Robust Subsampling*

Lorenzo Camponovo

University of Lugano

Olivier Scaillet

University of Geneva and Swiss Finance Institute

Fabio Trojani

University of Lugano and Swiss Finance Institute

First Version: November 2006. This Version: July 2011.

*We thank the editor (Takeshi Amemiya), the associate editor and two anonymous referees for useful comments. The authors acknowledge the financial support of the Swiss National Science Foundation (NCCR FINRISK and grants 101312-103781/1, 100012-105745/1, and PDFM1-114533). We thank participants at the ESEM 2006 in Vienna, the 2007 Meeting of the Swiss Society of Economics and Statistics in St. Gallen, the 2007 International Workshop on Computational and Financial Econometrics in Geneva, and the 2007 International Conference on Robust Statistics in Buenos Aires, for helpful comments. Part of this research was done when the first author was visiting Yale University as a post doctoral researcher. Correspondence address: Fabio Trojani, Faculty of Economics, University of Lugano, Via Buffi 13, CH-6900 Lugano, e-mail: Fabio.Trojani@usi.ch.

Abstract

We characterize the robustness of subsampling procedures by deriving a 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 when 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 can be applied instead. To overcome these robustness problems, we introduce 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 subsampling.

Keywords: Subsampling, bootstrap, breakdown point, robustness.

JEL: C12, C13, C15.

SC 2000: Primary 62F40; Secondary 62F35.

1 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, 1994) have widespread applicability, and are useful for a wide variety of inference problems in many fields. The bootstrap has been the object of a huge research in statistics and econometrics, since its introduction by Efron (1979). Subsampling procedures 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 in this paper 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), Mancini, Ronchetti and Trojani (2005), and La Vecchia and Trojani (2010). A natural question is this context is whether standard subsampling methods applied to robust statistics already imply a robust subsampling inference. The answer is no and motivates our interest for robust versions of the subsampling approach.

The following example, inspired by Singh (1998) in the bootstrap setting, illustrates the intuition for the failing robustness of the standard subsampling, even when applied to a robust statistic. Consider the 10% trimmed mean $T_n := T(X_1, \ldots, X_n; 0.1)$ (i.e., 5% trimming each side) on a random sample $(X_1, \ldots, X_n)$ of size $n = 200$. If there are 8 outliers in the upper tail, that is, the order statistics $\{X_{(193)}, ...X_{(200)}\}$ are extraordinarily large, T_n stays unaffected due to the trimming. Now suppose a subsample $(X_1^*, \ldots, X_m^*)$ of size $m = 20$ is drawn without replacement from this sample. One outlier out of $\{X_{(193)}, ...X_{(200)}\}$ could appear, two outliers could appear, or, in the most extreme case, the whole set of outliers could appear in the subsample. Consider the subsampling distribution of the trimmed mean. If one outlier appears, the subsampling statistic $T_{n,m}^* := T(X_1^*, \ldots, X_m^*; 0.1)$ is free ot it. If at least two outliers appear, it is not. The chances for the event that $T_{n,m}^*$ is free of a single outlier is $p_0 = P[X(200, 20, .04) \leq 1]$, which is about 80%, where $X(n, m, p)$ denotes an hypergeometrically distributed random variable, with parameters n, np, and m. This means that if the order statistics diverge to infinity, $100(1 - p_0)\%$ of all subsampling statistics diverge to infinity as well. In other words, the subsampling quantile $Q_{t,n,m}^*$ of $T_{n,m}^*$, where t ranges from 0 to 1, will go to infinity for all probability levels $t > p_0 \simeq 80\%$. This illustrates that the subsampling distribution of the trimmed mean is not immune from the breakdown of its upper quantile when $t > p_0$, and is thus not robust to the presence of outliers in the original sample. This simple example also indicates that the subsampling tends to be more fragile than its bootstrap counterpart. The chances for the event that the bootstrap statistic $T_n^* := T(X_1^*, \ldots, X_n^*; 0.1)$ computed on a bootstrap sample $(X_1^*, \ldots, X_n^*)$ of size $n = 200$ is free of 10 outliers is $P[B(200, .04) \leq 10] = p_0$, which is about 82%, where $B(n, p)$ denotes a binomial random variable, with parameters n and p. Moreover, the above arguments also show that lower subsampling block sizes strengthen the subsampling robustness problem. For example, it is easy to see that in the above setting a block size $m = 10$ implies a chance of only 66% for the event that $T_{n,m}^*$ does not depend on a single outlier.

We organize the rest of the paper as follows. In Section 2.1, we focus on global subsampling instability. We derive a formula for the breakdown point of subsampling quantiles, i.e., we compute the smallest fraction of outliers in the original sample such that the subsampling quantile diverges to infinity. When this occurs, inference based on subsampling distributions becomes meaningless. Indeed, critical values of subsampling-based tests diverge to infinity, and confidence intervals computed using subsampling distributions have an infinite length. It turns out that in this case both the size and the power of subsampling-based tests collapse to zero or one. Consequently, the quantile breakdown point is an important global robustness tool for describing up to which fraction of contaminations the subsampling distribution still provides some reliable information. We show that the subsampling quantile 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 even 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 2.2, 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 2.3. 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 2.4. Monte Carlo simulations and sensitivity analysis are presented in Section 3, for two settings where the bootstrap fails. In the first example, we study the inference about the squared mean of a random variable. In the second example, we study a linear regression model with a parameter of interest near a boundary. Andrews and Guggenberger (2009a,b, 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. 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. Section 4 gathers concluding remarks.

2 Subsampling Breakdown Point and Robust Subsampling

Let $(X_1, \ldots, X_n)$ be an iid random sample from a probability distribution H on the real line and $T_n := T(X_1, \ldots, X_n)$ be 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 $+\infty$ (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 $(X_1^*, \ldots, X_m^*)$ is drawn without replacement from the original sample $(X_1, \ldots, X_n)$. $T_{n,m}^* := T(X_1^*, \ldots, X_m^*)$ denotes the subsampling statistic. Given $t \in (0, 1)$, the t-quantile of $T_{n,m}^*$ is $Q_{t,n,m}^* := \inf\{x | P\left[T_{n,m}^* \leq x\right] \geq t\}$, where, by definition, $\inf(\emptyset) := \infty$.

Definition 1 *The upper breakdown point of the subsampling distribution t-quantile $Q_{t,n,m}^* = Q_{t,n,m}^*(X_1, \ldots, X_n)$ is defined by*

$$b_{t,n,m} := \inf\{p \in [1/n, b] : np \in \mathbb{N} \text{ and } Q_{t,n,m}^*(Z_1, \ldots, Z_n) \to +\infty\}, \tag{1}$$

where the sample $(Z_1, \ldots, Z_n)$ is obtained by replacing np data points $X_{i_1}, \ldots, X_{i_{np}}$ of the original sample $(X_1, \ldots, X_n)$ by values $Y_{i_1}, \ldots, Y_{i_{np}}$, with $Y_{i_j} \to \pm\infty$, $j = 1, \ldots, np$.

By definition, $b_{t,n,m}$ is the smallest fraction of outliers in the original sample $(X_1, .., X_n)$ such that the $t-$quantile of $T_{n,m}^*$ diverges to infinity. Intuitively, $b_{t,n,m}$ 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 m out of n bootstrap are equivalent for iid observations when $m^2/n \to 0$; 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.

2.1 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*

$$b_{t,n,m} = \inf\{p \in [1/n, b] : np \in \mathbb{N} \textit{ and } P\left[X(n,m,p) < mb\right] < t\}, \tag{2}$$

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

From formula (2), $b_{t,n,m}$ 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 $mb \in \mathbb{N}$. Moreover, $b_{t,n,m} = b$ 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 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.

Insert Table 1 about here

Theorem 2 implies that we can always obtain a target upper quantile breakdown point $\hat{b} \in (1/n, b]$ by selecting a suitable block size $\hat{m} = \hat{m}(n, t, b, \hat{b})$. 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 $\hat{b} \in (1/n, b]$ be such that $n\hat{b} \in \mathbb{N}$. The smallest block size $\hat{m}$ such that $b_{t,n,\hat{m}} \geq \hat{b}$ is given by*

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

Corollary 3 implies that, for $\hat{b} = b$, it is possible to obtain a breakdown point $b_{t,n,m}$ as large as the one of the statistic T_n. As highlighted by Table 1, in order to achieve this goal, it is not in general necessary to select a trivial block size $m = n$. For instance, when $b = .25$, for sample sizes $n = 40, 80, 120$, the maximal breakdown point is achieved with $m = 37, 77, 117$, respectively. When $b = .5$, and $n = 40, 80, 120$, it is instead achieved with $m = 39, 79, 119$, respectively.

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 $m/n \to 0$ should hold as $n, m \to \infty$ (see, for instance, Politis, Romano and Wolf, 1999). This means that the application of Corollary 3 is essentially relevant only for 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 $m/n \to r \in [0,1)$, $m, n \to \infty$. Then, $b_{t,n,m} = b - z_t\sqrt{b(1-b)(1-r)}/\sqrt{m} + O(1/m)$, for n large enough, where z_t is the $t-$quantile of the standard normal distribution.*

From Corollary 4, the subsampling breakdown point $b_{t,n,m}$ converges to the breakdown point of statistic T as $n, m \to \infty$. 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 large subsampling block sizes.

2.2 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 based on either a minimization of the confidence interval volatility index (MCIV) or the calibration method (CM); see Romano and Wolf (2001).

Given a sample of size n, we denote by $\mathcal{M} = \{m_{\min} \ldots, m_{\max}\}$ the set of admissible block sizes. Both MCIV (denoted by v) and CM (denoted by c) select the data-driven block size $m_u \in \mathcal{M}$, with $u = v, c$, as solution of a problem of the form

$$m_u = \arg\inf_{m \in \mathcal{M}} \{F_{u,1}(X_1, \ldots, X_n; m) : F_{u,2}(X_1, \ldots, X_n; m) \in I_u\} \ , \tag{3}$$

where by definition $\arg\inf(\emptyset) := \infty$, $F_{u,1}$, $F_{u,2}$ are two scalar functions of the original sample $(X_1, \ldots, X_n)$ and block size m, and I_u is a subset of $\mathbb{R}$; see equations (6) and (8) below for the explicit definitions of $F_{u,1}$, $F_{u,2}$ in the setting $u = v, c$.

We characterize the robustness properties of MCIV and CM by their respectively breakdown points. More precisely, 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. Consequently, we consider the following definition for the breakdown point:

Definition 5 *The breakdown point of $m_u := m_u(X_1, \ldots, X_n)$ is defined as*

$$b_t^u := \inf\{p \in [1/n, p] : np \in \mathbb{N} \text{ and } m_u(Z_1, \ldots, Z_n) \to +\infty\}, \tag{4}$$

where the sample $(Z_1, \ldots, Z_n)$ is obtained by replacing np data points $X_{i_1}, \ldots, X_{i_{np}}$ of the original sample $(X_1, \ldots, X_n)$ by values $Y_{i_1}, \ldots, Y_{i_{np}}$, with $Y_{i_j} \to \pm\infty$, $j = 1, \ldots, np$.

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

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

$$CIV(m) := \frac{1}{2k+1}\sum_{j=-k}^{j=k}\left(Q_t^*(m+j) - \overline{Q}_t^{*k}(m)\right)^2 . \tag{5}$$

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

$$m_v = \arg\inf_{m\in\mathcal{M}}\{CIV(m) : CIV(m) \in \mathbb{R}^+\} , \tag{6}$$

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). Moreoveor, 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

$$m_{\min} = \max(c_1 n^{\zeta}, \hat{m}), \tag{7}$$

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

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

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 7. 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.

2.2.2 Calibration Method

Another consistent method for a data driven choice of the block size m can be based on a calibration procedure in the spirit of Loh (1987). 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 $(X_1^*, \ldots, X_n^*)$ be a bootstrap sample from $(X_1, \ldots, X_n)$. 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*

$$m_c := \arg \inf_{m \in \mathcal{M}} \{ |t - P^* [T_n \leq Q_t^{**}(m)] | : P^* [Q_t^{**}(m) \in \mathbb{R}] > 1 - t \}, \tag{8}$$

where, by definition, $\arg\inf(\emptyset) := \infty$*, and P^* denotes 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 (8) 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*

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

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

Insert Table 2 about here

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).

2.3 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). Such a fast approach is put forward, among others, in Davidson and McKinnon (1999) and Andrews (2002) for the bootstrap. It can be used to extend in a convenient way the robust bootstrap procedure for fixed point estimators of 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.

We consider the class of robust M-estimators $\hat{\theta}_n$ for parameter $\theta \in \mathbb{R}^d$, defined by the solution of

$$\psi_n(\hat{\theta}_n) = \sum_{i=1}^{n} f(X_i, \hat{\theta}_n) = 0, \tag{9}$$

for some function $\psi_n : \mathbb{R}^d \to \mathbb{R}^d$ depending on the parameter θ and on the sample $(X_1, \ldots, X_n)$, where f is a known $\mathbb{R}^d$-valued function. In particular, we focus on robust M-estimators with bounded influence function, i.e., bounded estimating function f, see, e.g., Hampel, Ronchetti, Rousseeuw and Stahel (1986).

The boundedness of the estimating function is a key feature for developing our robust subsampling approach. 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, 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 (9) around the true parameter value $\theta_\star$: $\psi_n(\hat{\theta}_n) = \psi_n(\theta_\star) +$

$\nabla\psi_n(\theta_\star)(\hat{\theta}_n - \theta_\star) + o_P(1)$, where $\nabla\psi_n \in \mathbb{R}^{d\times d}$ is the matrix of partial derivatives with respect to parameter θ. This implies: $\hat{\theta}_n - \theta_\star = (-\nabla\psi_n(\theta_\star))^{-1}\psi_n(\theta_\star) + o_P(1)$. 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 $(X^*_1, \ldots, X^*_m)$.

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

$$L^R_{n,m}(x) = \frac{1}{N_{n,m}} \sum_{s=1}^{N_{n,m}} \mathbb{I}\left\{\tau_m(-\nabla\psi_n(\hat{\theta}_n))^{-1}\psi^*_{n,m,s}(\hat{\theta}_n) \leq x\right\}, \tag{10}$$

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

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

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

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

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

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

(1) If x is a continuity point of $J(., H)$, then $L^R_{n,m}(x) \to J(x, H)$ as $n \to \infty$.

(2) If $J(., H)$ is continuous, then $\sup_x |L^R_{n,m}(x) - J(x, H)| \to 0$ in probability as $n \to \infty$.

(3) Given $\alpha \in (0,1)$ *define* $c_{n,m}(1-\alpha) = \inf\{x : L^R_{n,m}(x) \geq 1-\alpha\}$ *and* $c(1-\alpha, H) = \inf\{x : J(x,H) \geq 1-\alpha\}$*. If* $J(.,H)$ *is continuous at* $c(1-\alpha,H)$*, then:*

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

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 $\sqrt{n}(\widehat{\theta}_n - \theta_\star)$. Statement 3 implies that the $(1-\alpha)-$quantile of $L^R_{n,m}$ 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 $\theta_\star$ with correct asymptotic size and coverage.

The robustness improvement provided by our robust approach is observable through the definition of the robust subsampling distribution (10). Indeed, in this definition we note that the subsampling quantile breakdown point is maximal if (i) $(-\nabla\psi_n(\hat{\theta}_n))^{-1}$ does not breakdown as long as $\hat{\theta}_n$ does not breakdown, 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. In Section 2.4, we characterize explicitly the breakdown point of robust subsampling quantiles in the linear regression setting based on MM-estimates.

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 $g_n : \mathbb{R}^d \to \mathbb{R}^d$ depending on the parameter θ and the sample $(X_1, \ldots, X_n)$. This follows from writing the fixed-point equations in the form $g_n(\tilde{\theta}_n) - \tilde{\theta}_n = 0$. This corresponds to equation (9) 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 Salibian-Barrera, Val Aelst and Willems (2007) to the subsampling setting.

2.4 Robust Subsampling in the Linear Regression Model

We consider the iid linear regression model:

$$Y_i = X_i'\beta + \sigma U_i, \quad i = 1, .., n, \tag{11}$$

where Y_i is a scalar random variable, X_i an $\mathbb{R}^d-$valued random variable, $\beta \in \mathbb{R}^d$, $\sigma \in \mathbb{R}^+$, $E[U_i] = 0$, $Var(U_i) = 1$, and $E[U_i X_i] = 0$. 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 $\{(y_i, x_i')' : i = 1, .., n\}$ be a sample of observations of model (11). The MM-estimate $\widehat{\beta}_n$ of β is defined by the implicit equation:

$$\frac{1}{n}\sum_{i=1}^{n} \nabla\rho_1 \left(\frac{y_i - x_i'\widehat{\beta}_n}{\widehat{\sigma}_n} \right) x_i = 0. \tag{12}$$

In equation (12), $\nabla\rho_1$ is the derivative of a continuously differentiable, bounded and symmetric function ρ_1, satisfying the assumption (A1)-(A4) below. The estimate $\widehat{\sigma}_n$ is a scale $S-$estimate that minimizes with respect to β the M$-$estimate $\widehat{\sigma}_n(\beta)$, defined implicitly by $\frac{1}{n}\sum_{i=1}^{n} \rho_0 \left(\frac{y_i - x_i'\beta}{\widehat{\sigma}_n(\beta)} \right) = B$, 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., $\widehat{\sigma}_n = \widehat{\sigma}_n(\tilde{\beta}_n)$. The choice of B determines the breakdown point of the estimators, which is maximal for $B = 0.5$ (see, e.g., Huber, 1981).

Note that the estimates $\hat{\beta}_n$, $\hat{\sigma}_n$, and $\tilde{\beta}_n$ satisfy the equations,

$$\frac{1}{n}\sum_{i=1}^{n} \nabla\rho_1 \left(\frac{y_i - x_i'\hat{\beta}_n}{\widehat{\sigma}_n} \right) x_i = 0, \quad \frac{1}{n}\sum_{i=1}^{n} \rho_0 \left(\frac{y_i - x_i'\tilde{\beta}_n}{\widehat{\sigma}_n} \right) = B, \quad \frac{1}{n}\sum_{i=1}^{n} \nabla\rho_0 \left(\frac{y_i - x_i'\tilde{\beta}_n}{\widehat{\sigma}_n} \right) x_i = 0. \tag{13}$$

Let $\hat{\theta}_n = (\hat{\beta}_n, \hat{\sigma}_n, \tilde{\beta}_n)$. Simple calculations allow us to write the equations in (13) as a unique fixed-point equation $g_n(\hat{\theta}_n) = \hat{\theta}_n$; see the proof of Theorem 13 for further details on the computation of function g_n. Therefore, as discussed in the previous section, we can apply our robust subsampling approach to this setting. We first introduce the following notation.

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

$$\widehat{\omega}_i^* = \nabla\rho_1(\widehat{r}_i^*/\widehat{\sigma}_n)/\widehat{r}_i^*, \qquad \tilde{v}_i^* = \frac{\widehat{\sigma}_n}{nB}\rho_0(\tilde{r}_i^*/\widehat{\sigma}_n)/\tilde{r}_i^*. \tag{14}$$

With these weights, define:

$$\widehat{\beta}_{n,m}^* = \left(\sum_{i=1}^{m} \widehat{\omega}_i^* x_i^* x_i^{*\prime}\right)^{-1} \sum_{i=1}^{m} \widehat{\omega}_i^* x_i^* y_i^*, \qquad \widehat{\sigma}_{n,m}^* = \sum_{i=1}^{m} \tilde{v}_i^*(y_i^* - x_i^{*\prime}\tilde{\beta}_n). \tag{15}$$

Note that the weights $\widehat{\omega}_i$ and $\tilde{v}_i$ are computed without recalculating the estimators $\widehat{\beta}_n$, $\tilde{\beta}_n$ and $\widehat{\sigma}_n$, and are kept fixed for each subsampling block. Because of this construction, the quantities $\widehat{\beta}_{n,m}^*$ and $\widehat{\sigma}_{n,m}^*$ in (15), which are only an approximation of the "true" point estimates $\widehat{\beta}_m^*$ and $\widehat{\sigma}_m^*$ implied by the subsampling block $\{(y_i^*, x_i^{*\prime}) : i = 1, .., m\}$, may not reflect the actual variability of $\hat{\beta}_n$ and $\hat{\sigma}_n$. We overcome this problem by applying the linear correction factors defined in (17) and (18) below. 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 the implied robust point estimate in each subsampling block, the robust subsampling in Definition 12 yields a computationally feasible resampling scheme. This allows us to compute robust confidence intervals for the regression parameter β in presence of the nuisance scale parameter σ.

Definition 12 *Let $\beta_\star$ be the true parameter value in the regression model (11) and $J_n(H)$ be the sampling distribution of $\sqrt{n}(\widehat{\beta}_n - \beta_\star)$, i.e., for any $x \in \mathbb{R}^n$: $J_n(x, H) = P\left[\sqrt{n}(\widehat{\beta}_n - \beta_\star) \leq x\right]$. The robust subsampling approximation of $J_n(x, H)$ is given by*

$$L_{n,m}^R(x) = \frac{1}{N_{n,m}} \sum_{s=1}^{N_{n,m}} \mathbb{I}\left\{M_n\sqrt{m}(\widehat{\beta}_{n,m,s}^* - \widehat{\beta}_n) + d_n\sqrt{m}(\widehat{\sigma}_{n,m,s}^* - \widehat{\sigma}_n) \leq x\right\}, \tag{16}$$

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

$$M_n = \widehat{\sigma}_n \left(\sum_{i=1}^{n} \nabla^2 \rho_1(\widehat{r}_i/\widehat{\sigma}_n) x_i x_i' \right)^{-1} \sum_{i=1}^{n} \widehat{\omega}_i x_i x_i', \tag{17}$$

$$d_n = \frac{nB}{\widehat{\sigma}_n^2 \sum_{i=1}^{n} \nabla \rho_0(\tilde{r}_i/\widehat{\sigma}_n) \tilde{r}_i/\widehat{\sigma}_n} \left(\sum_{i=1}^{n} \nabla^2 \rho_1(\widehat{r}_i/\widehat{\sigma}_n) x_i x_i' \right)^{-1} \sum_{i=1}^{n} \nabla^2 \rho_1(\widehat{r}_i/\widehat{\sigma}_n) \widehat{r}_i x_i. \tag{18}$$

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 $n \to \infty$.

(A6) The following limits in probability hold as $n \to \infty$: $\widehat{\beta}_n \to \beta_\star$, $\tilde{\beta}_n \to \tilde{\beta}_\star$, $\widehat{\sigma}_n \to \sigma_\star$, where parameters $\beta_\star, \tilde{\beta}_\star$ and $\sigma_\star$ are the unique solution of the set of moment conditions: $E\left[\nabla \rho_1((Y_1 - X_1'\beta)/\sigma)\right] = 0$, $E\left[\rho_0((Y_1 - X_1'\tilde{\beta})/\sigma)\right] = B$, $E\left[\nabla \rho_0((Y_1 - X_1'\tilde{\beta})/\sigma)\right] = 0$.

(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 $r = Y_1 - X_1'\beta_\star$. The following matrices exist and are finite:

$$E\left[\frac{\nabla \rho_1(r)}{r} X_1 X_1'\right]^{-1}, E\left[\nabla \rho_1(r) X_1 X_1'\right], E\left[\nabla^2 \rho_1(r) X_1 X_1'\right]^{-1}, \tag{19}$$

$$E\left[\nabla \rho_1(r) r X_1 X_1'\right], E\left[\frac{\nabla \rho_0(r)}{r} X_1 X_1'\right]^{-1}, E\left[\nabla \rho_0(r) r\right], \tag{20}$$

$$E\left[\nabla^2 \rho_0(r) X_1 X_1'\right], E\left[\nabla^2 \rho_0(r) r X_1\right], E\left[\nabla^2 \rho_1(r) r X_1\right]. \tag{21}$$

In addition, the third matrix in (20) is not zero.

(A9) The following functions are continuous:

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

A well-known example where assumptions (A1)-(A9) are typically satisfied arises for functions ρ in the Tukey's family:

$$\rho(u) = \begin{cases} 3(u/d)^2 - 3(u/d)^4 + (u/d)^6, & \text{if } |u| \leq d, \\ 1 & \text{if } |u| > d, \end{cases}$$

where $d > 0$ is a fixed constant; see also Salibian-Barrera and Zamar (2002). We use a MM-estimator based on such functions in the Monte Carlo experiments reported below for the linear regression. 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 $n, m \to \infty$ and $m/n \to 0$: $L^R_{n,m}(x) \to J(x, H)$.*

2. *If $J(\cdot, H)$ is continuous, then the following limit in probability holds as $n, m \to \infty$ and $m/n \to 0$: $\sup_x |L^R_{n,m}(x) - J(x, H)| \to 0$.*

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

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

1. *The breakdown point $b^R_{t,n,m}$ of the $t-$quantile of the robust subsampling in Definition 12 is given by*

$$b^R_{t,n,m} = \inf\{p \in [1/n, b] : np \in \mathbb{N} \text{ and } P\left[X(n, m, p) \leq m - d\right] < t\}, \tag{22}$$

where b is the breakdown point of the robust $MM-$regression estimator $\widehat{\beta}_n$.

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

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

The assumption on the general position of $\sqrt{\widehat{\omega}_1}x_1, .., \sqrt{\widehat{\omega}_n}x_n$ is also used in Salibian-Barrera and Zamar (2002), and is needed here to ensure that the approximation $\widehat{\beta}^*_{n,m}$ of the subsampling estimate $\widehat{\beta}^*_m$ is well-defined in every subsampling block. By comparing (22) with the breakdown formula (2) of the standard subsampling in Theorem 2, we note that for reasonable parameter choices $mb << m - d = m(1 - d/m)$. Therefore, $P\left[X(n, m, p) < mb\right] << P\left[X(n, m, p) \leq m - d\right]$ and $b^R_{t,n,m} >> b_{t,n,m}$. The numerical difference between the two breakdown points can be large. Table 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 $\widehat{\beta}_n$. 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.

Insert Table 3 about here

For $b = 0.225$ and $n = 40$, the robust subsampling breakdown point is $b^R_{t,n,m} = 0.225$ 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 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.

These results have implications also for the breakdown point of m_v in Corollary 7. 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.

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 16 then follows even for settings with fixed designs.

3 Monte Carlo Study and Sensitivity Analysis

We study through Monte Carlo simulations the statistical properties (size and power) of the subsampling and the robust subsampling in estimating the confidence interval of (i) the squared mean for an iid sample and (ii) a parameter of interest in the iid linear regression model (11) when this parameter is possibly near a boundary. In both settings, the bootstrap is inconsistent, but subsampling procedures are applicable.

3.1 Squared mean

The first example concerns the sampling distribution of the squared mean for an iid normal sample with known variance. This is an informative, albeit simple, design to measure the accuracy of subsampling techniques 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.

3.1.1 Model and Estimation

Let $(X_1, \ldots, X_n)$ be an iid sample with $X_i \sim N(\mu, 1)$, and take the parameter of interest as $\theta = \mu^2$. Let the nonrobust estimator $\hat{\theta}_n^{NR}$ be the squared sample average $\hat{\theta}_n^{NR} = \left(\frac{1}{n}\sum_{i=1}^n X_i\right)^2$, so that the asymptotic distribution of $n\hat{\theta}_n^{NR}$ is a non-central chi-squared distribution with one degree of freedom and noncentrality parameter μ^2. Let $(X_1^*, \ldots, X_m^*)$ be a random subsample such that the subsampling statistic is $\hat{\theta}_{n,m}^{NR,*} = \left(\frac{1}{m}\sum_{i=1}^m X_i^*\right)^2$. We denote by $\theta_\star$ the true parameter value. Then, the subsampling approximation of the distribution of $n(\hat{\theta}_n^{NR} - \theta_\star)$ is

$$L_{n,m}^{CS}(x) = \frac{1}{N_{n,m}} \sum_{s=1}^{N_{n,m}} \mathbb{I}\left\{m(\hat{\theta}_{n,m,s}^{NR,*} - \hat{\theta}_n^{NR}) \leq x\right\}. \tag{23}$$

We refer to using (23) as “classical subsampling (CS)” .

Let us now consider a robust estimator $\hat{\theta}_n^R$ based on the square of 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

$$\psi_n(\theta) = \frac{1}{n}\sum_{i=1}^n h_c(X_i - \mu), \tag{24}$$

and $h_c(x) = x \cdot \min(1, c/|x|)$ is the Huber function. We denote by $\hat{\theta}_{n,m}^{R,*}$ the robust estimator based on the random subsample. Then, the subsampling approximation of the distribution of $n(\hat{\theta}_n^R - \theta_\star)$ is

$$L_{n,m}^{SR}(x) = \frac{1}{N_{n,m}} \sum_{s=1}^{N_{n,m}} \mathbb{I}\left\{m(\hat{\theta}_{n,m,s}^{R,*} - \hat{\theta}_n^{R}) \leq x\right\}. \tag{25}$$

We refer to using (25) as “subsampling robust (SR)”. As discussed in the introduction, subsampling a robust estimator does not deliver a robust resampling procedure.

To robustify subsampling, we use the decomposition $m(\hat{\theta}_{n,m}^{R,*} - \hat{\theta}_n^R) = m((\bar{X}_n^{R,*} - \bar{X}_n^R)^2 + 2\bar{X}_n^R(\bar{X}_n^{R,*} - \bar{X}_n^R))$, where $\bar{X}_n^{R,*}$ is the subsampling robust estimator of the true mean $\mu_\star$. Then, using our robust approach, the idea is to consider $A_{n,m,s} = (-\nabla\psi_n(\bar{X}_n^R))^{-1}\psi_{n,m,s}^*(\bar{X}_n^R)$ as an approximation of $\bar{X}_n^{R,*} - \bar{X}_n^R$. The robust subsampling approximation of the distribution of $n(\hat{\theta}_n^R - \theta_\star)$ is obtained as

$$L_{n,m}^{RS}(x) = \frac{1}{N_{n,m}} \sum_{s=1}^{N_{n,m}} \mathbb{I}\left\{m(A_{n,m,s}^2 + 2\bar{X}_n^R A_{n,m,s}) \leq x\right\}. \tag{26}$$

We refer to using (26) as "robust subsampling (RS)".

3.1.2 Numerical Results

We consider a sample size $n = 120$. Results are almost identical for $n = 40, 80$. Since the calibration method is not applicable here (bootstrap inconsistency), we use a data driven block size obtained by minimizing the CIV index for $k = 2$. For $n = 120$, the minimal recommended choice in Romano and Wolf (2001) implies a lower bound $m_{\min} = 6$ and an upper bound $m_{\max} = 22$, for $c_1 = 0.5$ and $c_2 = 2$, respectively. Beside these selections, we consider also the block size $m = 3$ as extreme case.

We study the finite sample coverage implied by subsampling (CS, SR), and robust subsampling (RS). To this end, we test the null hypothesis $\mathcal{H}_0 : \theta_\star = 0$ against $\mathcal{H}_1 : \theta_\star > 0$, under the following data generating process for X_i:

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

for $\gamma = 0$ (no contamination), $\gamma = 0.05$ (5% of contaminated data) and $\gamma = 0.10$ (10% of contaminated data). We use the true value $\mu_\star = 0$ for the size, and $\mu_\star = .25, .5$ for the power. For each of the 2000 Monte Carlo replications, resampling distributions are computed based on 200 draws. For the subsampling approximations we generate the 200 draws, instead of the $N_{n,m}$ possibilities, by permuting the original sample and taking the first m observations of the random samples. Table 4 and Table 5 summarize the empirical frequencies of rejection of the null hypothesis $\mathcal{H}_0$ for a significance level $\alpha = .05$.

Let us first analyze the size issues summarized in Table 4. We find that the empirical frequencies for the robust subsampling are quite accurate and closer to the nominal frequencies than those of the other resampling methods. Under a model contamination, the underrejection of the null hypothesis using classical subsampling can be severe: We get an empirical rejection frequency less than .02 instead of the true .05 value when $\gamma = 0.10$. According to the theoretical results, Table 4 confirms that resampling a robust statistic does not yield a robust resampling method. Indeed, for $m = 3$, and $\gamma = 0.05, 0.10$, the results implied by SR are even worse than those implied by CS. The robustness problem of SR can be in part mitigated by the selection of a larger block size ($m = 22$), nevertheless our robust approach still outperforms SR. At the same time, the quantile implied by the classical subsampling virtually explodes in presence of contamination, leading to a virtually noninformative inference. For instance, when $\gamma = 0.10$, and $m = 3$, the median .95-quantile implied by the classical subsampling approximation is 71.55, while the one implied by the robust subsampling approximation is 5.75.

Finally, in Table 5 we can analyze the power of the procedures under investigation. When $\theta_\star > 0$, the proportion of rejection of $\mathcal{H}_0$ increases as expected. Without contamination all the procedures under investigation imply similar accurate results. When $\gamma > 0$, the proportion of rejections implied by the classic approach dramatically decreases. For instance, when $\theta_\star = 0.25^2$, $\gamma = 0.10$, and $m = 6$, the power of CS and SR are less than 6% and 20%, respectively. For our robust approach, in the same case, the power is instead larger than 54%.

Unreported Monte Carlo results for the bootstrap approach confirm the inconsistency of this method in the squared mean setting. In particular, we consider a nonrobust approximation similar to (23), and the robust approximations similar to (25) and (26), but based on the bootstrap instead of the subsampling approach. When $\theta_\star = 0$, tests based on these three bootstrap distributions never reject $\mathcal{H}_0$ both in presence or absence of contaminations.

Insert Table 4 and Table 5 about here

3.2 Linear Regression

In this section we consider the iid linear regression model (11) 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 classical and robust subsampling.

3.2.1 Model and Estimation

We consider the scalar regression parameter β, which is known to satisfy the constraint $\beta \geq 0$ in the iid linear regression model:

$$\begin{aligned} Y_i &= X_i'\theta + W_i\beta + \sigma U_i, \\ &= Z_i'\eta + \sigma U_i, \qquad i = 1, \ldots, n, \end{aligned} \tag{27}$$

where Y_i, W_i are scalars, 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. In order to construct confidence intervals for parameter β using the classic subsampling, we consider the constrained estimator $\hat{\beta}_n^{NR} = \max(0, \hat{\eta}_{n(2)}^{ols})$, where $\hat{\eta}_n^{ols}$ is the (unrestricted) OLS estimator of η. For the robust subsampling, we consider the constrained estimator $\hat{\beta}_n^{R} = \max(0, \hat{\eta}_{n(2)}^{rob})$, 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 $\eta_{(2)} \geq 0$.

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 $\sqrt{n}(\hat{\beta}_n^{NR} - \beta_\star)$ is then given by

$$L_{n,m}^{NR}(x) = \frac{1}{N_{n,m}} \sum_{s=1}^{N_{n,m}} \mathbb{I}\left\{\sqrt{m}(\hat{\beta}_{n,m,s}^{NR} - \hat{\beta}_n^{NR}) \leq x\right\}. \quad (28)$$

For the robust subsampling, we follow (15) and (16) 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 $\sqrt{n}(\hat{\beta}_n^R - \beta_\star)$ is then given by

$$L_{n,m}^{R}(x) = \frac{1}{N_{n,m}} \sum_{s=1}^{N_{n,m}} \mathbb{I}\left\{\sqrt{m}\,(\hat{sub\beta}_{n,m,s}^{*} - \hat{\beta}_n^{R}) \leq x\right\}. \quad (29)$$

By construction, the theoretical results in Section 2.4 for the upper quantile breakdown point of the robust subsampling distribution in Definition 12 hold also for (29).

3.2.2 Hybrid Procedures

Using (28) and (29), 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 (28) or (29). The corresponding hybrid quantile is:

$$c_{n,m}^{H}(1-\alpha) = \max(c_{n,m}(1-\alpha), c_\infty(1-\alpha)), \quad (30)$$

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 $c_\infty(1-\alpha)$, we 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).

3.2.3 Numerical Results

We consider the iid linear regression model (27) for $d = 3, 5$ and $n = 120$. Results are almost identical for $n = 40, 80$. The true parameter vector is $\eta_\star = (0, \beta_\star, 0)'$ and $\eta_\star = (0, \beta_\star, 0, 0, 0)'$, respectively, with $\beta_\star = .05, .25$. We analyze the accuracy of subsampling procedures for fixed and data driven block sizes. For our sample size $n = 120$, the maximal recommended choice in Romano and Wolf (2001) implies a lower bound $m_{\min} = 6$ and an upper bound $m_{\max} = 33$, for $c_1 = 0.5$ and $c_2 = 3$, respectively. Nevertheless, because of convergence problems of the algorithm, for $d = 5$ block sizes less than $m = 8$ are not applicable. Besides the selections recommended by Romano and Wolf (2001), we analyze also the accuracy of the block sizes implied by Theorem 14. In particular, we consider the block sizes $m = 11$ and $m = 16$, for $d = 3$ and $d = 5$, respectively, which imply a breakdown point of 50% for the .975-quantile. Finally, since the calibration method is not applicable here, we use a data driven block size obtained by minimizing the CIV index with $k = 2$. More precisely, for $d = 3$ we apply MCIV to the intervals $I_{(d=3,1)} = \{8, \ldots, 27\}$ and $I_{(d=3,2)} = \{11, \ldots, 27\}$, while for $d = 5$, we consider the intervals $I_{(d=5,1)} = \{10, \ldots, 27\}$ and $I_{(d=5,2)} = \{16, \ldots, 27\}$. The intervals $I_{(d=3,1)}$ and $I_{(d=5,1)}$ represent the average intervals proposed by Romano and Wolf (2001) for this setting. The intervals $I_{(d=3,2)}$ and $I_{(d=5,3)}$ represent instead the optimal intervals based on Theorem 14, which ensure a breakdown point of 50% for the .975-quantile.

We make use of functions ρ_0 and ρ_1 in Tukey family. The constant for the $MM-$regression estimator in our simulations is $B = 0.5$. For this choice, we obtain a breakdown point of $\widehat{\eta}_n$ satisfying $b \geq 0.47$; see Yohai (1987, Theorem 2.1).

We first study the finite sample coverage of confidence intervals implied by classical and robust subsampling methods (denoted by CS and RS, respectively). We do not report results concerning SR since its implementation is computationally prohibitive. The simulation results obtained in the previous squared mean regression example, as the theoretical results provided by our study, are convincing enough in showing that the application of the classical subsampling to the robust MM-estimator is not sufficient to imply a robust subsampling inference for this setting.

We consider parameter choices $\beta_\star = .05, .25$ under a contaminated distribution for U:

$$U \sim (1-\gamma)N(0,1) + \frac{\gamma}{2}(N(C,(0.1)^2) + N(-C,(0.1)^2)), \tag{31}$$

where $C = 5$, $\gamma = 0$ (no contamination), $\gamma = 0.15$ (15% of contaminated data) and $\gamma = 0.25$ (25% of contaminated data), as in Salibian-Barrera and Zamar (2002). For each of the 2000 Monte Carlo replications, resampling distributions are computed based on 500 draws. For the subsampling approximations we generate the 500 draws, instead of the $N_{n,m}$ possibilities, by permuting the original sample and taking the first m observations of the random samples. Table 6 and Table 7 summarize the empirical coverage for parameter choices $\beta_\star = .05$, and $\beta_\star = .25$, respectively, and the median confidence interval lengths for the nominal confidence level $1-\alpha = .95$.

Insert Table 6 and Table 7 about here

In all Monte Carlo simulation settings, the empirical coverage for the robust subsampling with data driven choice of the block size are accurate and closer to the nominal coverage than those of the classical subsampling. 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 $n = 120$, $d = 3$, $\beta_\star = .25$, the median confidence interval of the robust subsampling is approximately 9% higher than the median length of the subsampling. However, in presence of contamination, the robust subsampling produces clearly a more efficient inference with dramatically smaller median confidence interval lengths. For instance, for the case $n = 120$, $d = 3$, $\beta_\star = .25$, the median confidence interval of the robust subsampling is approximately 55% (53%)

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. Therefore, for the sake of brevity, we do not report a detailed power comparison of tests based on the two methods.

Unreported results for the inconsistent bootstrap and robust bootstrap yield empirical coverages between 51.35% to 56.50%, for parameter $\beta_\star = .05$. Similarly, the subsampling and robust subsampling without hybrid correction yield empirical coverages between 48.20% and 53.15%, which are not too far away from the theoretical distorted coverage $(1-\alpha)/2$ of equal-tailed confidence intervals; see Andrews and Guggenberger, 2010b. 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., $C = 4$, $d = 20$) produce similar results.

We complete this analysis with some results also for the naive bootstrap approach based on unconstrained estimators. More precisely, we compute a nonrobust bootstrap distribution by applying the classical bootstrap approach to the unconstrained OLS estimator of β. Similarly, we compute a robust bootstrap distribution by applying our robust approach to the robust unconstrained MM-estimator of β. Finally, we construct nonrobust and robust confidence intervals for the parameter of interest by truncating at 0 the confidence intervals implied by the unconstrained nonrobust and robust bootstrap distributions, respectively. As discussed in Andrews and Guggenberger (2010b) in relation to using the unconstrained OLS estimator, when $\beta_\star$ is "sufficiently" far from the boundary restriction, confidence intervals computed in this way may imply an empirical coverage quite close to the nominal confidence level. In contrast, for $\beta_\star$ "close" to the boundary this approach tends to produce too conservative confidence intervals. In our setting, for $\beta_\star = .25$, $d = 3$, the robust bootstrap shows a desirable stability, implying an empirical coverage between 94.4% and 96.1% both in presence and absence of contamination. The results for the nonrobust bootstrap confirm the lack of robustness of this method. Indeed, for $\beta_\star = .25$, the coverage ranges between 94.10% and 97.95%, for $\gamma = 0$ and $\gamma = 0.25$, respectively. As expected the accuracy of both procedures deteriorates for $\beta_\star = .05$. In this case, the coverage is larger than 96.65% for both nonrobust and robust bootstrap methods, even without contamination.

Finally, to analyze the robustness properties of classical and robust subsampling with larger sample sizes, we also consider the case $n = 500$. Table 8 summarizes the empirical coverage for parameter choices $\beta_\star = .05, .25$, and the median confidence interval lengths for the nominal confidence level $1 - \alpha = .95$.

Insert Table 8 about here

We consider the maximal recommended choice proposed by Romano and Wolf (2001), $m_{min} = 12$ and $m_{max} = 68$, respectively. Moreover, we also apply MCIV to the interval $I = \{23, \ldots, 45\}$, which represents the minimal interval proposed by Romano and Wolf (2001). In this setting, the block sizes $m = 11$ and $m = 16$ imply a breakdown point of 50% for the .975-quantile for $d = 3$ and $d = 5$, respectively.

The Monte Carlo results confirm the findings obtained for smaller sample sizes. In particular, we observe that classical and robust subsampling perform very similarly without contamination. Also, in this setting, the presence of contamination dramatically increases the length of classical subsampling confidence intervals. For instance, for $d = 3$ and $\beta_\star = .25$, the median confidence interval of the robust subsampling is approximately 55% lower than the median length of the classical subsampling when $\gamma = .25$.

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:

$$Y_{\max} = \arg\max_{Y_1,\ldots,Y_n} \{u(Y_i) | u(Y_i) = Y_i - Z_i'\eta, \text{under } \mathcal{H}_0 : \beta_\star = 0.25\}. \tag{32}$$

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 summarizes the results for $n = 120$.

Insert Figure 1 about here

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 (31) with $C = 15$, and have analyzed the average $p-$value variation with respect to the setting with no contamination ($\gamma = 0$). Figure 2 summarizes the results.

Insert Figure 2 about here

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.

4 Concluding Remarks

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.

Appendix: Proofs

Proof of Theorem 2. The quantile $Q^*_{t,n,m}$ 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 $(X^*_1,\ldots,X^*_m)$, when np is the number of outliers in the original sample $(X_1,\ldots,X_n)$. The random variable $X(n,m,p)$ follows a hypergeometric distribution with parameters n, np, and m. Consequently, $b_{t,n,m}$ is the smallest proportion p such that $np \in$ and $P\left[X(n,m,p) < mb\right] < t$, which is the stated result. ■

Proof of Corollary 3. Existence of $\hat{m}$ is ensured by Theorem 2. For a hypergeometrically distributed variable $X(n,m,p)$ such that $np \in \mathbb{N}$, the probability $P\left[X(n,m,p) < mb\right]$ is decreasing in p. Consequently, $b_{t,n,\hat{m}} \geq \hat{b}$. By definition, for every integer $m < \hat{m}$, $P\left[X(n,m,\hat{b}-1/n) < mb\right] < t$, and $b_{t,n,m} \leq \hat{b} - 1/n$. 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, $P\left[X(n,m,p) < mb\right] < t$, where $X(n,m,p)$ is a hypergeometric random variable with parameters n, np, and m. For n large enough and c small enough, $P\left[X(n,m,p) < mb\right] > t$. Therefore, $b_{t,n,m} = b - z_t\sqrt{b(1-b)(1-r)}/\sqrt{m} + O(1/m)$, as stated. ■

Proof of Corollary 7. 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 $\{X_1,\ldots,X_n\}$ satisfies $p \geq \min\{b_t(m-k), b_t(m-k+1), .., b_t(m+k-1), b_t(m+k)\}$. This concludes the proof. ■

Proof of Corollary 9. By definition, in order to get $m_c = \infty$ we must have $P\left[Q^{**}_t(m) = \infty\right] \geq t$ for all $m \in \mathcal{M}$. $Q^{**}_t(m) = \infty$ if the number of outliers in bootstrap sample $(X^*_1,\ldots,X^*_n)$ 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 $\hat{\theta}_n = (\widehat{\beta}_n', \widehat{\sigma}_n, \tilde{\beta}_n')'$ as the fixed point of the following system of equations:

$$\begin{aligned}
\widehat{\beta}_n &= A_n(\widehat{\beta}_n, \widehat{\sigma}_n)^{-1} V_n(\widehat{\beta}_n, \widehat{\sigma}_n), \\
\widehat{\sigma}_n &= \widehat{\sigma}_n U_n(\tilde{\beta}_n, \widehat{\sigma}_n), \\
\tilde{\beta}_n &= B_n(\tilde{\beta}_n, \widehat{\sigma}_n)^{-1} W_n(\tilde{\beta}_n, \widehat{\sigma}_n),
\end{aligned} \tag{33}$$

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 (33) can be written as $\hat{\theta}_n = F_n(\hat{\theta}_n)$ for an appropriate function

$F_n : \mathbb{R}^{2d+1} \to \mathbb{R}^{2d+1}$. A first order expansion of (33) gives

$$\sqrt{n}(\hat{\theta}_n - \theta_\star) = [I - \nabla F_n(\theta_\star)]^{-1} \sqrt{n}(F_n(\theta_\star) - \theta_\star) + o_P(1), \tag{34}$$

where $\theta_\star = (\beta_\star', \sigma_\star, \tilde{\beta}_\star')'$. The explicit computation of ∇F_n shows that $\tilde{\beta}_n$ does not enter (34) in the approximation of the first $d+1$ components of $\sqrt{n}(\hat{\theta}_n - \theta_\star)$, i.e., the approximation of $\sqrt{n}(\widehat{\beta}_n - \beta_\star)$ and $\sqrt{n}(\widehat{\sigma}_n - \sigma_\star)$. The $d \times d$ matrix M_n in (17) is the left upper diagonal block of $[I - \nabla F_n(\tau_n)]^{-1}$, and the vector d_n in (18) is the $d+1-$th upper $d-$dimensional column of this matrix. Summarizing, we obtain the approximation:

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

where $M_{n,\star}$ and $d_{n,\star}$ are the same matrix and the same vector as in (17) and (18), respectively, but evaluated at $\theta_\star$ instead of $\hat{\theta}_n$. Therefore, we have to show that the limit distribution of $\xi_n(\theta_\star)$ is the same as the limit distribution of

$$\begin{aligned} \xi^*_{n,m} &= M_n\sqrt{m}(A^*_{n,m}(\widehat{\beta}_n, \widehat{\sigma}_n)^{-1} V^*_{n,m}(\widehat{\beta}_n, \widehat{\sigma}_n) - \widehat{\beta}_n) \\ &\quad + d_n\sqrt{m}(\widehat{\sigma}_n U^*_{n,m}(\tilde{\beta}_n, \widehat{\sigma}_n) - \widehat{\sigma}_n) \\ &= M_n\sqrt{m}(\widehat{\beta}^*_{n,m} - \widehat{\beta}_n) + d_n\sqrt{m}(\widehat{\sigma}^*_{n,m} - \widehat{\sigma}_n). \end{aligned}$$

To this end, it is sufficient to prove that the limit distribution of $\zeta^*_{n,m}(\hat{\theta}_n) := \sqrt{m}(F^*_{n,m}(\hat{\theta}_n) - \hat{\theta}_n)$ is the same as the limit distribution of $\zeta_n(\theta_\star) := \sqrt{n}(F_n(\theta_\star) - \theta_\star)$. In order to obtain this, we only need to show that the U-statistic defined by $U_{n,m}(x) = \frac{1}{N_{n,m}} \sum_{s=1}^{N_{n,m}} \mathbb{I}\left\{\sqrt{m}(F^*_{n,m,s}(\hat{\theta}_n) - \theta_\star) \le x\right\}$ converges to the limit cumulative distribution of $\zeta_n(\theta_\star)$, 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 Salibian-Barrera and Zamar (2002) to show that, given a subsampling block of size m, the approximation $\widehat{\beta}^*_{n,m}$ 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:

$$P\left[X(n, m, p) > m - d\right] \geq 1 - t. \tag{35}$$

This proves statement (i) of Theorem 14, after taking complements of the event in (35). Statement (ii) now follows with the same arguments used to prove Corollary 3. ■

References

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

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

[3] Andrews, D., and P. Guggenberger 2009a. Hybrid and size-corrected subsample methods. Econometrica, 77, 721-762.

[4] 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.

[5] 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.

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

[7] 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.

[8] Chernozhukov, V., and I. Fernandez-Val, 2005. Subsampling inference on quantile regression processes. Sankhya, 67, 253–276.

[9] Cowell, F. A., and E. Flachaire, 2007. Income distribution and inequality measurement: the problem of extreme values. Journal of Econometrics, 141, 1044–1072.

[10] Datta, S., 1995. On a modified bootstrap for certain asymptotically non-normal statistics. Statistics and Probability Letters, 24, 91–98.

[11] Davidson, R., and E. Flachaire, 2007. Asymptotic and bootstrap inference for inequality and poverty measures. Journal of Econometrics, 141, 141–166.

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

[13] Efron, B., 1979. Bootstrap methods: Another look at the jackknife. Annals of Statistics, 7, 1–26.

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

[15] Gonzalo, J., and M. Wolf, 2005. Subsampling inference in threshold autoregressive models. Journal of Econometrics, 127, 201–224.

[16] Hall, P., 1992. The bootstrap and Edgeworth expansion. New York: Springer-Verlag.

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

[18] Hall, P., and Q. Yao, 2003. Inference in ARCH and GARCH models with heavy–tailed errors. Econometrica, 71, 285–317.

[19] Hampel, F., Ronchetti, E., Rousseeuw, P., and W. Stahel, 1986. Robust statistics. The approach based on influence functions. Wiley, New York.

[20] Heritier, S., and E. Ronchetti 1994, Robust bounded-influence tests in general parametric model. Journal of the American Statistical Association, 89, 897-904.

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

[22] Huber, P., 1981. Robust statistics. Wiley, New York.

[23] 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.

[24] La Vecchia, D., and F. Trojani, 2010. Infinitesimal robustness for diffusions. Journal of the American Statistical Association, 105, 703–712.

[25] 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.

[26] 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.

[27] 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.

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

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

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

[31] 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.

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

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

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

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

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

[37] Salibian-Barrera, M., 2006a. The asymptotic of MM-estimators for linear regression with fixed designs. Metrika, 63, 283–294.

[38] Salibian-Barrera, M., 2006b. Bootstrapping MM-estimators for linear regression with fixed designs. Statistics & Probability Letters, 76, 1287–1297.

[39] 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.

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

[41] Salibian-Barrera, M., and R. Zamar, 2002. Bootstrapping robust estimates of regression. Annals of Statistics, 30, 556–582.

[42] Singh, K., 1998. Breakdown theory for bootstrap quantiles. Annals of Statistics, 26, 1719–1732.

[43] Wu, C.F., 1990. On the asymptotic poperties of the jacknife histogram. Annals of Statistics, 18, 1438–1452.

[44] Yohai, V, 1987. High breakdown-point and high efficiency robust estimates for regression. Annals of Statistics, 15, 642–656.

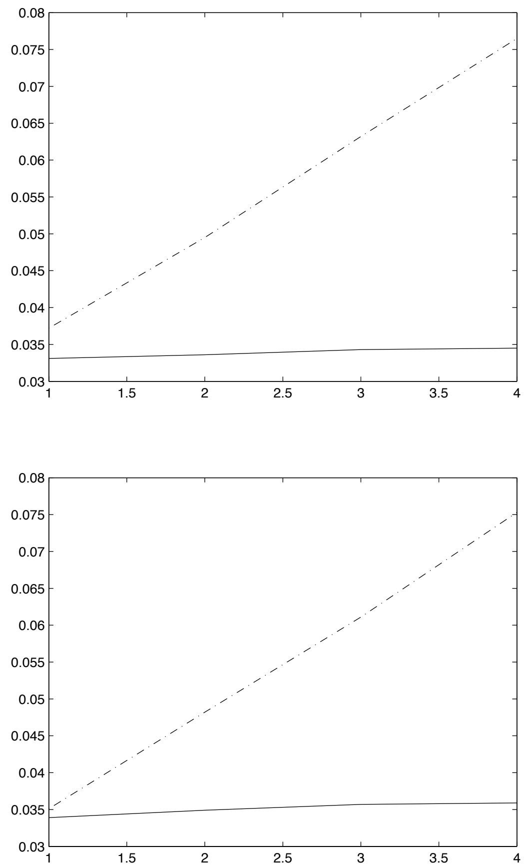

Figure 1: **Sensitivity analysis.** Sensitivity plots of the absolute variation of the empirical $p-$value average, for a test of the null hypothesis $\mathcal{H}_0 : \beta_\star = 0.25$, with respect to variations of $Y_{\max}$, in each Monte Carlo sample, within the interval $[1, 4]$. The random samples are generated under $\mathcal{H}_0$ and, from the top to the bottom, with $n = 120$, and $d = 3$, $d = 5$, respectively. We consider the classical subsampling (dash-dotted line) and the robust subsampling (straight line), using the MCIV method for the block size selection.

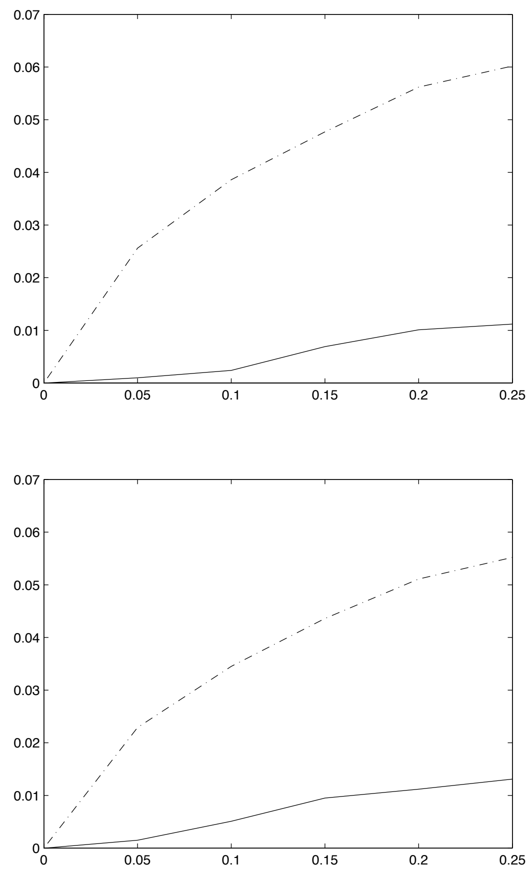

Figure 2: **Breakdown point analysis.** Sensitivity plots of the empirical $p-$value average for a test of the null hypothesis $\mathcal{H}_0 : \beta_\star = 0.25$. Each $p-$value average is computed using Monte Carlo samples generated with contamination probabilities $\gamma \in [0, 0.25]$. The graphs 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 $n = 120$, and $d = 3$, $d = 5$, respectively. 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
CB	.1750	.1500	.1250	**CB**	.4000	.3750	.3250
CS				**CS**			
$m=5$	.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
CB	.1875	.1750	.1500	**CB**	.4250	.4125	.3750
CS				**CS**			
$m=10$	.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
CB	.2000	.1917	.1667	**CB**	.4417	.4250	.3917
CS				**CS**			
$m=10$	.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: **Breakdown point of subsampling and bootstrap quantiles.** t-quantile upper breakdown point of the classical bootstrap (CB) and the classical subsampling (CS) 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) results. Subsampling breakdown points are computed using Theorem 2. The smallest subsample size such that the $t-$quantile subsampling breakdown point equals the breakdown point of statistic T for all given confidence levels is 37, 77, and 117 (when b=0.25), and 39, 79 and 119 (when b=0.5), for sample sizes 40, 80, and 120, 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 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).

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

Table 3: **Breakdown point of robust subsampling and robust bootstrap quantiles.** t-quantile breakdown point of the robust bootstrap and the robust subsampling, denoted "RB" and "RS", respectively, 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.

$n=120$, $\theta_\star=0$	$\gamma=0\%$	$\gamma=5\%$	$\gamma=10\%$
CS $m=3$	.0530(3.6728)	.0425(34.1359)	.0195(71.5518)
CS $m=6$	.0475(3.6751)	.0200(35.5514)	.0135(62.2192)
CS $m=22$	.0360(3.3840)	.0135(24.4441)	.0125(46.1002)
CS MCIV	.0385(3.5610)	.0140(26.8490)	.0130(50.7207)
SR $m=3$	.0555(3.6784)	.0070(27.4814)	.0020(71.3789)
SR $m=6$	.0485(3.6821)	.0220(6.2247)	.0140(17.0751)
SR $m=22$	.0380(3.4263)	.0330(4.3636)	.0315(5.5619)
SR MCIV	.0415(3.5236)	.0315(4.5493)	.0295(5.7227)
RS $m=3$	.0540(3.6776)	.0540(4.6349)	.0530(5.7500)
RS $m=6$	.0475(3.6763)	.0465(4.6267)	.0460(5.7412)
RS $m=22$	.0375(3.4188)	.0345(4.3154)	.0345(5.3183)
RS MCIV	.0390(3.5930)	.0370(4.5302)	.0360(5.6653)

Table 4: **Proportion of rejections of $\mathcal{H}_0 : \theta_\star = 0$, using the classical subsampling, the subsampling applied to a robust estimator, and the robust subsampling** (CS, SR, and RS, respectively). The true value of $\theta_\star$ is equal to 0 (size analysis). We report the proportion of rejection in the squared mean testing setting, for significance level $\alpha = .05$, for the sample size $n = 120$, and contamination probabilities $\gamma = 0\%, 5\%, 10\%$. We denote the block size by m. The data driven block size selection procedure based on the MCIV index is denoted by "MCIV". In brackets, we give the median of the .95-quantile of the correspondently subsampling distributions. The number of replications is 2000.

$n = 120$, $\theta_\star = 0.25^2$	$\gamma = 0\%$	$\gamma = 5\%$	$\gamma = 10\%$
CS $m = 3$	.7635(4.1476)	.1910(31.7474)	.0535(73.8258)
CS $m = 6$	.7490(4.5221)	.1375(35.6029)	.0590(62.1055)
CS $m = 22$	.6760(5.4951)	.1335(26.1528)	.0590(47.5607)
CS MCIV	.7170(5.0817)	.1350(28.1353)	.0575(51.4249)
SR $m = 3$	.7690(4.1514)	.1885(27.2040)	.0280(70.5696)
SR $m = 6$	.7500(4.5252)	.4370(7.2949)	.1970(17.9760)
SR $m = 22$	.6795(5.5407)	.5615(6.5782)	.4680(7.6559)
SR MCIV	.7260(4.9687)	.5535(6.1515)	.4595(7.7612)
RS $m = 3$	.7670(4.1536)	.6660(5.1016)	.5695(6.1756)
RS $m = 6$	.7500(4.5383)	.6420(5.5291)	.5415(6.5852)
RS $m = 22$	.6785(5.5402)	.5695(6.5089)	.4850(7.5393)
RS MCIV	.7190(5.1322)	.6015(6.0631)	.4895(7.1599)

$n = 120$, $\theta_\star = 0.5^2$	$\gamma = 0\%$	$\gamma = 5\%$	$\gamma = 10\%$
CS $m = 3$	.9995(5.3078)	.4670(32.0986)	.2000(70.4190)
CS $m = 6$	.9995(6.3159)	.4360(36.0152)	.2170(63.6958)
CS $m = 22$	.9995(8.9073)	.4740(30.7224)	.2340(52.4371)
CS MCIV	.9995(7.8590)	.4535(31.1614)	.2260(54.4384)
SR $m = 3$	.9995(5.3093)	.5230(26.2870)	.1565(70.2665)
SR $m = 6$	.9995(6.3324)	.8195(9.4694)	.5445(20.5761)
SR $m = 22$	.9995(8.9026)	.9905(10.3072)	.9630(11.8511)
SR MCIV	.9995(7.7826)	.9895(9.2190)	.9615(11.0778)
RS $m = 3$	.9995(5.3135)	.9965(6.3198)	.9850(7.4703)
RS $m = 6$	.9995(6.3356)	.9955(7.4686)	.9815(8.6930)
RS $m = 22$	.9995(8.9353)	.9935(10.2753)	.9685(11.7832)
RS MCIV	.9995(7.9353)	.9950(9.0888)	.9760(10.4156)

Table 5: **Proportion of rejections of $\mathcal{H}_0 : \theta_\star = 0$, using the classical subsampling, the subsampling applied to a robust estimator, and the robust subsampling** (CS, SR, and RS, respectively). The true value of $\theta_\star$ is equal to 0.25^2 and 0.5^2, respectively (power analysis). We report the proportion of rejection in the squared mean testing setting, for significance level $\alpha = .05$, for the sample size $n = 120$, and contamination probabilities $\gamma = 0\%, 5\%, 10\%$. We denote the block size by m. The data driven block size selection procedure based on the MCIV index is denoted by "MCIV". In brackets, we give the median of the .95-quantile of the correspondently subsampling distributions. The number of replications is 2000.

$n = 120$, $d = 3$, $\beta_\star = .05$	$\gamma = 0\%$	$\gamma = 15\%$	$\gamma = 25\%$
CS $m = 6$	.9990(0.2356)	.9995(0.4620)	.9995(0.5598)
CS $m = 11$	.9840(0.2341)	.9935(0.4637)	.9955(0.5585)
CS $m = 33$	.9465(0.2351)	.9660(0.4574)	.9675(0.5573)
CS MCIV1	.9625(0.2377)	.9750(0.4619)	.9755(0.5592)
CS MCIV2	.9615(0.2368)	.9740(0.4583)	.9745(0.5571)
RS $m = 6$	1.0000(0.2559)	1.0000(0.2650)	1.0000(0.3460)
RS $m = 11$	.9905(0.2536)	.9920(0.2616)	.9955(0.3418)
RS $m = 33$	.9490(0.2507)	.9535(0.2605)	.9625(0.3360)
RS MCIV1	.9630(0.2531)	.9635(0.2616)	.9715(0.3410)
RS MCIV2	.9615(0.2530)	.9615(0.2602)	.9685(0.3406)
$n = 120$, $d = 3$, $\beta_\star = .25$	$\gamma = 0\%$	$\gamma = 15\%$	$\gamma = 25\%$
CS $m = 6$	.9725(0.4172)	.9950(0.6365)	.9965(0.7237)
CS $m = 11$	.9560(0.3744)	.9925(0.6355)	.9930(0.7219)
CS $m = 33$	.9205(0.3338)	.9620(0.6100)	.9655(0.7126)
CS MCIV1	.9430(0.3522)	.9745(0.6213)	.9755(0.7146)
CS MCIV2	.9410(0.3517)	.9710(0.6186)	.9715(0.7151)
RS $m = 6$	.9730(0.4470)	.9745(0.4572)	.9835(0.5168)
RS $m = 11$	.9640(0.4096)	.9650(0.4321)	.9830(0.4949)
RS $m = 33$	.9245(0.3616)	.9290(0.3790)	.9490(0.4450)
RS MCIV1	.9440(0.3836)	.9445(0.3999)	.9575(0.4662)
RS MCIV2	.9425(0.3818)	.9430(0.3982)	.9550(0.4632)

Table 6: **Empirical coverage of the classical subsampling and the robust subsampling** (CS and RS, respectively). Simulated empirical coverage in the linear regression testing setting with $\beta_\star = .05, .25$, $d = 3$, for confidence levels $t = 0.95$, for the sample size $n = 120$ and contamination probabilities $\gamma = 0\%, 15\%, 25\%$. We denote the block size by m. Data driven block size selection procedures based on the MCIV are denoted by "MCIV". In particular, for "MCIV1" we consider the set $\mathcal{M}_1 = \{8, \ldots, 27\}$, while for "MCIV2" we consider the set $\mathcal{M}_2 = \{11, \ldots, 27\}$. In brackets we report the median of the confidence interval lengths. The number of replications is 2000.

$n = 120$, $d = 5$, $\beta_\star = .05$	$\gamma = 0\%$	$\gamma = 15\%$	$\gamma = 25\%$
CS $m = 8$	.9980(0.2372)	1.0000(0.4661)	1.0000(0.5689)
CS $m = 16$	.9815(0.2388)	.9920(0.4579)	.9955(0.5605)
CS $m = 33$	.9530(0.2373)	.9635(0.4571)	.9710(0.5613)
CS MCIV1	.9720(0.2367)	.9820(0.4523)	.9835(0.5567)
CS MCIV2	.9635(0.2353)	.9785(0.4510)	.9815(0.5592)
RS $m = 8$	.9985(0.2713)	1.0000(0.2814)	1.0000(0.3596)
RS $m = 16$	.9830(0.2701)	.9915(0.2735)	.9955(0.3580)
RS $m = 33$	.9545(0.2687)	.9610(0.2742)	.9705(0.3567)
RS MCIV1	.9730(0.2712)	.9775(0.2802)	.9815(0.3618)
RS MCIV2	.9640(0.2709)	.9715(0.2777)	.9785(0.3578)
$n = 120$, $d = 5$, $\beta_\star = .25$	$\gamma = 0\%$	$\gamma = 15\%$	$\gamma = 25\%$
CS $m = 8$	.9720(0.4308)	.9985(0.6309)	1.0000(0.7201)
CS $m = 16$	.9545(0.3788)	.9900(0.6308)	.9905(0.7298)
CS $m = 33$	.9255(0.3434)	.9615(0.6150)	.9625(0.7259)
CS MCIV1	.9505(0.3655)	.9755(0.6301)	.9770(0.7288)
CS MCIV2	.9420(0.3610)	.9745(0.6281)	.9760(0.7282)
RS $m = 8$	.9760(0.4569)	.9785(0.4679)	.9895(0.5405)
RS $m = 16$	.9610(0.4191)	.9635(0.4475)	.9825(0.5167)
RS $m = 33$	.9325(0.3850)	.9340(0.4021)	.9565(0.4730)
RS MCIV1	.9525(0.4042)	.9570(0.4204)	.9710(0.4952)
RS MCIV2	.9475(0.4035)	.9530(0.4167)	.9670(0.4914)

Table 7: **Empirical coverage of the classical subsampling and the robust subsampling** (CS and RS, respectively). Simulated empirical coverage in the linear regression testing setting with $\beta_\star = .05, .25$, $d = 5$, for confidence levels $t = 0.95$, for the sample size $n = 120$ and contamination probabilities $\gamma = 0\%, 15\%, 25\%$. We denote the block size by m. Data driven block size selection procedures based on the MCIV are denoted by "MCIV". In particular, for "MCIV1" we consider the set $\mathcal{M}_1 = \{10, \ldots, 27\}$, while for "MCIV2" we consider the set $\mathcal{M}_2 = \{16, \ldots, 27\}$. In brackets we report the median of the confidence interval lengths. The number of replications is 2000.

$n = 500$, $d = 3$, $\beta_\star = .05$	$\gamma = 0\%$	$\gamma = 15\%$	$\gamma = 25\%$
CS $m = 12$	.9925(0.1436)	.9935(0.2430)	.9950(0.2909)
CS $m = 68$	.9635(0.1412)	.9670(0.2417)	.9760(0.2812)
CS MCIV	.9780(0.1423)	.9780(0.2420)	.9795(0.2888)
RS $m = 12$	.9930(0.1514)	.9965(0.1558)	.9980(0.1825)
RS $m = 68$	.9640(0.1510)	.9645(0.1536)	.9690(0.1790)
RS MCIV	.9805(0.1509)	.9835(0.1543)	.9835(0.1787)
$n = 500$, $d = 3$, $\beta_\star = .25$	$\gamma = 0\%$	$\gamma = 15\%$	$\gamma = 25\%$
CS $m = 12$	.9640(0.1951)	.9685(0.4029)	.9765(0.4679)
CS $m = 68$	.9385(0.1726)	.9430(0.3596)	.9550(0.4311)
CS MCIV	.9530(0.1782)	.9560(0.3713)	.9595(0.4496)
RS $m = 12$	.9660(0.2209)	.9670(0.2365)	.9685(0.2903)
RS $m = 68$	.9375(0.1885)	.9365(0.1951)	.9380(0.2360)
RS MCIV	.9530(0.1947)	.9565(0.2029)	.9575(0.2487)
$n = 500$, $d = 5$, $\beta_\star = .05$	$\gamma = 0\%$	$\gamma = 15\%$	$\gamma = 25\%$
CS $m = 12$	.9975(0.1464)	1.0000(0.2494)	1.0000(0.2944)
CS $m = 68$	.9685(0.1454)	.9725(0.2477)	.9765(0.2861)
CS MCIV	.9785(0.1455)	.9820(0.2478)	.9840(0.2927)
RS $m = 12$	.9990(0.1559)	1.0000(0.1572)	1.0000(0.1873)
RS $m = 68$	.9715(0.1514)	.9775(0.1534)	.9785(0.1857)
RS MCIV	.9800(0.1545)	.9825(0.1552)	.9850(0.1845)
$n = 500$, $d = 5$, $\beta_\star = .25$	$\gamma = 0\%$	$\gamma = 15\%$	$\gamma = 25\%$
CS $m = 12$	.9735(0.2115)	.9740(0.4200)	.9750(0.4810)
CS $m = 68$	.9435(0.1745)	.9455(0.3617)	.9470(0.4302)
CS MCIV	.9565(0.1823)	.9590(0.3764)	.9630(0.4505)
RS $m = 12$	.9785(0.2447)	.9790(0.2764)	.9835(0.3386)
RS $m = 68$	.9505(0.1931)	.9555(0.2006)	.9580(0.2409)
RS MCIV	.9605(0.2022)	.9605(0.2109)	.9635(0.2549)

Table 8: **Empirical coverage of the classical subsampling and the robust subsampling** (CS and RS, respectively). Simulated empirical coverage in the linear regression testing setting with $\beta_\star = .05, .25$, $d = 3, 5$, for confidence levels $t = 0.95$, for the sample size $n = 500$ and contamination probabilities $\gamma = 0\%, 15\%, 25\%$. We denote the block size by m. Data driven block size selection procedures based on the MCIV are denoted by "MCIV". In brackets we report the median of the confidence interval lengths. The number of replications is 2000.