.7312)	.9920(44.4842)
Subs. MCIV	.9715(3.5285)	.9880(25.5487)	.9900(48.4229)
R. Subs. $m = 8$.9575(3.6729)	.9580(4.6322)	.9590(5.7238)
R. Subs. $m = 27$.9670(3.3121)	.9695(4.2125)	.9700(5.1608)
R. Subs. MCIV	.9665(3.5040)	.9695(4.6120)	.9695(5.5682)

Table 1.4: **Empirical coverage of the subsampling and the robust subsampling** (Subs. and R. Subs., respectively). Simulated empirical coverage in the square of the sample average testing setting, for confidence levels $t = 0.95$, for the sample sizes $n = 40, 80, 120$ and contamination probabilities $\gamma = 0\%, 5\%, 10\%$. We denote the block size by m . Data driven block size selection procedure based on the MCIV index is denoted by “MCIV”. In brackets, we give the median .95-quantile of the robust and nonrobust subsampling distributions, respectively. The number of replications is 2000.

$n = 40, d = 3$	$\gamma = 0\%$	$\gamma = 15\%$	$\gamma = 25\%$
Subs. $m = 9$.9610(0.6837)	.9850(1.4588)	.9860(1.7875)
Subs. $m = 16$.9240(0.5950)	.9470(1.2326)	.9500(1.5326)
Subs. MCIV	.9370(0.6215)	.9600(1.2774)	.9645(1.6039)
R. Subs. $m = 9$.9665(0.7767)	.9750(0.9325)	.9790(1.4024)
R. Subs. $m = 16$.9250(0.6792)	.9330(0.7601)	.9410(1.1120)
R. Subs. MCIV	.9415(0.7091)	.9465(0.7959)	.9540(1.1592)
$n = 40, d = 5$	$\gamma = 0\%$	$\gamma = 15\%$	$\gamma = 25\%$
Subs. $m = 13$.9720(0.6976)	.9805(1.4556)	.9830(1.7979)
Subs. $m = 18$.9405(0.6292)	.9515(1.2920)	.9535(1.6035)
Subs. MCIV	.9465(0.6359)	.9665(1.3177)	.9680(1.6515)
R. Subs. $m = 13$.9730(0.8012)	.9750(0.9607)	.9770(1.5002)
R. Subs. $m = 18$.9430(0.7023)	.9480(0.8360)	.9505(1.3181)
R. Subs. MCIV	.9495(0.7351)	.9565(0.8770)	.9600(1.3570)
$n = 60, d = 3$	$\gamma = 0\%$	$\gamma = 15\%$	$\gamma = 25\%$
Subs. $m = 9$.9705(0.5728)	.9855(1.2209)	.9880(1.5005)
Subs. $m = 19$.9360(0.4919)	.9565(1.0349)	.9580(1.2756)
Subs. MCIV	.9470(0.5148)	.9695(1.0771)	.9735(1.3353)
R. Subs. $m = 9$.9720(0.6499)	.9785(0.7444)	.9800(1.0961)
R. Subs. $m = 19$.9390(0.5489)	.9450(0.5998)	.9505(0.8580)
R. Subs. MCIV	.9485(0.5720)	.9545(0.6248)	.9605(0.8838)
$n = 60, d = 5$	$\gamma = 0\%$	$\gamma = 15\%$	$\gamma = 25\%$
Subs. $m = 13$.9635(0.5717)	.9850(1.2244)	.9865(1.5075)
Subs. $m = 23$.9275(0.4979)	.9490(1.0498)	.9550(1.2967)
Subs. MCIV	.9375(0.5127)	.9690(1.0889)	.9705(1.3510)
R. Subs. $m = 13$.9660(0.6602)	.9750(0.7591)	.9800(1.1549)
R. Subs. $m = 23$.9375(0.5730)	.9420(0.6276)	.9490(0.9152)
R. Subs. MCIV	.9430(0.5966)	.9560(0.6598)	.9605(0.9521)

Table 1.5: **Empirical coverage of the subsampling and the robust subsampling** (Subs. and R. Subs., respectively). Simulated empirical coverage in the linear regression testing setting with $d = 3, 5$, for confidence levels $t = 0.95$, for the sample sizes $n = 40, 60$ and contamination probabilities $\gamma = 0\%, 15\%, 25\%$. We denote the block size by m . Data driven block size selection procedure based on the MCIV index is denoted by “MCIV”. In brackets the median of the confidence interval lengths. The number of replications is 2000.

Chapter 2

Robust Resampling Methods for Time Series

2.1 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 simulation 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.

2.2 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 (2009a). 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 (2009a), 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 (2009a). 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, which makes it applicable to non linear 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. Finally, we investigate by Monte Carlo simulation the performance of our robust resampling approach in the benchmark context of the estimation of the autoregressive parameter in an AR(1) model. 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.

The paper is organized as follows. Section 2.3 outlines the main setting and introduces the quantile breakdown point formulas of different block resampling procedures. In Section 2.4 we develop our robust approach and derive the relevant expression for its associated quantile breakdown point formula. We show that, under weak conditions, the resulting quantile breakdown point is maximal. In Section 2.5, 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 and some sensitivity analysis are presented in Section 2.6. Section 2.7 concludes.

2.3 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.

2.3.1 Definition

Let $X_{(n)} = (X_1, \dots, X_n)$ be a sample from a real valued stationary process $\mathcal{X} = \{X_t, t \in \mathbb{Z}\}$ defined on the probability space (Ω, \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 $X_{(k)}^* = (X_1^*, \dots, X_k^*)$ 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:

$$Q_t^* = \inf\{x | P^*(T_{n,k}^* \leq x) \geq t\}, \quad (2.1)$$

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

We characterize the robustness of quantile (2.1) via its breakdown point, i.e., the smallest fraction of outliers in the original sample such that Q_t^* degenerates, making inference based on (2.1) meaningless. Different than in the iid case, in time series we can consider different possible models of contamination by outliers: 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}_p^ζ 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 ζ_1, \dots, ζ_p to be all equal to ζ , but we rather assume existence of constants c_1, \dots, 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 $T(X_{(n)} + Z_{nb}^\zeta) = +\infty$ for some $Z_{nb}^\zeta \in \mathcal{Z}_{nb}^\zeta$. 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 simulation 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 15 *The upper breakdown point of the t -quantile Q_t^* is given by:*

$$b_t = \frac{1}{n} \cdot \left[\inf_{\{1 \leq p \leq \lceil n/2 \rceil\}} \{p \mid \text{there exists } Z_p^\zeta \in \mathcal{Z}_p^\zeta \text{ such that } Q_t^*(X_{(n)} + Z_p^\zeta) = +\infty\} \right], \quad (2.2)$$

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

2.3.2 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 them. For brevity, we denote by b_t^K , $K = OS, NB, OB$, 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).

Subsampling

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

Theorem 16 *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:*

$$b_t^{OS} \leq \inf_{\{p \in \mathbb{N}, p \leq r-1\}} \left\{ p \cdot \frac{\lceil mb \rceil}{n} \mid p > \frac{(1-t)(n-m+1) + \lceil mb \rceil - 1}{m} \right\}. \quad (2.3)$$

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 3.1 confirms this basic intuition. Using Theorem 16 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 imply 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 (2009a). 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.

Moving Block Bootstrap

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

overlapping moving block bootstrap selects randomly with replacement r overlapping blocks $X_{(m),i}^{O*}$, $i = 1, \dots, r$. Then, it applies statistic T to the n -sample $X_{(n)}^{O*} = (X_{(m),1}^{O*}, \dots, X_{(m),r}^{O*})$.

Theorem 17 *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:*

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

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

Similar to the findings for the subsampling, the right part of (i) and (ii) are similar for large $n \gg m$. 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 17, 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 3.1, 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.

2.4 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 (2009a).

2.4.1 Definition

Given the original sample $X_{(n)} = (X_1, \dots, X_n)$, 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:

$$\psi_n(X_{(n)}, \hat{\theta}_n) := \frac{1}{n - q + 1} \sum_{i=q}^n g(X_{i-q+1}, \dots, X_i; \hat{\theta}_n) = 0, \quad (2.4)$$

where $\psi_n(X_{(n)}, \cdot) : \mathbb{R}^d \rightarrow \mathbb{R}^d$ 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 imply 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 (2.4) around the true parameter θ_0 :

$$\hat{\theta}_n - \theta_0 = -[\nabla_{\theta} \psi_n(X_{(n)}, \theta_0)]^{-1} \psi_n(X_{(n)}, \theta_0) + o_p(1), \quad (2.5)$$

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 $-\nabla_{\theta} \psi_n(X_{(n)}, \hat{\theta}_n)^{-1} \psi_n(X_{(n)}, \hat{\theta}_n)$ as an approximation of $\hat{\theta}_n - \theta_0$ 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:

$$L_{n,m}^{RF*}(x) = \frac{1}{N} \sum_{s=1}^N \mathbb{I}(\tau_k(-[\nabla_{\theta}\psi_n(X_{(n)}, \hat{\theta}_n)]^{-1}\psi_k(X_{(k),s}^*, \hat{\theta}_n)) \leq x), \quad (2.6)$$

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 (2.6) 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).

2.4.2 Robust Resampling Methods and Quantile Breakdown Point

In the computation of (2.6) 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 18 *Let b be the breakdown point of the robust M -estimator $\hat{\theta}_n$ defined by (2.4). The t -quantile breakdown point of resampling distribution (2.6) is given by $b_t = \min(b, b_{\nabla\psi})$, where:*

$$b_{\nabla\psi} = \frac{1}{n} \cdot \inf_{1 \leq p \leq \lceil n/2 \rceil} \{p \mid \text{there exists } Z_p^{\zeta} \in \mathcal{Z}_p^{\zeta} \text{ such that } \det(\nabla_{\theta}\psi_n(X_{(n)} + Z_p^{\zeta}, \hat{\theta}_n)) = 0\}. \quad (2.7)$$

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 (2.6) is maximal, independent of confidence level t .

2.5 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:

$$b_t^u := \frac{1}{n} \cdot \inf_{1 \leq p \leq \lceil n/2 \rceil} \{p \mid \text{there exists } Z_p^\zeta \in \mathcal{Z}_p^\zeta \text{ such that } m_u(X_{(n)} + Z_p^\zeta) = \infty\}. \quad (2.8)$$

2.5.1 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.

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 19 *Let $m_{\min} < m_{\max}$ and $k \in \mathbb{N}$ be fixed. For $m \in \{m_{\min} - k, \dots, m_{\max} + k\}$, 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 $m \in \{m_{\min}, \dots, m_{\max}\}$ by*

$$CIV(m) := \frac{1}{2k+1} \sum_{i=-k}^{i=k} \left(Q_t^*(m+i) - \overline{Q}_t^{*k}(m) \right)^2. \quad (2.9)$$

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

$$m_{MCIV} = \arg \inf_{m \in \mathcal{M}} \{CIV(m) : CIV(m) \in \mathbb{R}^+\}, \quad (2.10)$$

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 16, the formula for the breakdown point of m_{MCIV} is given in the next corollary.

Corollary 20 *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 16, as a function of the block*

size $m \in \mathcal{M}$. It then follows:

$$b_t^{OS, MCIV} = \sup_{m \in \mathcal{M}} \inf_{j \in \{-k, \dots, k\}} b_t^{OS}(m + j). \quad (2.11)$$

The dependence of the breakdown point formula for the MCIV on the breakdown point of subsampling quantiles is identical in 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.

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). We present this method for the case of one-sided confidence intervals only. The modifications for two-sided intervals are straightforward.

Definition 21 Fix $t \in (0, 1)$ and let (X_1^*, \dots, 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

$$m_{CM} := \arg \inf_{m \in \mathcal{M}} \{ |t - P^* [\hat{\theta}_n \leq Q_t^{**}(m)]| : P^* [Q_t^{**}(m) \in \mathbb{R}] > 1 - t \}, \quad (2.12)$$

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 21, 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 22 *Let b be the breakdown point of estimator $\hat{\theta}_n$, $t \in (0, 1)$, and define:*

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

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

$$b_t^{OS,CM} \leq \sup_{m \in \mathcal{M}} \{b_t^{OS^{**}}(m)\}. \quad (2.13)$$

Because of the use of the moving block bootstrap instead of the standard iid bootstrap in the CM for time series, equation (2.13) is quite different from the formula for the iid case in Camponovo, Scaillet and Trojani (2009a). Similar to the iid case, the theoretical results in Table 2.2 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.

2.5.2 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 23 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 $l \in \{l_{\min}, \dots, l_{\max}\}$ 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, \dots, X_{i+m-1}) , $1 \leq i \leq n - m + 1$. $\bar{Q}_t^*(m, l) := \frac{1}{n - m + 1} \sum_{i=1}^{n - m + 1} Q_t^*(m, l, i)$ 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 $l \in \{l_{\min}, \dots, l_{\max}\}$ define the MSE index is defined as:

$$MSE(l) := \left(\bar{Q}_t^*(m, l) - Q_t^*(n, l') \right)^2 + \frac{1}{n - m + 1} \sum_{i=1}^{n - m + 1} (Q_t^*(m, l, i) - Q_t^*(n, l'))^2, \quad (2.14)$$

and set:

$$l_{HHJ} = \arg \inf_{l \in \{l_{\min}, \dots, l_{\max}\}} \{MSE(l) : MSE(l) \in \mathbb{R}^+\}, \quad (2.15)$$

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

$$m_{HHJ} = l_{HHJ} \left(\frac{n}{m} \right)^{1/5}. \quad (2.16)$$

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 24 *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 16, as a function of the block size $l \in \{l_{\min}, \dots, l_{\max}\}$ and a size m of the initial sample. It then follows for $K = NS, OS$:*

$$b_t^{K,MCIV} = \frac{m}{n} \cdot \sup_{l \in \{l_{\min}, \dots, l_{\max}\}} b_t^{K,m}(l). \quad (2.17)$$

The computation of the optimal block size l_{HHJ} based on smaller subsamples of size $l \ll m < n$, causes a large instability in the computation of m_{HHJ} . Because of this effect, the MSE index in (2.14) 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 2.2 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.

2.6 Monte Carlo Simulations

We compare through Monte Carlo simulation 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) model of the form:

$$X_t = \theta X_{t-1} + \epsilon_t, \quad X_0 \sim N\left(0, \frac{1}{1 - \theta^2}\right), \quad (2.18)$$

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:

$$\psi_n^{OLS}(X_{(n)}, \hat{\theta}_n^{OLS}) := \frac{1}{n-1} \sum_{t=2}^n X_{t-1}(X_t - \hat{\theta}_n^{OLS} X_{t-1}) = 0. \quad (2.19)$$

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

$$\psi_n^{ROB}(X_{(n)}, \hat{\theta}_n^{ROB}) := \frac{1}{n-1} \sum_{t=2}^n h_c(X_{t-1}(X_t - \hat{\theta}_n^{ROB} X_{t-1})) = 0, \quad (2.20)$$

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 replacement outliers random samples $(\tilde{X}_1, \dots, \tilde{X}_n)$ generated according to:

$$\tilde{X}_t = (1 - p_t)X_t + p_t \cdot X_{1.5max}, \quad (2.21)$$

where $X_{1.5max} = 1.5 \cdot \max(X_1, \dots, X_n)$ and p_t is an iid 0–1 random sequence, independent of process (2.18) 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.

2.6.1 The Standard Strictly Stationary Case

We construct symmetric resampling confidence intervals for 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 (2009a, 2010a) and Mikusheva (2007) also show that because of a lack of uniformity in pointwise asymptotics, non symmetric subsampling confidence

intervals for autoregressive models can imply a distorted asymptotic size, which is instead correct for symmetric confidence intervals.

Using OLS estimator (2.19), 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 (2.20), 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 (2.20) are computationally too expensive.

We generate $N=1000$ samples of size $n = 180$ according to model (2.18) 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 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 3.1 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 simulation (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:

$$X_{\max} = \arg \max_{X_1, \dots, X_n} \{u(X_i) | u(X_i) = X_i - \theta X_i, \text{ under } \mathcal{H}_0\}, \quad (2.22)$$

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 2.1, we plot the resulting empirical p -values. As expected, our robust fast approach shows a desirable stability for both subsampling and bootstrap methods.

2.6.2 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 (2009a, 2010b) and Mikusheva (2007). Therefore, we focus exclusively on the latter method.

Consider the OLS estimator (2.19) 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 (2.18) for the parameter choices $\theta_0 = 0.8, 0.85, 0.9, 0.95$, and simulate contaminated samples $(\tilde{X}_1, \dots, \tilde{X}_n)$ according to (3.8) 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 $\mathcal{M} = \{9, 10, 12, 15, 18\}$.

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 2.3 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 simulation (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 non-monotonic, 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 $\theta_0 = 0.95$.

2.7 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 fast robust resampling approach, which is applicable to a wide class of block resampling methods, and we show that it 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.2 Proofs

Proof of Theorem 16. Denote by $X_{(m),i}^N = (X_{(i-1)m+1}, \dots, X_{im})$, $i = 1, \dots, r$ and $X_{(m),i}^O = (X_i, \dots, X_{i+m-1})$, $i = 1, \dots, n-m+1$ the nonoverlapping and overlapping blocks of size m respectively. Given the original sample $X_{(n)}$, for the first nonoverlapping block $X_{(m),1}^N$, consider following type of contamination:

$$X_{(m),1}^N = (X_1, \dots, X_{m-\lceil mb \rceil}, Z_{m-\lceil mb \rceil+1}, \dots, Z_m), \quad (2.23)$$

where X_i , $i = 1, \dots, m - \lceil mb \rceil$ and Z_j , $j = m - \lceil mb \rceil + 1, \dots, m$, denote the non contaminated and contaminated points respectively. By construction, the first $m - \lceil mb \rceil + 1$ overlapping blocks $X_{(m),i}^O$, $i = 1, \dots, m - \lceil mb \rceil + 1$, contain $\lceil mb \rceil$ outliers. Consequently, $T(X_{(m),i}^O) = +\infty$, $i = 1, \dots, m - \lceil mb \rceil + 1$. Assume that the first $p < r - 1$ nonoverlapping blocks $X_{(m),i}^N$, $i = 1, \dots, p$ have the same contamination as in (3.8). 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,

$$b_t^{OS} \leq \inf_{\{p \in \mathbb{N}, p \leq r-1\}} \left\{ p \cdot \frac{\lceil mb \rceil}{n} \left| \frac{mp - \lceil mb \rceil + 1}{n - m + 1} > 1 - t \right. \right\} \quad \blacksquare$$

Proof of Theorem 17. *Case (i): Nonoverlapping Moving Block Bootstrap.* Consider $X_{(m),i}^N$, $i = 1, \dots, r$. 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 $X \sim \text{BIN}(r, \frac{p_2}{r})$.

By Definition 15, $Q_t^* = +\infty$ 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,

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

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

By Definition 15, $Q_t^* = +\infty$ 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,

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

Proof of Corollary 18. Consider the robust fast approximation of $(\hat{\theta}_k^* - \hat{\theta}_n)$ given by:

$$-[\nabla_{\theta} \psi_n(X_{(n)}, \hat{\theta}_n)]^{-1} \psi_k(X_{(k),s}^*, \hat{\theta}_n), \quad (2.24)$$

where $k = n$ or $k = m$. Assuming a bounded estimating function, expression (2.24) 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 $b_t = \min(b, b_{\nabla\psi})$. \blacksquare

Proof of Corollary 20. 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 $\{X_1, \dots, X_n\}$ satisfies $p \geq \min\{b_t^{OS}(m-k), b_t^{OS}(m-k+1), \dots, b_t^{OS}(m+k-1), b_t^{OS}(m+k)\}$. This

concludes the proof. ■

Proof of Corollary 22. By definition, in order to get $m_{CM} = \infty$ we must have $P[Q_t^{**}(m) = \infty] \geq t$ 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 (3.8). Let X be the random variable which denotes the number of contaminated blocks in the nonoverlapping moving block bootstrap sample. As in (i), $X \sim \text{BIN}(r, q/r)$. 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 24. By definition, in order to get $m_{HHJ} = \infty$ we must have $l_{HHJ} = \infty$, i.e. $MSE(l) = \infty$, for all $l \in \{l_{min} \dots, l_{max}\}$. For l fixed, $MSE(l) = \infty$ if just a single $Q_t^*(m, l, i)$, $i = 1, \dots, n - m + 1$ diverges to infinity. This concludes the proof. ■

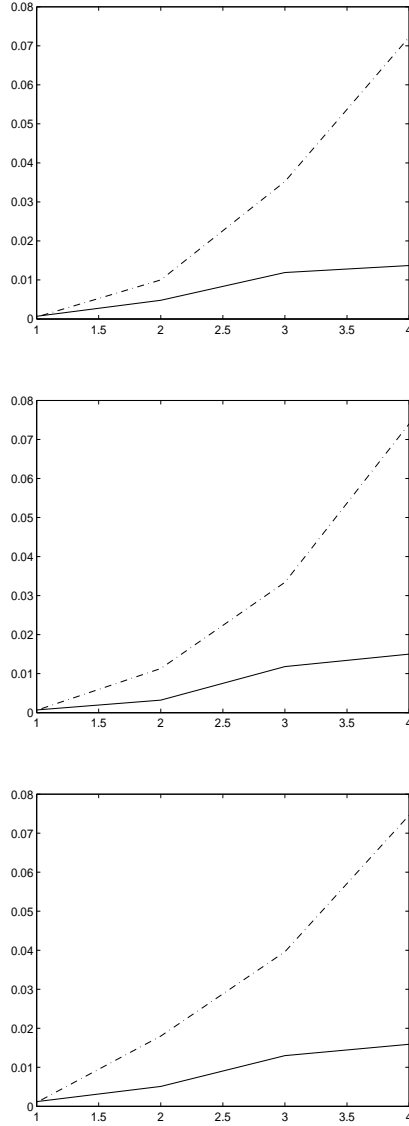


Figure 2.1: **Sensitivity analysis.** Sensitivity plots of the variation of the empirical p -value average, for a test of the null hypothesis $\mathcal{H}_0 : \theta_0 = 0.5$, with respect to variations of X_{\max} , in each Monte Carlo sample, within the interval $[1, 4]$. The random samples were generated 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).

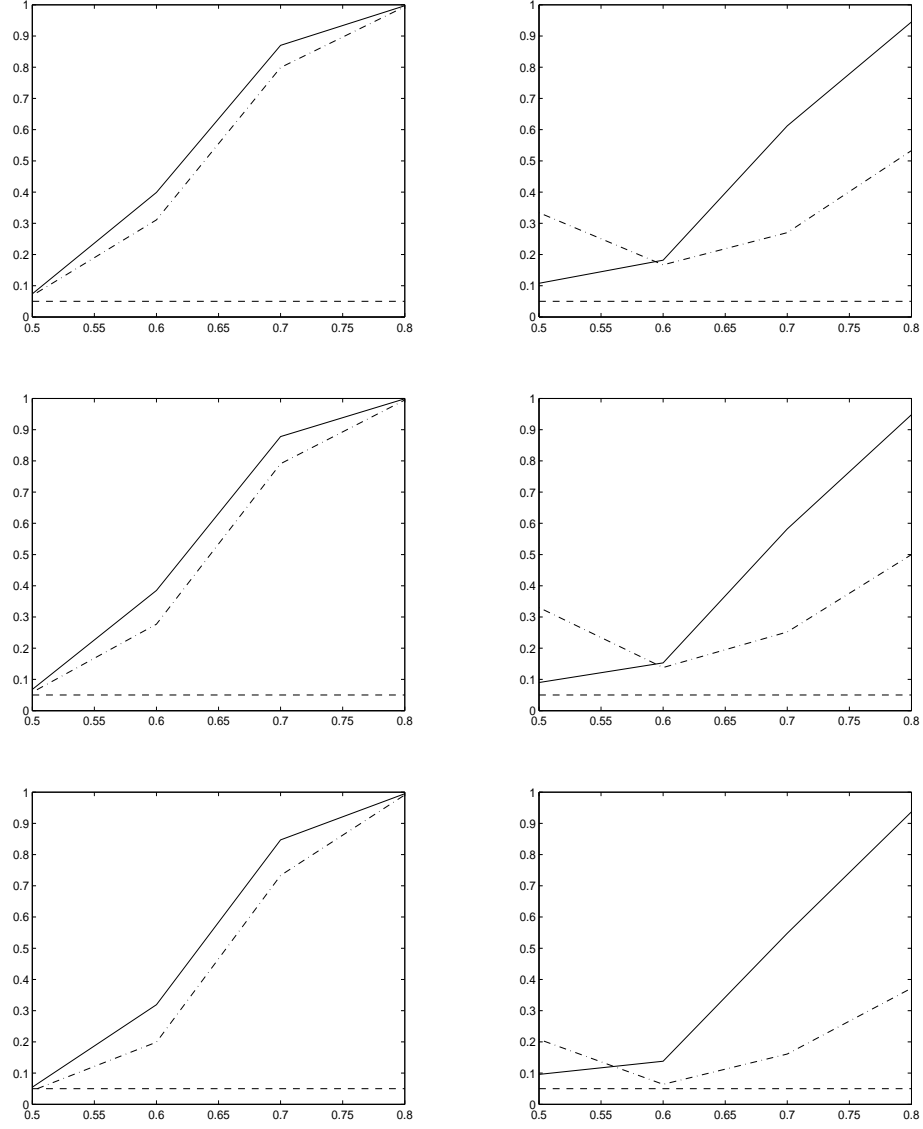


Figure 2.2: **Power curves in the standard strictly stationary case.** We plot the proportion of rejections of the null hypothesis $\mathcal{H}_0 : \theta_0 = 0.5$, when the true parameter value is $\theta_0 \in [0.5, 0.8]$. From the top to the bottom, we present the overlapping subsampling with MCIV, 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 $\eta = 1.5\%$.

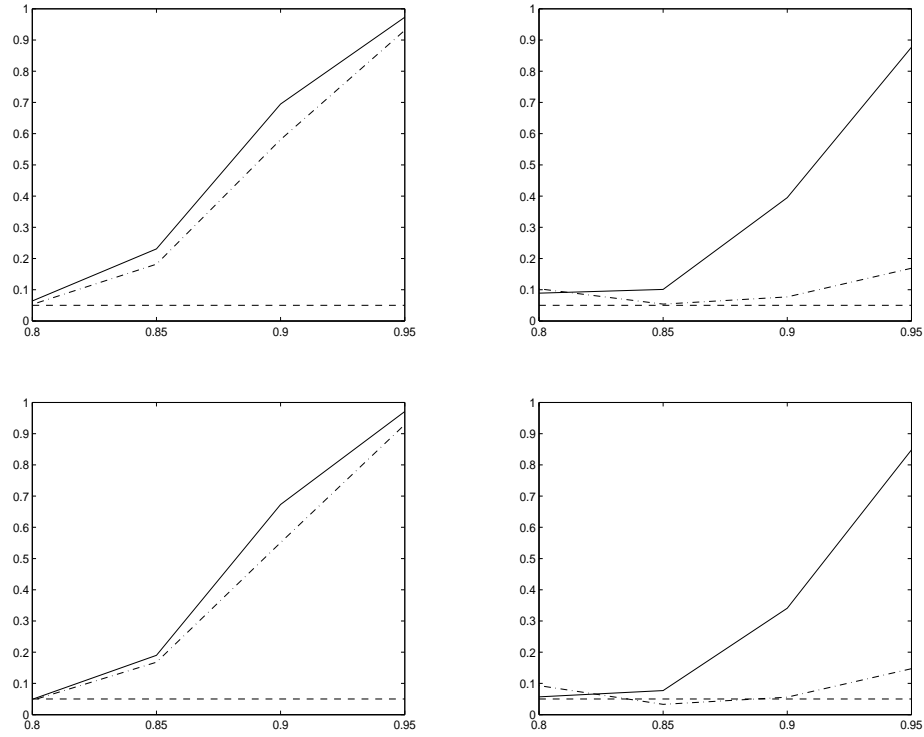


Figure 2.3: **Power curves in the near-to-unit-root case.** We plot the proportion of rejections of the null hypothesis $\mathcal{H}_0 : \theta_0 = 0.8$, when the true parameter value is $\theta_0 \in [0.8, 0.95]$. From the top to the bottom, we present the overlapping subsampling with MCIV and CM. We 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 2.1: **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 16 and 17.

$n = 120$	$t = 0.95$	$t = 0.99$
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 2.2: **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 20, 22 with $\mathcal{M} = \{6, 8, 10, 12, 15\}$. For (HHJ) we use Corollary 24 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 $n = 120$.

Chapter 3

Robust Predictive Regression and Stock Returns Predictability

3.1 Abstract

A large literature studies the predictability of stock returns by other lagged financial variables in a predictive regression setting. A common feature of widely used testing procedures is a failing statistical robustness, which may lead to misleading conclusions determined by the particular features of a small subfraction of the data. We propose a new general method to deal with this problem based on the robust subsampling approach. The method implies robust confidence intervals and inference results. It is applicable both in the multi-predictor context and in settings with nearly integrated regressors. Simulation evidence confirms the higher accuracy and efficiency of our robust testing approach for typical applications in which the data may follow only approximately the predictive regression model. We apply our approach to US equity data from 1961 to 2008 and find that it yields a stronger evidence in favor of predictability than a number of other (nonrobust) tests in the literature.

3.2 Introduction

A number of studies has investigated whether stock returns can be predicted by economic variables such as, e.g., the price-dividend ratio or the interest rate; see, for instance, Rozeff (1984), Fama and French (1988), Campbell and Shiller (1988), Nelson and Kim (1993), Goetzmann and Jorion (1995), Campbell and Yogo (2006), Jansson and Moreira (2006), Polk, Thompson and Vuolteenaho (2006). The econometric approach to test for predictability is mostly based on a predictive regression of stock returns onto a set of lagged financial variables; see, e.g., Stambaugh (1999). Important distinctions between testing procedures in the literature arise because of the different test statistics and asymptotic theories used to test the null hypothesis of no predictability. These differences lead in a number of cases to diverging results and conclusions.

Mankiw and Shapiro (1986) and Stambaugh (1986) note that in a setting with endogenous predictor and correlated innovations standard asymptotic theory causes small sample biases that may imply an over rejection of the hypothesis of no predictability. To mitigate the problem, a first approach proposes tests based on bias-corrected estimators of predictive regressions. For instance, Stambaugh (1999) and Amihud, Clifford and Wand (2008) introduce bias-corrected OLS estimators for the univariate and the multi-predictor setting, respectively. Another strand of early literature applies a VAR simulation approach or a bootstrap scheme to compute finite sample confidence intervals; see, among others, Nelson and Kim (1993) and Goetzmann and Jorion (1993).¹

Recent work has considered the issue of endogenous integrated or nearly integrated predictors, following the evidence in Torous, Valkanov and Yan (2004) that various variables assumed to predict stock returns follow a local-to-unit root autoregressive process. Lewellen (2004), Torous et al. (2004) and Campbell and Yogo (2006) introduce new testing procedures and more accurate unit-root and

¹These methods deliver consistent predictability tests under particular assumptions that may be hardly verifiable in applications. Standard asymptotic consistency proofs for the VAR approach assume knowledge of the true data generating model. Consistency of the bootstrap approach in Goetzmann and Jorion (1993) follows if returns are iid and dividends are not stochastic.

local-to-unit root asymptotics for the predictive regression model with a single persistent predictor and correlated innovations. Torous et al. (2004) and Campbell and Yogo (2006) tests are based on Bonferroni methods and are thus difficult to extend to multi-predictor predictive regression settings. Following Politis and Romano (1994a) and Wolf (2000), Choi and Chue (2007) show that the subsampling is a convenient resampling scheme to define valid confidence intervals and tests for predictive regression models with multiple, potentially nearly integrated, regressors. In contrast to the bootstrap, the subsampling satisfies weaker consistency conditions that are typically satisfied by many empirical predictive regression models.

A common denominator of all above testing approaches is their dependence on procedures that can be heavily influenced by a small fraction of particular observations in the data. For standard OLS estimators of linear regression models with iid data this problem is well-known since a long time; see Huber (1981) for a review. Following Huber’s seminal work, several authors have emphasized the potentially even worse robustness features of level and power of asymptotic tests of parametric hypotheses in time series settings; see Heritier and Ronchetti (1994), Ronchetti and Trojani (2001), Mancini, Ronchetti and Trojani (2005) and Gagliardini, Trojani and Urga (2005), among others.² Finally, recent research has shown that confidence intervals and tests of resampling schemes like the bootstrap and the subsampling may be easily inflated by a small fraction of outliers in the data; see Singh (1998), Salibian-Barrera and Zamar (2002) and Camponovo, Scaillet and Trojani (2009a,b). Intuitively, this problem arises because the fraction of outliers generated by resampling schemes like the bootstrap and the subsampling is often much higher than the fraction of outliers in the original data. This feature may imply unreliable test results even when estimating the predictive regression model with robust regression techniques.

²For instance, t -type statistics and confidence intervals for testing the null of no predictability require estimation of both the slope parameter in the predictive regression model and the variance of the error term. Since standard variance estimators can be very sensitive to outliers or other particular structures of a small subset of the data the resulting test may imply unreliable level and power properties.

The main contributions of this paper are (i) a new general robust method for hypothesis testing in the predictive regression setting and (ii) a robust analysis on the conclusions of a number of predictability tests in the empirical literature. Our method is applicable both in multivariate and near-to-unity regression settings, it ensures robustness of level and power of the implied tests, and it allows the researcher to identify from the data potentially problematic observations that may invalidate the inference derived from nonrobust test procedures in the literature. We build our tests by using well-known robust M-estimators for linear regression models available in the statistical literature, and by computing test critical values and confidence intervals with the robust subsampling approach in Camponovo, Scaillet and Trojani (2009b). While the subsampling allows us to consider both multivariate and near-to-unit-root settings, using the robust subsampling allows us to avoid the potentially damaging effects of a few particular data points in the data on the test conclusions. We apply our robust approach to US equity data from 1961 to 2008 and find that it yields a stronger evidence in favor of predictability than a variety of other tests in the literature. This evidence suggests that conventional approaches in the literature may feature a low power in applications, due to the presence of abnormal structures in a small subset of the data.

The paper is organized as follows. Section 3.3 introduces the relevant multivariate predictive regression model and some recent procedures proposed for testing the null hypothesis of no predictability. Section 3.4 explains how robust predictability tests are obtained using the robust subsampling approach in Camponovo, Scaillet and Trojani (2009b). Moreover, it produces Monte Carlo evidence on the accuracy and robustness properties of our tests relative to classical methods. In Section 3.5, we apply our robust testing approach to US equity data and examine the empirical evidence for predictability. Section 3.6 concludes.

3.3 Predictive Regression Model

We denote by $\{y_t\}$ the response variables and by $\{x_t\}$ the explanatory variables of the predictive regressions model:

$$y_t = \alpha + \beta' x_{t-1} + u_t \quad (3.1)$$

$$x_t = \mu + \rho x_{t-1} + v_t \quad (3.2)$$

where, for $t = 1, \dots, n$, y_t , α and u_t are scalars, x_t , μ and v_t are p -dimensional vectors and ρ is a $p \times p$ matrix with eigenvalues $|\lambda_i| < 1$, $i = 1, \dots, p$. We assume that the errors $\{v_t\}$ are iid with mean zero and covariance matrix Σ , while $u_t = \phi' v_t + e_t$, where ϕ is a p -dimensional vector and $\{e_t\}$ are iid random variables with mean zero and independent of $\{x_t\}$ and $\{v_t\}$. We use the subscript 0 to indicate the true values β_0 and ρ_0 of the parameters β and ρ .

It is well known that in this setting inference based on standard asymptotic theory suffers from small sample biases, which may imply an over rejection of the hypothesis of no predictability, $\mathcal{H}_0 : \beta_0 = 0$; see e.g. Mankiw and Shapiro (1986) and Stambaugh (1986). Moreover, as emphasized in Torous et al. (2004), various state variables considered as predictors are well approximated by a nearly integrated process. Consequently, this suggests a local-to-unit framework $\lambda_i = 1 + c/n$ for the autoregressive model (3.2), which may imply a nonstandard asymptotic distribution for the classical estimators of ρ and β .

Recently, several tests procedures have been proposed in order to overcome these problems. In this paper, we focus on the more recent approaches based on the Bonferroni method proposed in Torous et al. (2004) and Campbell and Yogo (2006), the bias-corrected estimators method used in Amihud et al. (2008) and the subsampling procedures proposed in Wolf (2000) and Choi and Chue (2007). We first study the accuracy of these procedures under small violations of the model assumptions.

Indeed, all these approaches are based on statistical tools which are very sensitive to contamination by outliers or more general deviations from the assumed predictive regression model. Consequently, these testing procedures may become easily inefficient or biased even with a small fraction of outliers in the data. We first briefly introduce the different approaches and finally we analyze through Monte Carlo simulations their accuracy, both in presence and absence of small deviations from the model assumptions.

3.3.1 Bonferroni Method

Consider the predictive regressions model (3.1)-(3.2) with $p = 1$ and $\rho = \lambda_1 = 1 + c/n$. In order to test the null Hypothesis $\mathcal{H}_0 : \beta_0 = b$, Campbell and Yogo (2006) introduced the Q -statistic:

$$Q(b, \rho) = \frac{(\hat{\beta}_n - b) - \beta_{uv}(\hat{\rho}_n - \rho)}{\sigma_u(1 - \delta^2)^{1/2}(\sum_{t=1}^n x_{t-1}^{\mu^2})^{-1/2}} \quad (3.3)$$

where $\sigma_u^2 = \text{Var}(u_t)$, $\sigma_v^2 = \text{Var}(v_t)$, $\sigma_{uv} = \text{Cov}(u_t, v_t)$, $\beta_{uv} = \sigma_{uv}/\sigma_v^2$, $\delta = \sigma_{uv}/\sigma_u\sigma_v$, $x_{t-1}^\mu = x_{t-1} - (n^{-1} \sum_{t=1}^n x_{t-1})$ and $\hat{\beta}_n$ and $\hat{\rho}_n$ are the OLS estimators of β and ρ , respectively.³

The Q -statistic depends on the unknown parameter ρ . Consequently, tests based on (3.3) become unfeasible. To solve this problem, Campbell and Yogo (2006) propose a Bonferroni approach. More precisely, for fixed $\alpha_1 = \bar{\alpha}_1 + \underline{\alpha}_1$, they first compute $\bar{\rho}(\bar{\alpha}_1)$ and $\underline{\rho}(\underline{\alpha}_1)$ such that $\bar{\alpha}_1 = P(\rho > \bar{\rho}(\bar{\alpha}_1))$ and $\underline{\alpha}_1 = P(\rho < \underline{\rho}(\underline{\alpha}_1))$. Then, for fixed α_2 with $\alpha = \alpha_1 + \alpha_2$, the Bonferroni confidence interval for β is given by:

$$C_\beta(\alpha) = [\underline{\beta}(\bar{\rho}(\bar{\alpha}_1), \alpha_2), \bar{\beta}(\underline{\rho}(\underline{\alpha}_1), \alpha_2)] \quad (3.4)$$

where $\underline{\beta}(\rho, \alpha_2) = \beta(\rho) - z_{\frac{\alpha_2}{2}} \sigma_u \left(\frac{1 - \delta^2}{\sum_{t=1}^n x_{t-1}^{\mu^2}} \right)^{1/2}$, $\bar{\beta}(\rho, \alpha_2) = \beta(\rho) + z_{\frac{\alpha_2}{2}} \sigma_u \left(\frac{1 - \delta^2}{\sum_{t=1}^n x_{t-1}^{\mu^2}} \right)^{1/2}$, $\beta(\rho) = \frac{\sum_{t=1}^n x_{t-1}^\mu [y_t - \beta_{uv}(x_t - \rho x_{t-1})]}{\sum_{t=1}^n x_{t-1}^{\mu^2}}$, and $z_{\frac{\alpha_2}{2}}$ denotes the $1 - \alpha_2/2$ quantile of the standard normal distribu-

³Campbell and Yogo (2006) introduce a t -statistic as well. However, because of the better accuracy of the Q -statistic, showed through theoretical results and Monte Carlo simulations, we focus on the Q -statistic.

tion. By Bonferroni's inequality, the confidence interval (3.4) has coverage of at least $100(1 - \alpha)\%$. However, confidence interval (3.4) is typically conservative. Campbell and Yogo (2005, 2006) provide a refinement of the Bonferroni method and a procedure to implement it. In our study, we consider this last approach.

3.3.2 Bias Correction Methods

Consider the predictive regressions model (3.1)-(3.2) with $p \geq 1$. For testing the hypothesis of predictability, Amihud et al. (2008) consider the augmented regressions:

$$y_t = \alpha + \beta' x_{t-1} + \phi' v_t + u_t. \quad (3.5)$$

Since $\{v_t\}$ is unknown, it has to be replaced by a sample estimate. A simple candidate is $\{\hat{v}_t\}$, the residuals from the VAR(1) model (3.2) computed using the OLS estimator $\hat{\rho}_n$ of ρ in (3.2). However, a key issue is that $\{\hat{v}_t\}$ inherits the small sample bias of $\hat{\rho}_n$; see Nicholls and Pope (1988). To overcome this problem, Amihud et al. (2008) consider first the expression for the bias of the OLS estimator $\hat{\rho}_n$ in Nicholls and Pope (1998) given as:

$$E[\hat{\rho}_n - \rho] = -b/n + O(n^{-3/2}) \quad (3.6)$$

Then, they provide an iterative procedure in order to estimate b and, using equation (3.6), they compute a bias corrected estimator $\hat{\rho}_n^c$ of ρ . Consequently, they calculate the reduced bias residuals $\{\hat{v}_t^c\}$, with $\hat{\rho}_n^c$ instead of $\hat{\rho}_n$ in (3.2). Their bias corrected estimate $\hat{\beta}_n^c$ of β is the OLS estimator in the regression model (3.5) with $\{\hat{v}_t^c\}$ instead of $\{v_t\}$. Finally, in order to make inference on the parameter β , they estimate:

$$E[(\hat{\beta}_n^c - \beta)(\hat{\beta}_n^c - \beta)'] \quad (3.7)$$

and approximate $cov(\hat{\beta}_n^c) = E[(\hat{\beta}_n^c - E[\hat{\beta}_n^c])(\hat{\beta}_n^c - E[\hat{\beta}_n^c])']$ through expression (3.7). Using this approximation, they test the null hypothesis of no predictability with the conventional t and Wald tests; see Amihud et al. (2008) for more details.

3.3.3 Subsampling Method

The tests defined in Wolf (2000) and Choi and Chue (2007) are based on the subsampling theory developed since Politis and Romano (1992, 1994a). The subsampling is a resampling procedure, which aims to approximate the sampling distribution of a statistic. More precisely, let $\mathcal{X}_{(n)} = ((Y_1, Z'_0), \dots, (Y_n, Z'_{n-1}))$, with $Z'_i = (1, X'_i)$, be an observation sample of process (3.1)-(3.2) and consider a statistic $T_{(n)} = T(\mathcal{X}_{(n)})$. For $i = 1, \dots, n$, the subsampling approach splits the original sample in overlapping blocks $\mathcal{X}_{(m),i} = ((Y_i, Z'_{i-1}), \dots, (Y_{i+m-1}, Z'_{i+m-2}))$ of size $m < n$. Then, it applies the statistic T to the so generated blocks, $T_{(m),i} = T(\mathcal{X}_{(m),i})$. Finally, using the empirical distribution of $T_{(m),i}$ as an approximation of the distribution of $T_{(n)}$, it computes confidence intervals for the parameter of interest.⁴

In order to test the predictability hypothesis in the stationary case, Wolf (2000) considers the statistics $T_{(n)}^W = \sqrt{n}(\hat{\beta}_n - \beta_0)$ and $T_{(m),i}^W = \sqrt{m}(\hat{\beta}_{m,i} - \hat{\beta}_n)$, where $\hat{\beta}_n$ and $\hat{\beta}_{m,i}$ are the OLS estimators of β based on the whole sample and the subsampling blocks of size m , respectively. In order to extend the subsampling approach to the local-to-unit framework, Choi and Chue (2007) propose instead the studentized statistics $T_{(n)}^{CC} = \sqrt{n}(\hat{\beta}_n - \beta_0)/\hat{\sigma}_n$ and $T_{(m),i}^{CC} = \sqrt{m}(\hat{\beta}_{m,i} - \hat{\beta}_n)/\hat{\sigma}_{m,i}$, where $\hat{\sigma}_n$ and $\hat{\sigma}_{m,i}$ are estimates of the standard deviation of $\hat{\beta}_n$ and $\hat{\beta}_{m,i}$, respectively.

Andrews and Guggenberger (2009a, 2010a,b) and Mikusheva (2007) show that confidence intervals based on the subsampling distribution of the statistic $T_{(m),i}^{CC}$ may imply a distorted size, which is instead correct using the symmetric statistic $T_{(n)} = \sqrt{n}|\hat{\beta}_n - \beta_0|/\hat{\sigma}_n$. Therefore, in our robust analysis we consider the subsampling distribution based on $T_{(n)}$. Finally, a main issue in the application of

⁴Alternatively, it is possible to consider nonoverlapping blocks as well.

subsampling procedures is the choice of the block size m , since accuracy of the subsampling distribution depends on this parameter. In our experiments, we apply a data driven calibration method (CM) for the selection of the block size, introduced in Romano and Wolf (2001). Details are provided in Appendix A.4.⁵

3.3.4 The Robustness Problem of Standard Tests of Predictability

We compare through Monte Carlo simulations the accuracy and robustness of the Bonferroni approach, the bias-corrected method and the subsampling procedures for testing the predictability hypothesis.

We generate $N = 1000$ samples $\mathcal{X}_{(n)} = ((Y_1, Z'_0), \dots, (Y_n, Z'_{n-1}))$, with $Z'_{t-1} = (1, X'_{t-1})$, $t = 1, \dots, n$, of size $n = 192$, according to model (3.1)-(3.2) for the parameter choices $p = 1$, $\alpha = \mu = 0$, $\rho = 0.95$, $\Sigma = 1$, $\phi = -0.75$, and $\beta_0 = 0, 0.025, 0.05, 0.075, 0.1$. To study the robustness of the different methods under investigation, we consider replacement outliers random samples $\tilde{\mathcal{X}}_{(192)} = ((\tilde{Y}_1, Z'_0), \dots, (\tilde{Y}_{192}, Z'_{191}))$ generated according to:

$$\tilde{Y}_t = (1 - p_t)Y_t + p_t \cdot Y_{3max}, \quad (3.8)$$

where $Y_{3max} = 3 \cdot \max(Y_1, \dots, Y_{192})$ and p_t is an iid $0 - 1$ random sequence, independent of process (3.1)-(3.2) and such that $P[p_t = 1] = \eta$. The probability of contamination is set to $\eta = 5\%$, which is a small contamination of the original sample. For the subsampling procedures, we select the block size applying CM to the set of admissible block sizes $\mathcal{M} = \{12, 16, 20, 24\}$. The significance level is $1 - \alpha = 0.9$.

We first analyze the finite sample coverage and the power of the methods under investigation in a test of the null hypothesis $\mathcal{H}_0 : \beta_0 = 0$. The first 3 panels of Figure 3.1 plots the empirical frequencies

⁵Romano and Wolf (2001) introduce two data driven methods for the selection of the block size: The minimum confidence index volatility (MCIV) and the calibration method (CM). Because of the better accuracy and higher robustness of the CM, showed in Camponovo et al. (2009b), we focus on the CM only.

of rejection of the null hypothesis $\mathcal{H}_0 : \beta_0 = 0$ for different values of the alternative hypothesis: $\beta_0 = 0, 0.025, 0.05, 0.075, 0.1$.

Without contamination (straight line), when $\beta_0 = 0$, the size of all methods considered is close to the nominal level $\alpha = 0.1$: For all procedures, the difference with the nominal level is less than 1%. As expected, for increasing values of β_0 the power of the tests increases. When $\beta_0 = 0.05$, all methods have a frequency of rejection larger than 65% and when $\beta_0 = 0.075$ larger than 90%. The accuracy of all these procedures dramatically deteriorates when applied to contaminated samples (dashed line). When $\beta_0 = 0$, the Bonferroni approach and the bias corrected methods tend to overreject \mathcal{H}_0 . In contrast, the classic subsampling underrejects \mathcal{H}_0 , implying a size lower than 5%. Moreover, when $\beta_0 > 0$ the power of all procedures is much lower than that in the case without contamination. When $\beta_0 = 0.05$, all methods have a proportion of rejection less than 30%, which is for $\beta_0 = 0.075$ less than 45%.

3.4 Robust Predictive Regression

To overcome the lack of robustness of existing tests for predictability, a possibility could be to directly develop robust versions of these approaches. However, this task may be hard to achieve in general. The Bonferroni approach introduced in Campbell and Yogo (2006) is based on the nonrobust OLS estimators $\hat{\beta}_n$ and $\hat{\rho}_n$. To robustify it, it would first be necessary to consider instead robust estimators of β and ρ , respectively. However, for making inference on parameter β the asymptotic distribution of these estimators in the local to-unit-root case has to be computed. This modification extremely complicates the construction of Bonferroni confidence intervals and makes the procedures introduced in Campbell and Yogo (2005) hard to implement. Indeed, this objective required a not obvious extension to the predictive regression setting of the robust local to unit root asymptotic in Lucas (1995, 1997).

In the bias corrected method of Amihud et al. (2008), the reduce bias residuals $\{\hat{v}_t^c\}$ are computed using the expression for the bias of the nonrobust OLS estimator in Nicholls and Pope (1988). To robustify this method, it would first be necessary to compute a similar expression for the bias of a robust estimator. Moreover, in the augmented regression model (3.5), we have to consider also a robust estimator of β as well. Finally, for making inference on parameter β , a robust procedure for the computation of the covariance of these estimators is required.

To robustify the subsampling approach an obvious first idea is to apply the standard subsampling method to a robust statistic, instead of a nonrobust one. However, this approach does not provide a robust subsampling distribution. Indeed, even when applied to maximal breakdown point, i.e. maximally robust, statistics, subsampling and more general resampling procedures suffer from a low robustness of their resampling quantiles, especially in a time series setting; see Singh (1998) and Salibian-Barrera and Zamar (2002) for the iid bootstrap, Camponovo et al. (2009a) for the iid subsampling and Camponovo et al. (2009b) for the general time series case. Moreover, as discussed in Salibian-Barrera and Zamar (2002) and Camponovo et al (2009a,b), the application of standard resampling methods in nonlinear robust time series settings can become easily computationally unfeasible.

In order to test the hypothesis of predictability in an accurate and robust way, we propose the robust fast subsampling procedure introduced in Camponovo et al. (2009b). Similar to the studentized subsampling, the robust fast subsampling constructs an empirical distribution function that approximates the distribution of $T_{(n)}^R = \sqrt{n}|\hat{\beta}_n^R - \beta_0|/\hat{\sigma}_n^R$, where nonrobust OLS estimator $\hat{\beta}_n$ of β is replaced by a robust estimator $\hat{\beta}_n^R$ and $\hat{\sigma}_n^R$ is a robust scale estimate. However, the crucial difference between the studentized classic and robust fast subsampling consists in the computation of the block statistics $T_{(m),i}^R$. More precisely, instead of exactly computing these statistics, the robust fast subsampling replaces them by a first order approximations based on a Taylor expansion of $T_{(m),i}^R$. Consequently,

because of the approximations of the block statistics, the fast subsampling method provide very fast and computationally feasible procedures, which requires only computation of estimators for statistic $T_{(n)}^R$. Moreover, as shown in Camponovo et al. (2009b), when applied to robust statistics $T_{(n)}^R$ this approach implies robust resampling procedures as well. Indeed, the robust fast subsampling inherits directly the degree of robustness of the robust estimators used to compute $T_{(n)}^R$. In Appendix A.3, we explain in detail how to apply this method to the model (3.1)-(3.2).

Through Monte Carlo simulations, we can study the accuracy and robustness of our robust fast subsampling tests for predictability. With the same Monte Carlo setting samples of Section 3.3.4, we analyze the finite sample coverage and the power of the robust fast subsampling in a test of the null hypothesis $\mathcal{H}_0 : \beta_0 = 0$. The bottom panel in Figure 3.1 plots the empirical frequencies of rejection of the null hypothesis $\mathcal{H}_0 : \beta_0 = 0$, implied by our robust approach for different values of the alternative hypothesis: $\beta_0 = 0, 0.025, 0.05, 0.075, 0.1$.

Without contamination (straight line), when $\beta_0 = 0$, the size of the robust fast subsampling is 0.109, which is very close to the nominal level $\alpha = 0.1$. Moreover, for large values of β_0 the frequencies of rejections increase with power values similar to those observed for the nonrobust methods. For $\beta_0 = 0.05$ the power is 65.7%, while for $\beta_0 = 0.075$ the frequency of rejection is 92.9%. For the nonrobust methods the power ranges from 64.5% to 70.0% and from 92.4% to 96.3% for $\beta_0 = 0.05$ and $\beta_0 = 0.075$, respectively. Finally, the robust fast subsampling implies accurate results even with contaminated samples (dashed line). When $\beta_0 = 0$, the size is 0.093 and consequently close to the nominal level. For $\beta_0 > 0$, the power curve of the robust fast subsampling remains similar to the power curve observed without contaminations, indicating that our robust approach clearly outperforms classical methods. In particular, when $\beta_0 = 0.05$ and $\beta_0 = 0.075$ the power is close to 45% and 70%, respectively. For the nonrobust methods, the proportion of rejections ranges instead from 16.1% to 29.4% and from 26.9% to 44.7% for $\beta_0 = 0.05$ and $\beta_0 = 0.075$, respectively. This

corresponds to an increase of at least 50% of the power of our robust method relative to all classical procedures.

3.5 Predictability of Stock Returns

We implement the procedures introduced in the previous sections on US equity data. We consider monthly S&P 500 index data (1871-2008) from Shiller (2000). In our analysis, we study the forecast ability of dividend yields for future stock returns. We define the one-period real total return as:

$$R_t = (P_t + d_t)/P_{t-1} \quad (3.9)$$

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

$$D_t = d_t + (1 + r_t)d_{t-1} + (1 + r_t)(1 + r_{t-1})d_{t-2} + \cdots + (1 + r_t)(1 + r_{t-1} \cdots (1 + r_{t-10}))d_{t-11} \quad (3.10)$$

where r_t is the one-month Treasury-bill rate. For $t = 1, \dots, n$, we consider the predictive regression model:

$$\ln(R_t) = \alpha + \beta \left(\frac{D_{t-1}}{P_{t-1}} \right) + \epsilon_t \quad (3.11)$$

Finally, using the Bonferroni approach, the bias corrected method, the classical subsampling and our robust fast subsampling, we test the null hypothesis of no predictability $\mathcal{H}_0 : \beta_0 = 0$.

We apply the procedures under investigation to the period 1961-2008, consisting of 576 observations and the subperiods 1961-1992 and 1993-2008, consisting of 384 and 192 observations, respectively. In order to test $\mathcal{H}_0 : \beta_0 = 0$, for each method and period, we construct 90% confidence intervals for parameter β . Table 3.1 reports our empirical results.

In the whole period 1961-2008, all procedures under investigations provide very similar confidence intervals and provide evidence in favor of the predictability of stock returns. Similar findings arise for the subperiod 1961-1992: Also in this case, the Bonferroni approach, the bias corrected method, the classic subsampling and our robust fast subsampling find evidence in favor of predictability. Finally, in the subperiod 1993-2008, only our robust approach produces significant evidence of predictability. In particular, the nonrobust procedures provide larger confidence intervals, which imply a non-rejection of \mathcal{H}_0 . This is an interesting finding. A possible source of the divergent conclusions of nonrobust methods and our approach is the larger proportion of anomalous observations in the subperiod 1993-2008. As shown through Monte Carlo simulations, the presence of outliers in the data can dramatically decrease the power and the accuracy of nonrobust procedures. The finding for the subperiod 1993-2008 seems to confirm this result. Indeed, the year 2008 is characterized by several unusual observations linked to the recent credit crisis: For instance we find that observation October 2008 is the most influential data point for the whole period 1961-2008. Such outliers deteriorate the accuracy of classical methods and the non-rejection of \mathcal{H}_0 due to the large confidence intervals provided by nonrobust methods suggests a low power of these approaches in these cases.

The results in Table 3.1 represent an initial empirical finding in the predictability analysis based on our data set. We are aware that in order to provide more general findings, a wider empirical analysis based on several data sets combined with a more detailed theoretical analysis of the tests procedures are required. Nevertheless, through our theoretical and our preliminary empirical results, we find that the robustness and reliability of our approach are very important aspects to consider for testing the predictability.

3.6 Conclusions

A large literature studies the forecasting ability of a variety of explanatory variables for future stock returns. In particular, several recent testing procedures have been proposed, in univariate and multivariate predictive regression models with correlated errors and nearly integrated regressors. All these methods produce desirable inference properties under the strict model assumptions. However, we find that even small deviations from the assumptions of the strict predictive regression model can dramatically deteriorate the accuracy of these tests. To overcome this problem, we propose a robust fast subsampling test for predictability. Our method works under weak consistency conditions in a multivariate framework and in presence of potentially nearly integrated regressors. Monte Carlo simulations confirm the fragility of classical tests of predictability when applied to data which only approximatively follow the predictive regression model assumptions. In contrast, our robust approach shows a desirable stability even for different contaminations by outliers. Finally, in the application to US equity data, our procedure shows stronger evidence in favor of predictability of stock returns by dividend yields than classic methods. In the whole period 1960-2008 and in the subperiods 1960-1992 and 1993-2008, the robust fast subsampling always rejects the hypothesis of no predictability. In contrast, in the subperiod 1993-2008 the classic methods provide larger confidence intervals, which imply a non-rejection of \mathcal{H}_0 . This finding seems to confirm the lack of power of nonrobust methods highlighted in our Monte Carlo simulations. Indeed, the presence of an even moderate proportion of anomalous observations in this subperiod deteriorates the accuracy of nonrobust methods. These findings indicate that robustness is a key aspect to consider when testing the predictability in predictive regression models.

A.3 Robust Fast Subsampling and Predictive Regression Model

We compute the robust fast subsampling distribution for the multivariate predictive regression model (3.1)-(3.2). For $t = 1, \dots, n$, we write regression model (3.1) as:

$$y_t = \theta' z_{t-1} + u_t \quad (3.12)$$

where $\theta = (\alpha, \beta')$ and $z_{t-1} = (1, x'_{t-1})'$. Let $\mathcal{X}_{(n)} = ((Y_1, Z'_0), \dots, (Y_n, Z'_{n-1}))$ be an observation sample according to (3.12). Furthermore, for $m < n$ and $i = 1, \dots, n - m + 1$, denote by $\mathcal{X}_{(m),i} = ((Y_i, Z'_{i-1}), \dots, (Y_{i+m-1}, Z'_{i+m-2}))$ the $n - m + 1$ overlapping blocks of size m .

We consider the robust estimator $\hat{\theta}_n^R$ of true parameter $\theta_0 = (\alpha_0, \beta'_0)'$, defined as solution of the equation $\psi_{n,c}(\mathcal{X}_{(n)}, \hat{\theta}_n^R) = 0$, where:

$$\psi_{n,c}(\mathcal{X}_{(n)}, \theta) = \frac{1}{n} \sum_{i=1}^T g_c(Y_i, Z_{i-1}, \theta) \quad (3.13)$$

and

$$g_c(Y_i, Z_{i-1}, \theta) = (Y_i - \theta Z_{i-1}) Z_{i-1} \cdot \min \left(1, \frac{c}{\|(Y_i - \theta Z_{i-1}) Z_{i-1}\|} \right). \quad (3.14)$$

For $j = 1, \dots, p$, we denote by $\theta_0^{(j+1)} = \beta_0^{(j)}$ the j -th component of parameter β_0 . Let $\hat{\sigma}_n^{R,j}$ be a robust estimator of the standard deviation of $\hat{\beta}_n^{R(j)}$. The robust fast subsampling approximates the distribution of $\frac{|\hat{\beta}_n^{R(j)} - \beta_0^{(j)}|}{\hat{\sigma}_n^{R,j}}$ by the empirical distribution of:

$$\frac{\left| \left(-[\nabla_{\theta} \psi_{m,c}(\mathcal{X}_{(m),i}, \hat{\theta}_n^R)]^{-1} \psi_{m,c}(\mathcal{X}_{(m),i}, \hat{\theta}_n^R) \right)^{(j)} \right|}{\hat{\sigma}_{m,i}^{R,j}} \quad (3.15)$$

where $\nabla_{\theta} \psi_{m,c}(\mathcal{X}_{(m),i}, \hat{\theta}_n^R)$ denotes the derivative of function $\psi_{m,c}$ with respect to θ and $\hat{\sigma}_{m,i}^{R,j}$ denotes the estimated standard deviation of $\hat{\beta}_m^{R(j)}$, using the subsampling block $\mathcal{X}_{(m),i}$; see also Hong and

Scaillet (2006).

Finally, given $t \in (0, 1)$ let Q_t be the upper quantile of the empirical distribution implied by the robust fast subsampling statistics (3.15). The symmetric robust fast subsampling symmetric t -confidence interval for parameter $\beta_0^{(j)}$ is given by:

$$CI_t = [\hat{\beta}_n^{R(j)} - \hat{\sigma}_n^{R,j} Q_t, \hat{\beta}_n^{R(j)} + \hat{\sigma}_n^{R,j} Q_t] \quad (3.16)$$

A.4 Block Size Selection

An important issue in the application of subsampling procedures is the choice of the block size m , since accuracy of the resampling distribution depends on this parameter. In this section, we propose a data driven selection method, introduced in Romano and Wolf (2001), based on a calibration approach.

Fix $t \in (0, 1)$, $i \in \{1, \dots, p\}$ and let $\beta_0^{(i)}$ be the parameter of interest. We denote by $\mathcal{M} = \{m_{min}, \dots, m_{max}\}$ the set of admissible block sizes. Let $\mathcal{X}_{(n)}^*$ be a nonoverlapping moving block bootstrap sample generated from $\mathcal{X}_{(n)}$ with block size m . For each bootstrap sample, denote by $CI_t^*(m)$ the subsampling t -confidence interval according to block size m ; see Appendix A.3 and equation (3.16) for details on the construction of $CI_t^*(m)$. The data driven block size according to the calibration method is defined by

$$m_{CM} := \arg \inf_{m \in \mathcal{M}} \{ |t - P^* [\hat{\beta}_n^{(i)} \in CI_t^*(m)]| : P^* [CI_t^*(m) \in \mathbb{R}] > 1 - t \} \quad (3.17)$$

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

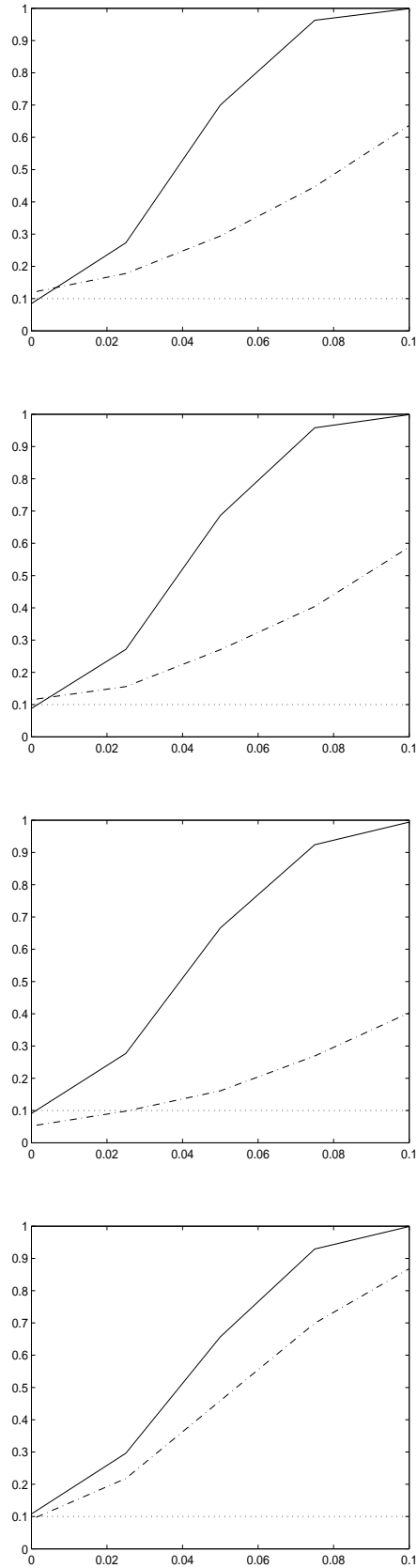


Figure 3.1: **Power Curves.** We plot the proportion of rejections of the null hypothesis $\mathcal{H}_0 : \beta_0 = 0$, when the true parameter value is $\beta_0 \in [0, 0.1]$. From the top to the bottom, we present power curves for the Bonferroni approach, the bias-corrected method, the subsampling and the robust fast subsampling. We consider non contaminated samples (straight line) and contaminated samples (dashed line).

	(1961 – 2008) (576)	(1961 – 1992) (384)	(1993 – 2008) (192)
Bonferroni	[0.0281, 0.0420]	[0.0185, 0.0384]	[−0.0045, 0.1209]
Bias Corrected	[0.0266, 0.0423]	[0.0160, 0.0398]	[−0.0028, 0.1075]
Subsampling	[0.0242, 0.0498]	[0.0147, 0.0485]	[−0.0104, 0.1317]
R.F.Subsampling	[0.0269, 0.0468]	[0.0179, 0.0447]	[0.0327, 0.1128]

Table 3.1: **Stock Returns Predictability.** We report 90% confidence intervals for the parameter β in model (3.11). We consider the Bonferroni approach, the bias corrected method, the classic subsampling and our robust fast subsampling for the period 1961-2008 (576 observations) and the subperiods 1961-1992 and 1993-2008 (384 and 192 observations, respectively).

Bibliography

- [1] Amihud, Y., Hurvich C.M., and Y. Whang, 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, No 3, 673–700.
- [5] Andrews, D., and P. Guggenberger 2009a. Hybrid and size-corrected subsample methods. *Econometrica*, 77, forthcoming.
- [6] Andrews, D., and P. Guggenberger, 2009b. Incorrect asymptotic size of subsampling procedures based on post-consistent model selection estimators. *Journal of Econometrics*, 152, 19–27.
- [7] 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.

- [8] Andrews, D., and P. Guggenberger, 2010b. Application of subsampling, hybrid and size-correction methods. *Journal of Econometrics*, forthcoming.
- [9] Bickel, P., Gotze, F., and W. van Zwet, 1997. Resampling fewer than n observations: Gains, losses, and remedies for losses. *Statistica Sinica*, 7, 1–31.
- [10] Bühlmann, P., and H. Künsch, 1999. Block length selection in the bootstrap for time series. *Computational Statistics and Data Analysis*, 31, 295–310.
- [11] Campbell, J.Y., and R.J. Shiller 1988. The dividend ratio model and small sample bias: a Monte Carlo study. *Economics Letters*, 29, 325–331.
- [12] Campbell, J.Y., and M. Yogo 2005. Implementing the econometric methods in "Efficient tests of stock return predictability". Unpublished working paper. University of Pennsylvania.
- [13] Campbell, J.Y., and M. Yogo 2006. Efficient tests of stock return predictability. *Journal of Financial Economics*, 81, 27–60.
- [14] Camponovo, L., O. Scaillet, and F. Trojani 2009a. Robust subsampling, Working paper.
- [15] Camponovo, L., O. Scaillet, and F. Trojani 2009b. Robust resampling methods for time series, Working paper.
- [16] Camponovo, L., O. Scaillet, and F. Trojani 2009c. Robust predictive regression and stock returns predictability, Working paper.
- [17] 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.
- [18] Chernozhukov, V., and I. Fernandez-Val, 2005. Subsampling inference on quantile regression processes. *Sankhya*, 67, 253–276.

- [19] Choi, I., and T.K. Chue, 2007. Subsampling-based tests of stock-return predictability, Working paper.
- [20] Cowell, F. A., and E. Flachaire, 2007. Income distribution and inequality measurement: the problem of extreme values. *Journal of Econometrics*, 141, 1044–1072.
- [21] Datta, S., 1995. On a modified bootstrap for certain asymptotically non-normal statistics. *Statistics and Probability Letters*, 24, 91–98.
- [22] Davidson, R., and E. Flachaire, 2007. Asymptotic and bootstrap inference for inequality and poverty measures. *Journal of Econometrics*, 141, 141–166.
- [23] Davidson, R., and J.G. MacKinnon, 1999. Bootstrap testing in nonlinear models. *International Economic Review*, 40, 487–508.
- [24] Efron, B., 1979. Bootstrap methods: Another look at the jackknife. *Annals of Statistics*, 7, 1–26.
- [25] Efron, B., and R. Tibshirani, 1993. *An Introduction to the bootstrap*. New York: Chapman and Hall.
- [26] Fama, E., and K. French 1988. Dividend yields and expected stock returns. *Journal of Financial Economics*, 22, 3-25.
- [27] Gagliardini P., Trojani F., and G. Urga (2005). Robust GMM tests for structural breaks. *Journal of Econometrics*, 129, 139-182.
- [28] 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.

- [29] Goetzmann, W.N., and P. Jorion 1993. Testing the predictive power of dividend yields. *Journal of Finance*, 48, No.2, 663-679.
- [30] Goetzmann, W.N., and P. Jorion 1995. A longer look at dividend yields. *Journal of Business*, 68, No.4.
- [31] Goncalves, S., and H. White (2004). Maximum likelihood and the bootstrap for nonlinear dynamic models. *Journal of Econometrics*, 119, 199–220.
- [32] Gonzalo, J., and M. Wolf, 2005. Subsampling inference in threshold autoregressive models. *Journal of Econometrics*, 127, 201–224.
- [33] Hall, P., 1985. Resampling a coverage process. *Stochastic Process Applications*, 19, 259–269.
- [34] Hall, P., 1988. On symmetric bootstrap confidence intervals. *Journal of the Royal Statistical Society, Ser.B*, 50, 35–45.
- [35] Hall, P., 1992. *The Bootstrap and Edgeworth Expansion*. New York: Springer-Verlag.
- [36] Hall, P., and J. Horowitz, 1996. Bootstrap critical values for tests based on Generalized-Method-of-Moment estimators. *Econometrica*, 64, 891–916.
- [37] Hall, P., Horowitz J., and B.-Y. Jing 1995. On blocking rules for the bootstrap with dependent data. *Biometrika*, 82, 561–574.
- [38] Hall, P., and Q. Yao, 2003. Inference in ARCH and GARCH models with heavy-tailed errors. *Econometrica*, 71, 285–317.
- [39] Hampel, F., Ronchetti, E., Rousseeuw, P., and W. Stahel, 1986. *Robust statistics. The approach based on influence functions*. Wiley, New York.

- [40] Heritier, S., and E. Ronchetti 1994, Robust bounded-influence tests in general parametric model. *Journal of the American Statistical Association*, 89, 897-904.
- [41] Hong, H., and O. Scaillet, 2006. A fast subsampling method for nonlinear dynamic models. *Journal of Econometrics*, 133, 557–578.
- [42] Hu, F., and J. Kalbfleisch, 2000. The estimating function bootstrap. *Canadian Journal of Statistics*, 28, 449–499
- [43] Huber, P., 1981. *Robust Statistics*. Wiley, New York.
- [44] Jansson, M., and M.J. Moreira 2006. Optimal inference in regression models with nearly integrated regressors. *Econometrica*, 74, 681–714.
- [45] Kim, T. H., White H., and D. Stone, 2005. Asymptotic and bayesian confidence intervals for sharpe-style weights. *Journal of Financial Econometrics*, 3, 315–343.
- [46] Künsch, H., 1984. Infinitesimal robustness for autoregressive processes. *Annals of Statistics*, 12, 843–863.
- [47] Künsch, H., 1989. The jackknife and the bootstrap for general stationary observations. *Annals of Statistics*, 17, 1217–1241.
- [48] Lee, S., and M. Pun, 2006. On m out of n bootstrapping for nonstandard M-estimation with nuisance parameters. *Journal of the American Statistical Association*, 101, 1185–1197.
- [49] Linton, O., Massoumi E., and Y. Whang, 2005. Consistent testing for stochastic dominance under general sampling schemes. *The Review of Economic Studies*, 72, 735–765.
- [50] Loh, W. Y., 1987. Calibrating confidence coefficients. *Journal of the American Statistical Association*, 82, 155–162.

- [51] Lucas, A., 1995. Unit root tests based on M-estimators. *Econometric Theory*, 11, 331–346.
- [52] Lucas, A., 1997. Cointegration testing using pseudolikelihood ratio tests. *Econometric Theory*, 13, 149–169.
- [53] 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.
- [54] Mankiv, N.G., and M.D. Shapiro 1986. Do we reject too often? Small sample properties of tests of rational expectation models. *Economics Letters* 20, 139-145.
- [55] Mikusheva, A., 2007. Uniform inference in autoregressive models. *Econometrica*, 75, 1411–1452.
- [56] Muler, N., and V. Yohai, 2008. Robust estimates for GARCH models. *Journal of Statistical Planning and Inference*, 138, 2918–2940.
- [57] Nelson, C.R., and M.J. Kim 1993. Predictable stock returns: the role of small sample bias. *Journal of Finance*, 48, No.2, 641-661.
- [58] Ortelli, C., and F. Trojani, 2005. Robust efficient method of moments. *Journal of Econometrics*, 128, 69–97.
- [59] Politis, D., and J. Romano, 1992. A general theory for large sample confidence regions based on subsamples under minimal assumptions. Technical Report 399, Dept of Statistics, Stanford University.
- [60] Politis, D., and J. Romano, 1994a. Large sample confidence regions based on subsamples under minimal assumptions. *Annals of Statistics*, 22, 203–2050.
- [61] Politis, D. N., and J. P. Romano, 1994b. The stationary bootstrap. *Journal of the American Statistical Association*, 89, 1303–1313.

- [62] Politis, D., Romano, J., and M. Wolf, 1999. *Subsampling*. Springer, New York.
- [63] Polk, C., S. Thompson, and T. Vuolteenaho 2006. Cross-sectional forecast of the equity premium. *Journal of Financial Economics*, 81, 101-141.
- [64] Romano, J., and M. Wolf, 2001. Subsampling intervals in autoregressive models with linear time trend. *Econometrica*, 69, 1283-1314.
- [65] Ronchetti, E., and F. Trojani, 2001. Robust inference with GMM estimators, *Journal of econometrics*, 101, 37–69.
- [66] Rozeff, M. 1984. Dividend yields are equity risk premium. *Journal of Portfolio Management*, 11, 68-75.
- [67] Sakata, S., and H. White, 1998. High breakdown point conditional dispersion estimation with application to S&P500 daily returns volatility. *Econometrica*, 66, 529–567.
- [68] Salibian-Barrera, M., 2006a. The asymptotic of MM-estimators for linear regression with fixed designs. *Metrika*, 63, 283–294.
- [69] Salibian-Barrera, M., 2006b. Bootstrapping MM-estimators for linear regression with fixed designs. *Statistics & Probability Letters*, 76, 1287–1297.
- [70] 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.
- [71] Salibian-Barrera, M., Van Aelst, S., and G. Willems, 2007. Fast and robust bootstrap. *Statistical Methods and Applications*, 17, 41–71.
- [72] Salibian-Barrera, M., and R. Zamar, 2002. Bootstrapping robust estimates of regression. *Annals of Statistics*, 30, 556–582.

- [73] Shao, J., and D. Tu, 1995. *The jackknife and bootstrap*. Springer, New York.
- [74] Shiller, R.J. 2000. *Irrational Exuberance*. Princeton University Press, Princeton, New York.
- [75] Singh, K., 1998. Breakdown theory for bootstrap quantiles. *Annals of Statistics*, 26, 1719–1732.
- [76] Stambaugh, R.F. 1986. Bias in regressions with lagged stochastic regressors. Graduate School of Business, University of Chicago, Working Paper No. 156.
- [77] Stambaugh, R.F. 1999. Predictive regressions. *Journal of Financial Economics*, 54, 375–421.
- [78] Torous, W., R. Valkanov, and S. Yan 2004. On predicting stock returns with nearly integrated explanatory variables. *Journal of Business*, 77, 937–966.
- [79] 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.
- [80] Wu, C.F., 1990. On the asymptotic properties of the jackknife histogram. *Annals of Statistics*, 18, 1438–1452.
- [81] Yohai, V, 1987. High breakdown-point and high efficiency robust estimates for regression. *Annals of Statistics*, 15, 642–656.