J. R. Statist. Soc. B (2003)
65, *Part* 1, *pp.* 81–94

Comprehensive definitions of breakdown points for independent and dependent observations

Marc G. Genton

North Carolina State University, Raleigh, USA

and André Lucas

Vrije Universiteit, Amsterdam, the Netherlands

[Received October 2000. Final revision July 2002]

Summary. We provide a new definition of breakdown in finite samples, with an extension to asymptotic breakdown. Previous definitions centre on defining a critical region for either the parameter or the objective function. If for a particular outlier configuration the critical region is entered, breakdown is said to occur. In contrast with the traditional approach, we leave the definition of the critical region implicit. Our proposal encompasses previous definitions of breakdown in linear and non-linear regression settings. In some cases, it leads to a different and more intuitive notion of breakdown than other procedures that are available. An important advantage of our new definition is that it also applies to models for dependent observations where current definitions of breakdown typically fail. We illustrate our suggestion by using examples from linear and non-linear regression, and time series.

Keywords: Bias curve; Linear regression; Non-linear regression; Outliers; Statistical robustness; Time series

1. Introduction

Qualitative robustness and especially the definition of breakdown have evolved considerably over the last three decades. Informally, a breakdown of a statistical procedure means that the procedure no longer conveys useful information on the data-generating mechanism. Available definitions mainly differ in their formalization of what is meant by 'no longer conveys useful information'. Hampel (1971) operationalized the breakdown point in the setting of parameter estimation as the fraction of contamination (or outliers) that suffices to drive the estimator beyond all bounds (or to the edge of the parameter space if the latter is bounded). In that case the estimator no longer bears any resemblance to the true parameter value. Since the original introduction of the concepts of breakdown and the breakdown point by Hampel (1971), the breakdown point has been extended to finite samples (Donoho and Huber, 1983), dependent observations (Martin and De Jong, 1977; Martin, 1980; Genton, 1998; Ma and Genton, 2000), test statistics (He *et al.*, 1990; He, 1991) and non-linear regression models (Stromberg and Ruppert, 1992; Sakata and White, 1995, 1998). Especially Stromberg and Ruppert (1992) and Sakata and White (1995) convincingly argued that the bias of the parameter estimators as originally proposed by Hampel is not always a good criterion to assess the breakdown of an estimator. Instead, Stromberg and Ruppert proposed to consider the fraction of contamination

Address for correspondence: M. G. Genton, Department of Statistics, North Carolina State University, Box 8203, Raleigh, NC 27695-8203, USA.
E-mail: genton@stat.ncsu.edu

 1369–7412/03/65081

that drives at least one of the fitted values to its supremum or infimum. Sakata and White argued that the fitted value may sometimes not be a satisfactory criterion either and proposed several alternative criterion functions to assess breakdown.

Although these alternative definitions cover a wide range of models and estimators, we can easily construct examples that are not covered by them. Consider the ordinary least squares estimator and the autoregressive time series model of order 1, AR(1),

$$Y_i = \phi Y_{i-1} + e_i, \qquad i = 1, \ldots, n, \tag{1}$$

with bounded parameter space $\phi \in (-1, 1)$ and with e_i independent and identically distributed innovations. Suppose that Y_i is observed with error as $\tilde{Y}_i = Y_i + Z_i$, where $Z_i = \zeta$ when $i = i_0$ for a single $i_0 \in \{2, \ldots, n\}$, and $Z_i = 0$ otherwise. The ordinary least squares estimator of ϕ based on the contaminated sample $\tilde{Y}_1, \ldots, \tilde{Y}_n$ is

$$\hat{\phi} = \frac{\sum_{i=2}^{n} \tilde{Y}_i \tilde{Y}_{i-1}}{\sum_{i=2}^{n} \tilde{Y}_{i-1}^2} = \frac{\zeta(Y_{i_0-1} + Y_{i_0+1}) + \sum_{i=2}^{n} Y_i Y_{i-1}}{\zeta^2 + 2\zeta Y_{i_0} + \sum_{i=2}^{n} Y_{i-1}^2}. \tag{2}$$

Clearly, as $\zeta \to \infty$, $\hat{\phi} \to 0$, whatever the realization of the uncontaminated sample $Y_1, \ldots, Y_n$. The ordinary least squares estimator then no longer conveys any useful information on ϕ and the estimate is totally determined by the contamination. In this simple time series model, therefore, the ordinary least squares estimator breaks down to 0 with one extreme outlier. Note that $\phi = 0$ is at the *centre* of the parameter space. This form of breakdown typically rules out the classical definition of Hampel (1971), because the estimator does not diverge to either ∞ or to the edge of the parameter space. This simple example does not fit the more recent definitions of breakdown either. In particular, following the definition of He and Simpson (1992, 1993), breakdown occurs if the supremum bias is reached. Here, however, the supremum bias is reached when $\hat{\phi}$ tends to 1 or -1, depending on whether ϕ is negative or positive respectively. Alternatively, Stromberg and Ruppert (1992) and also Sakata and White (1995, 1998) defined breakdown as the point where the model's fit, $\hat{\phi} Y_{i-1}$, or some other criterion function, tends to either its supremum or its infimum for some observation in the sample. Clearly, this would again induce breakdown to ± 1 given the restricted parameter space, and not breakdown to 0.

Given the drawbacks of the previous definitions available, we introduce a new concept of breakdown. All previous definitions make explicit use of a criterion function combined with a critical region. For example, Hampel's (1971) original definition uses the absolute bias as the criterion function and ∞ (or the edge of the parameter space) as the critical region. If the criterion function (bias) enters the critical region (∞) for a certain fraction of outliers, breakdown is said to have occurred. Following Sakata and White (1995), we consider a general model measure of badness as our criterion function. This encompasses the definitions of Hampel (1971) as well as Stromberg and Ruppert (1992). In contrast with previous work, however, we leave the definition of the critical region implicit.

The intuition for our new definition is as follows. A crucial property of an estimator is that it takes different values for different sample realizations. If a continuum of sample realizations is possible and the estimator is continuous in the sample, we expect a continuum of possible values for the estimator. For example, if we consider $\mathbb{R}^n$ as the sample space, the sample mean can take any value in $\mathbb{R}$, i.e. a continuum of values. We now look for the fraction of contamination for which this property is lost. In particular, we look for the number of outliers (and corresponding outlier configuration) such that the estimator, or more specifically the measure of badness, can

take only a *finite* number of different values despite a continuum of possible uncontaminated sample realizations. If this happens, we conclude that the estimator no longer conveys useful information on the true parameters, as it can no longer distinguish clearly between alternative uncontaminated samples. In the same example of the sample mean as before, place one outlier at ∞. Then the sample mean can only take the value of ∞, even though the sample space is the continuum $\mathbb{R}^n$. As the uncontaminated observations no longer affect the estimator, we conclude that the estimator has been broken.

Our approach accommodates all of the earlier definitions of breakdown and, in several cases, leads to a more intuitive notion of breakdown. In addition, we also cover situations of breakdown that are not covered by the earlier definitions. Specifically, we provide a *unified* definition of breakdown that is applicable to settings with either independent or dependent observations. Previous definitions typically relate to one of these settings only.

We illustrate the main issues with examples from linear and non-linear regression as well as time series. In regular cases, the breakdown point according to our definition coincides with its conventional counterpart. We therefore focus on cases where our definition of breakdown gives a breakdown point that is different from that from the available definitions. We provide a typical example in the non-linear regression context, confronting our breakdown point with that of Stromberg and Ruppert (1992). The estimator may already have lost all its informational value on the uncontaminated sample, while still giving rise to a bounded model fit over a prespecified domain of interest. This implies that the estimator may already have broken in our definition, while still being unbroken in the definition of Stromberg and Ruppert.

The remainder of the paper is set up as follows. In Section 2 we introduce the basic notation and our new definitions of breakdown for finite samples as well as for the asymptotic case. We provide various illustrative examples of breakdown point computations for location in Section 3, for linear and non-linear regression in Section 4 and for time series in Section 5. We conclude in Section 6.

2. Definition of breakdown

We consider a sample of size n, $Y = (Y_1, \ldots, Y_n)$, consisting of either independent or dependent observations. The estimator of the K-dimensional parameter vector θ is defined as a function of the sample, $\hat{\theta} = \hat{\theta}(Y)$. We represent the contaminated sample by

$$Y + Z_m^\zeta, \tag{3}$$

where the outliers are introduced through a contaminating sample Z_m^ζ. When the observations are independent, m will be the number of non-zero components of Z_m^ζ, but with dependent observations the number of non-zero components of Z_m^ζ may or may not be equal to m, depending on the context (see the discussion in the next paragraph). The index $\zeta \in \bar{\mathbb{R}}$, an extended real number, indicates the magnitude of the outliers. In the case of multiple outliers ($m > 1$), these need not be identical though they are characterized by a single index ζ. A simple example is outliers $c_1\zeta, \ldots, c_m\zeta$ for fixed finite constants $c_1, \ldots, c_m$.

For instance, in the context of estimating the location of an independent and identically distributed sample Y, the contaminated sample $Y + Z_m^\zeta$ is such that Z_m^ζ typically contains $n - m$ zeros and m non-zeros. This concept of a contaminated sample is inspired from the measurement error literature (see for example Stefanski (2000)) and has the natural property of reducing to the uncontaminated sample Y when $\zeta = 0$. By using the concept of a contaminating sample rather than individual outliers, we can also consider outlier patterns that exhibit more structure than

in the regression setting. For example, in a time series setting we distinguish between additive outliers, replacement outliers and innovation outliers. The first two can be considered as pure measurement errors, whereas innovation outliers are exceptional shocks that satisfy the feed-through mechanism of the time series process. For instance, consider a simple autoregression as in model (1). An additive outlier can be studied by specifying

$$Z_1^{\zeta} = (0, \ldots, 0, \zeta, 0, \ldots, 0),$$

for some $\zeta \in \bar{\mathbb{R}}$, whereas an innovation outlier corresponds to

$$Z_1^{\zeta *} = (0, \ldots, 0, \zeta, \phi\zeta, \phi^2\zeta, \ldots).$$

In the latter case, a single ($m = 1$) outlier ζ contaminates all subsequent observations through the autoregressive mechanism. We define $\mathcal{Z}_m^{\zeta}$ as the set of allowable outlier constellations. For example, for $\phi \neq 0$ the set of allowable innovation outliers typically includes $Z_1^{\zeta *}$, but not Z_1^{ζ} as defined above, since Z_1^{ζ} cannot represent an outlier in the innovations of the autoregressive mechanism.

As mentioned in Section 1, defining breakdown in terms of bias in the estimator is not always satisfactory. Especially if the model is non-linear, criteria other than parameter stability may be more relevant. See also the arguments raised in Stromberg and Ruppert (1992) and Sakata and White (1995). Following Sakata and White (1995), therefore, we introduce a measure of badness

$$R(\theta, Y) \in \bar{\mathbb{R}}^{+}.$$

The measure of badness may equal some measure of model fit, a distance $|\hat{\theta} - \theta|$ between the estimator and the parameter, the likelihood value, and so on. Its precise interpretation depends on the context at hand.

Sakata and White (1995) defined breakdown as the fraction of contamination that drives the measure of badness to its supremum. As the measure of badness may coincide with either the distance $|\hat{\theta} - \theta|$ or the (norm of the) estimator $\hat{\theta}$ itself, this encompasses Donoho and Huber's (1983) definition of breakdown. The main disadvantage of Sakata and White's definition is that estimators may already have broken down before the measure of badness reaches its supremum. An example was given in Section 1 for the ordinary least squares estimator of the AR(1) parameter, where badness is taken as the distance $|\hat{\theta} - \theta|$. The failure of the definition of Sakata and White is due to the fact that the breakdown region is made explicit, in particular, the *supremum* badness. In our definition, we leave the breakdown region implicit.

We also introduce a neighbourhood $\mathcal{Y}$ of the uncontaminated sample Y. The set $\mathcal{Y}$ gives the set of allowable alternative realizations of the uncontaminated sample. For example, in the case of location estimation, the uncontaminated sample might be any $Y \in \mathbb{R}^n$, and therefore $\mathcal{Y} = \mathbb{R}^n$. We assume that $\mathcal{Y}$ is non-discrete. As mentioned in Section 1, a crucial property of an estimator is that it takes different values for different elements of $\mathcal{Y}$. Consequently, different elements of $\mathcal{Y}$ also correspond to different realizations of the measure of badness. Given a continuum of possible samples, i.e. a non-degenerate $\mathcal{Y}$, we therefore expect a continuum of possible values of badness. To formalize this, we define the badness set

$$R^*(\hat{\theta}, Y, Z_m^{\zeta}, \mathcal{Y}) = \bigcup_{Y' \in \mathcal{Y}} \{R(\hat{\theta}(Y' + Z_m^{\zeta}), Y)\}. \tag{4}$$

The badness set gives the range of possible values of badness for alternative realizations Y' of the uncontaminated sample. In the same location example as above with badness $R(\theta, Y)$ equal to the distance $|\hat{\theta}(Y) - \theta|$, and no outliers ($m = 0$ or $\zeta = 0$), we have for any sensible estimator $\hat{\theta}$

that $R^*(\hat{\theta}, Y, Z_0^0, \mathbb{R}^n)$ equals $\mathbb{R}^+$: alternative uncontaminated samples lead to alternative values of badness. As all samples in $\mathbb{R}^n$ are possible, badness can take any value in $\mathbb{R}^+$.

The breakdown point is now given by the fraction of contamination at which the badness set collapses from an uncountable to a finite set. The estimator can then no longer clearly distinguish between all possible alternative uncontaminated samples. For example, if $\hat{\theta}(Y)$ is the sample mean and $\mathcal{Y}$ equals $\mathbb{R}^n$, take Z_1^ζ as the standard isolated outlier configuration with one entry ζ and the remaining entries equal to 0. For every $Y' \in \mathbb{R}^n$ we then have

$$R(\hat{\theta}(Y' + Z_1^\infty), Y) = |\hat{\theta}(Y' + Z_1^\infty) - \hat{\theta}(Y)| = \infty,$$

such that $R^*(\hat{\theta}, Y, Z_1^\infty, \mathbb{R}^n)$ collapses to the singleton (or finite set) $\{\infty\}$ for every $Y \in \mathbb{R}^n$. This is formalized in the following definition.

Definition 1. The breakdown point $\varepsilon_n(\hat{\theta}, R, \mathcal{Z}_m^\zeta, \mathcal{Y})$ of the estimator $\hat{\theta}$ of θ under the measure of badness R for the set of allowable outlier constellations $\mathcal{Z}_m^\zeta$ and alternative uncontaminated samples $\mathcal{Y}$ is

$$\varepsilon_n(\hat{\theta}, R, \mathcal{Z}_m^\zeta, \mathcal{Y}) = \min \left\{ \frac{m-1}{n} \,\middle|\, \exists Z_m^\zeta \in \mathcal{Z}_m^\zeta \text{ and } \zeta \in \bar{\mathbb{R}} \text{ such that } R^*(\hat{\theta}, Y, Z_m^\zeta, \mathcal{Y}) \text{ is a finite set for every } Y \in \mathcal{Y} \right\}.$$

Alternatively, we could opt for a stricter definition by considering the fraction of contamination for which the badness set collapses to a singleton. The least median of squares for the AR(1) example in Section 5, however, shows that the estimator may become virtually useless long before that point. The breakdown point ε_n as introduced in definition 1 depends significantly on the estimator $\hat{\theta}$, on the measure of badness R and, in the context of dependent observations, on the set of allowable outlier constellations $\mathcal{Z}_m^\zeta$. It also depends on the set of alternative uncontaminated samples $\mathcal{Y}$, which is shown in Sections 4 and 5, and the example below. The breakdown point does not depend on the sample Y, as breakdown only occurs if *one* particular outlier configuration makes the badness set collapse for *every* allowable sample. This makes our definition different from the usual definition (Hampel *et al.*, 1986). For example, consider the univariate sample $Y_i = (i-2)\pi$ for $i = 1, 2, 3$ and the (redescending) M-estimator of location

$$\hat{\theta} = \arg\min_\theta \left[\sum_{i=1}^{3} \{1 - \cos(Y_i) \mathbf{1}_{\{|Y_i - \theta| < \pi\}}\} \right], \tag{5}$$

with $\mathbf{1}_A$ the indicator function for the set A. Because of the wide sample dispersion and the flat objective function outside $[-\pi, \pi]$, only one outlier is needed to drive one of the solutions of equation (5) to ∞, such that the estimator is broken in the Hampel or Donoho and Huber sense. This is not so much a property of the estimator as a property of the sample. For, if the sample were $Y_1 = Y_2 = Y_3 = 0$, two (instead of one) outliers would be needed to break the estimator. This type of sample dependence is therefore absent from our new definition. See also Zhang and Li (1998) for a more detailed analysis of the sampling properties of breakdown points. The breakdown point does depend on $\mathcal{Y}$ here as, for $\mathcal{Y} = \mathbb{R}^3$, the breakdown point is $\frac{1}{3}$, whereas for $\mathcal{Y} = \cup_{y \in \mathbb{R}} \{(y, y + \pi, y + 2\pi)\}$ the breakdown point is 0.

Note that $\zeta = \infty$ need not imply that the outliers equal (plus or minus) ∞. For example, in the context of scale estimation it can be worthwhile to consider outliers for which the non-zero elements of Z_m^ζ are either ζ or $\zeta^{-1} - Y_i$. These alternative types of outlier can be used to check for explosion ($\zeta = \infty$) or implosion ($1/\zeta = 0$) of the scale estimator respectively. Note also that

we leave the absolute magnitude of the badness as well as the relative magnitude with respect to the supremum badness implicit in our definition. The estimator may already have broken down according to our definition, whereas the badness is still below its supremum (and may increase further on adding additional outliers). We provide examples in Sections 4 and 5.

To extend our definition to the asymptotic case, we introduce the stochastic processes $Y = \{Y_i\}_{i=1}^{\infty}$ and $Z_p^{\zeta} = \{Z_i^{\zeta}\}_{i=1}^{\infty}$. The subscript p for the contaminating process indicates the fraction or probability of contamination. For example, for regression outliers Z_p^{ζ} typically is an independent and identically distributed process with $P(Z_i^{\zeta} = \zeta) = 1 - P(Z_i^{\zeta} = 0) = p$. The measure of badness is now a functional $R(\theta, Y)$, and the badness set is

$$R^*(\hat{\theta}, Y, Z_p^{\zeta}, \mathcal{Y}) = \bigcup_{Y' \in \mathcal{Y}} \{R(\hat{\theta}(Y' + Z_p^{\zeta}), Y)\}, \tag{6}$$

with $\mathcal{Y}$ the set of alternative uncontaminated processes allowed.

Definition 2. The asymptotic breakdown point $\varepsilon(\hat{\theta}, R, \mathcal{Z}_p^{\zeta}, \mathcal{Y})$ of the estimator $\hat{\theta}$ of θ under the measure of badness R for the set of allowable contaminating processes $\mathcal{Z}_p^{\zeta}$ and alternative uncontaminated processes $\mathcal{Y}$ is

$$\varepsilon(\hat{\theta}, R, \mathcal{Z}_p^{\zeta}, \mathcal{Y}) = \inf\{p |\ \exists Z_p^{\zeta} \in \mathcal{Z}_p^{\zeta} \text{ and } \zeta \in \bar{\mathbb{R}} \text{ such that } R^*(\hat{\theta}, Y, Z_p^{\zeta}, \mathcal{Y}) \text{ is a finite set for every } Y \in \mathcal{Y}\}.$$

3. Examples: location

In this section we present some simple illustrative examples of breakdown point computation for location models. In this setting our definition resembles that of Hampel (1971) or Donoho and Huber (1983) if the badness equals the distance $|\hat{\theta} - \theta|$. Therefore we can derive the breakdown point of large classes of estimators, e.g. M-estimators, with the usual techniques.

For example, consider the mean with one ($m = 1$) extreme outlier at ∞, $\mathcal{Y} = \mathbb{R}^n$ and the set $\mathcal{Z}_1^{\infty}$ of additive outlier contaminations. Denote the outlier by Z_1^{∞}. If badness $R(\theta, Y)$ is defined as $|\hat{\theta}(Y) - \theta|$, then $R^*(\hat{\theta}, Y, Z_1^{\infty}, \mathbb{R}^n)$ equals $\{\infty\}$, which is a finite set, for every $Y \in \mathcal{Y}$. Therefore, the breakdown point of the mean is given by $\varepsilon_n\{\hat{\theta}, |\hat{\theta}(Y) - \theta|, \mathcal{Z}_1^{\infty}, \mathbb{R}^n\} = (1 - 1)/n = 0$. Similarly, for the class of non-redescending M-estimators, we can use the conventional Donoho and Huber (1983) breakdown point to obtain a badness set of $\{\infty\}$, such that the breakdown point of Donoho and Huber coincides with ours. Introduce $\lfloor x \rfloor$ as the integer part of x. For redescending M-estimators, we can use the line of reasoning from Section 2 to show that the breakdown point is $\lfloor (n - 1)/2 \rfloor / n$, rather than a sample-dependent value; compare Zhang and Li (1998). This is most easily seen by noting that in our definition the estimator also must break for the (alternative) sample $Y = (0, \ldots, 0)$, which clearly requires $\lfloor (n + 1)/2 \rfloor$ outliers to drive the estimator to ∞.

To illustrate our definition in the asymptotic case, consider the mean as a location estimator in the independent and identically distributed setting,

$$\hat{\theta}(Y) = \int Y_i \, P(\mathrm{d}Y_i). \tag{7}$$

As a measure of badness, consider the distance $|\hat{\theta} - \theta|$. This should give us the Hampel (1971) breakdown point of the mean. Let $\mathcal{Y}$ denote the set of processes with finite mean. For the independent and identically distributed contaminating process $P(Z_i^{\zeta} = \zeta) = 1 - P(Z_i^{\zeta} = 0) = p$ with $\zeta = \infty$, it is then easy to show that, for every positive p arbitrarily close to 0, the badness set R^* equals $\{\infty\}$, which is a finite set. Consequently, the asymptotic breakdown point of the

mean is 0. A similar line of argument can be used to show that the median has an asymptotic breakdown point of $\frac{1}{2}$.

4. Examples: linear and non-linear regression

In this section, we provide examples of breakdown point computations in the linear and non-linear regression setting. As in the previous sections, in most regular cases we come up with the same breakdown point as traditional definitions for a wide range of estimators, e.g. GM-estimators. Ways to compute the breakdown point of these estimators are well documented elsewhere. In the present section, we therefore focus on the settings where our definition produces a breakdown point that is different from that proposed by Stromberg and Ruppert (1992) or Sakata and White (1995, 1998). Apart from the examples below, we can also think of redescending GM-estimators for regression as an extension of the results that were presented for location in the previous section.

4.1. Ordinary least squares estimator

Consider the independent and identically distributed regression model

$$Y_i = X_i\beta + e_i, \tag{8}$$

with $Y = ((Y_1, X_1), \ldots, (Y_n, X_n))$. Define badness as $R(\theta, Y) = |\hat{\theta}(Y) - \theta|$, with $\theta = \beta$. Also let $\mathcal{Z}_m^\zeta \in \mathbb{R}^{n\times 2}$ denote contaminating samples (of magnitude ζ) with $n - m$ pairs of zeros $(0, 0)$ and m pairs of non-zeros. Any reasonable estimator for β will take a continuum of values depending on the sample Y. Therefore, the only way to obtain a finite badness set R^* is to place outliers at ∞. The Donoho and Huber (1983) definition of the breakdown point looks for the fraction of contamination such that one extreme outlier configuration can be found for which $\|\beta\| = \infty$. For the ordinary least squares estimator it suffices to consider Z_1^∞ for breakdown in the Donoho and Huber sense. If the sample space is $\mathcal{Y} = \mathbb{R}^{n\times 2}$, the same outlier configuration makes the badness set $R^*(\hat{\theta}, Y, Z_1^\infty, \mathbb{R}^{n\times 2})$ collapse to the finite set $\{\infty\}$ for every $Y \in \mathcal{Y}$. So in this case the Donoho and Huber breakdown point provides an upper bound for breakdown as defined in Section 2. That the bound may be strict can be illustrated if Y is the realization of a qualitative response model $Y_i = F(X_i\beta)$ for some distribution function $F(\cdot)$. If we are interested in estimating a linear regression based on Y instead of the correct qualitative response model, realistic outlier configurations restrict the value of Y_i between 0 and 1, even if outliers are added. This implies that the ordinary least squares regression estimator can no longer diverge to ∞, rendering the Donoho and Huber definition inapplicable. By adding a single outlier of the form $(0, \zeta)$ with $\zeta = \infty$, the contaminated sample satisfies the natural bounds on the dependent variable. The ordinary least squares estimator then collapses to $\{0\}$, whatever the value of $Y \in [0, 1] \times \mathbb{R}$. So the ordinary least squares estimator still has a zero breakdown point in this setting according to the definition in Section 2. By contrast, the Donoho and Huber definition results in a breakdown point of 1.

4.2. Non-linear least squares estimator

The next example treats the non-linear regression setting. It further illustrates the crucial differences between our definition and the definition of breakdown of Stromberg and Ruppert (1992). In particular, using our new definition, an estimator has broken if variations in the uncontaminated observations no longer have an effect on the estimator. This may happen even

if the model's fit is still below its supremum or above its infimum, which is needed for breakdown according to Stromberg and Ruppert. In our view, it is more natural to say that the estimator has been broken if its value is totally dictated by the outliers in the sample, even if the maximum bias has not yet been reached. As the definition of Sakata and White (1995) is very similar to that of Stromberg and Ruppert, similar comments apply to it.

Consider the Michaelis–Menten model

$$Y_i = \frac{VX_i}{K + X_i} + e_i = \frac{\alpha K X_i}{K + X_i} + e_i,$$

where $\alpha = V/K$, $V, K > 0$, and $X_i > 0$. For simplicity, we assume that our region of interest for X is bounded; in particular $X \in [0, 3]$. It is easy to see that the functional form of the Michaelis–Menten model is non-decreasing. The main point of Stromberg and Ruppert (1992) in their discussion of this model is that, if outliers are such that the estimator for K diverges while that for α remains constant, the estimator is broken in the Donoho and Huber (1983) sense. The model's fit (over the range mentioned), however, is still bounded and tends to $\hat{\alpha}X$. This led Stromberg and Ruppert to their alternative definition of breakdown. We restrict attention to their type of outliers and show that the definition of Stromberg and Ruppert and our definition of breakdown do not coincide. We define $\mathcal{Z}_m^{\zeta}$ by considering outliers of the form $Z_{i,m}^{\zeta} = (\beta\zeta, \zeta)$ for the non-zero elements in Z_m^{ζ}, where β is a fixed bounded constant and $Z_{i,m}^{\zeta}$ is the ith element of Z_m^{ζ}. So the outliers lie on a specific ray from the origin. We consider the non-linear least squares estimator and let ζ tend to ∞. Clearly, if there is one extreme outlier only, the least squares regression curve must pass through it, implying that

$$Y_j + \beta\zeta = \frac{\hat{\alpha}\hat{K}(X_j + \zeta)}{\hat{K} + X_j + \zeta} \Leftrightarrow \hat{\alpha} = \frac{(Y_j + \beta\zeta)(\hat{K} + X_j + \zeta)}{(X_j + \zeta)\hat{K}}, \tag{9}$$

for the outlier pair $(Y_j + \beta\zeta, X_j + \zeta)$. Using this, the least squares objective function becomes

$$\min_{\hat{K}} \left[\sum_{i \neq j} \left\{ Y_i - \frac{(Y_j + \beta\zeta)(\hat{K} + X_j + \zeta)\hat{K}X_i}{(X_j + \zeta)\hat{K}(\hat{K} + X_i)} \right\}^2 \right]. \tag{10}$$

As the numerator of the second term within brackets is of order ζ^2, whereas the denominator is of order ζ, it is clear that the optimal $\hat{K}$ satisfies $\hat{K} = \rho\zeta^{\kappa} + o_p(\zeta^{\kappa})$ for some $\rho > 0$ and $\kappa \geqslant 1$ when $\zeta \to \infty$. If $\kappa > 1$, $\hat{\alpha} = \beta + O(\zeta^{-1})$. Similar derivations can be followed if there are two outliers, i.e. $Z_{i,m}^{\zeta} = (\beta\zeta_1, \zeta_1)$ for $i = j_1$ and $Z_{i,m}^{\zeta} = (\beta\zeta_2, \zeta_2)$ for $i = j_2$. By letting $\zeta_1, \zeta_2 \to \infty$ and $\zeta_1/\zeta_2 \to c$ with $1 > c \geqslant 0$, we easily derive that the Michaelis–Menten curve collapses to a straight line through the origin with slope coefficient β, i.e. $\hat{\alpha} = \beta$ and $\hat{K}/\max(\zeta_1, \zeta_2) \to 0$. This also holds if there are more than two outliers of the type above. If we measure badness by the fit at a specific $X \in [0, 3]$, we see that there is an outlier configuration with two outliers such that $R^*(\hat{\alpha}, Y, Z_2^{\infty}, \mathcal{Y}) = \{3\beta X\}$, where $\mathcal{Y} = \mathbb{R}^{n \times 2}$, for every uncontaminated sample $Y \in \mathcal{Y}$. Therefore, an upper bound on the breakdown point of the non-linear least squares estimator is $1/n$. For fixed β the fit at $X \in [0, 3]$ always remains bounded. Thus, the least squares estimator has not broken in the sense of Stromberg and Ruppert (1992). Indeed, their definition leads to a breakdown point of 1 under outlier contamination of this type. Of course, if unrestricted outlier configurations are allowed, we can let β diverge to ∞ and obtain a zero breakdown point under both the definition of Stromberg and Ruppert and our definition.

To answer whether the breakdown point according to the definition of Section 2 is $1/n$ or 0, substitute $\hat{K} = \rho\zeta^{\kappa}$ into expression (10) to obtain

$$\min_{\rho>0,\kappa\geqslant 1}\left[\sum_{i\neq j}\left\{Y_i-\frac{\beta\rho\zeta^{1+\kappa}(\rho\zeta^{\kappa}+\zeta)X_i}{\rho^2\zeta^{1+2\kappa}}\right\}^2\right]+O(\zeta^{-2}). \tag{11}$$

For $\kappa > 1$, the value of ρ is immaterial and $\hat{\alpha} = \beta$. For $\kappa = 1$, we can rewrite expression (11) as

$$\min_{\theta\geqslant 1}\left\{\sum_{i\neq j}(Y_i-\theta\beta X_i)^2\right\}+O(\zeta^{-2}), \tag{12}$$

with $\theta = (1+\rho)/\rho$. We can interpret the boundary solution $\theta = 1$ as indicating $\kappa > 1$, such that expression (12) is the only relevant case to look at. Given that β is fixed, the unrestricted optimal $\hat{\theta}$ is the ordinary least squares estimator based on the uncontaminated (except the jth) observations divided by β. If the uncontaminated sample can be any vector $Y \in \mathbb{R}^n$, the estimate of θ (and therefore of ρ) varies over $\mathcal{Y} = \mathbb{R}^n$. From result (9) we have $\hat{\alpha} = \beta\hat{\theta}$, such that also the badness as measured by the fit at a specific $X \in [0, 3]$, i.e. $\hat{\alpha}X$, varies over $\mathcal{Y}$. Therefore, the estimator has not broken down with one outlier only, and the breakdown point of the non-linear least squares estimator for this model is given by $1/n > 0$. This contrasts with the breakdown points that we find with the Stromberg and Ruppert (1992) definition, i.e. 0 (for $\beta \to \infty$) or 1 (for fixed β).

5. Examples: time series

So far, we have concentrated on independent data. We now show that our new definition also applies directly to the time series setting. There, it gives rise to some remarkable results. Martin and De Jong (1977) and Martin (1980) also defined breakdown for the (autoregressive) time series setting. In particular, they considered breakdown of the autoregressive parameter towards 0 or to ± 1 as the relevant notions of breakdown. Although we completely agree with their analysis, their approach is rather focused on the time series setting and only discusses asymptotic robustness. Our definition, by contrast, naturally accommodates both the regression setting and the time series setting, whereas theirs does not. Moreover, we also provide a definition of breakdown that can be used in finite samples. As our definition encompasses that of Martin and De Jong (1977), their results can be used to compute the breakdown point of (non-redescending) GM-estimators. We focus instead on high breakdown estimators.

Other previous definitions of breakdown points for dependent observations have been studied by Genton (1998) for spatial statistics and Ma and Genton (2000) in the context of time series. Both approaches extend the traditional breakdown point to a spatial and temporal breakdown point for variogram and autocovariance estimators respectively. However, both definitions are local in the sense that they describe breakdown only for a fixed lag vector in space or time. It is also worth mentioning Boente *et al.* (1987) on qualitative robustness for dependent observations. They considered the continuity of estimators under outlier contamination in an asymptotic context. In particular, discontinuity of the estimator is regarded as a signal that the estimator is not qualitatively robust. Their concept of continuity of the estimator is, however, difficult to implement in finite samples.

In this section, we provide three examples of the computation of breakdown points for time series examples, where the crucial differences between our definition and the previous ones are most noticeable. In particular, we present the derivations for the least squares, the deepest regression and the least median of squares estimators. In the asymptotic context, neither the least median of squares nor the deepest regression estimators have explicit expressions. The estimators can be seen as the solutions to estimating equations. It is known that in simple regression

the least median of squares and deepest regression both have a breakdown point of $\frac{1}{2}$. In the time series setting, however, deepest regression and least median of squares behave markedly differently. The least median of squares estimator has a breakdown point that is far below the familiar $\frac{1}{2}$, whereas the breakdown point of the deepest regression estimator remains at $\frac{1}{2}$. Such an assessment of breakdown properties in time series was impossible with previous definitions.

5.1. Ordinary least squares estimator

Consider the AR(1) model (1), with one additive outlier. In Section 1 we argued that the ordinary least squares estimator here tends to 0 for extreme outliers ($\zeta = \infty$), independent of the sample. Using our definition, badness equal to $|\hat{\phi} - \phi|$ and the set $\mathcal{Z}_m^\zeta$ of additive outlier contaminations, the breakdown point of the ordinary least squares estimator in this setting equals 0 as we have $R^*(\hat{\phi}, Y, Z_1^\infty, \mathcal{Y}) = \{|\hat{\phi}(Y) - 0|\}$, which is a finite set, for every $Y \in \mathcal{Y} = \mathbb{R}^n$. In contrast, consider the outlier configuration $Z_m^\zeta = (0, \ldots, 0, \zeta, \zeta, \ldots, \zeta, 0, \ldots, 0)$, where the patch of ζs has length m. If badness is $|\hat{\phi} - \phi|$, it is easy to check that this results in a badness of $|\hat{\phi}(Y) - (m-1)/m|$ for ζ diverging to ∞. Similarly, if we have an alternating outlier patch $\zeta, -\zeta, \zeta, \ldots$ of length m, the badness is $|\hat{\phi}(Y) + (m-1)/m|$. Supremum badness is reached if $m = n$, such that the breakdown point following Sakata and White (1995, 1998) in this case is close to 1. This seems inappropriate, as the ordinary least squares estimator $\hat{\phi}_{\mathrm{OLS}}$ has already lost all its informational content under one extreme (patch of) outlier(s). Our definition in Section 2 automatically detects this loss of information and assigns the more intuitive value of 0 to the breakdown point of $\hat{\phi}_{\mathrm{OLS}}$.

In the asymptotic case, consider the set $\mathcal{Z}_p^\zeta$ of independent and identically distributed additive outlier processes

$$P(Z_i^\zeta = \zeta) = P(Z_i^\zeta = -\zeta) = \{1 - P(Z_i^\zeta = 0)\}/2 = p/2. \tag{13}$$

Under this contaminating process, badness equal to $|\hat{\phi} - \phi|$ and the set of alternative uncontaminated processes ($\phi' \in (-1, 1)$), we have

$$\hat{\phi}_{\mathrm{OLS}} = \frac{\mathbb{E}(Y_i Y_{i-1}) + \mathbb{E}(Z_i^\zeta Z_{i-1}^\zeta)}{\mathbb{E}(Y_{i-1}^2) + \mathbb{E}\{(Z_{i-1}^\zeta)^2\}} = O(\zeta^{-2}) \overset{\zeta \to \infty}{\longrightarrow} 0,$$

for arbitrary small positive p. The badness set R^* for arbitrary small but positive p and $\zeta = \infty$ is $\{|\phi - 0|\}$, which is a finite set for every $Y \in \mathcal{Y}$ (or $\phi \in (-1, 1)$). Therefore, the breakdown point of $\hat{\phi}_{\mathrm{OLS}}$ is 0 in the asymptotic AR(1) time series setting as well.

5.2. Deepest regression estimator

Consider the AR(1) model (1). The deepest regression estimator (Van Aelst and Rousseeuw, 2000) then is

$$\hat{\phi}_{\mathrm{DR}} = \mathrm{median}\left(\frac{Y_2}{Y_1}, \ldots, \frac{Y_n}{Y_{n-1}}\right). \tag{14}$$

If Y_i is observed with error as $\tilde{Y}_i = Y_i + Z_i$, where $Z_i = \zeta$ for some indices i and $Z_i = 0$ otherwise, the ratios in equation (14) can tend to ∞, 0 or 1, depending on whether ζ appears in the numerator, in the denominator or in both. Here the worst pattern of additive contamination takes the alternating form $Z_m^\zeta = (0, \zeta, 0, \zeta, \ldots)$ because each ζ contaminates two ratios in equation (14). Therefore, if badness is measured as $|\hat{\phi} - \phi|$, we have $R^*(\hat{\phi}_{\mathrm{DR}}, Y, Z_m^\infty, \mathbb{R}^n) = \{|\hat{\phi}_{\mathrm{DR}}(Y) - 0|, \infty\}$

for $m = \lfloor n/2 \rfloor$, which is a finite set for every $Y \in \mathcal{Y}$. The breakdown point of the deepest regression estimator is thus $(\lfloor n/2 \rfloor - 1)/n$ in this setting.

In the asymptotic setting, consider a Gaussian AR(1) process. Gaussianity is not essential, but it simplifies our exposition. The deepest regression estimator is given by the value of x that solves $F(x) = \frac{1}{2}$, where

$$F(x) = p(1-p)\mathbf{1}_{\{x \geqslant 0\}} + p^2 \mathbf{1}_{\{x \geqslant 1\}} + p(1-p)\mathbf{1}_{\{x \geqslant \infty\}} + (1-p)^2 \int_{-\infty}^{x} \frac{1}{\pi\{1 + (t-\phi)^2/(1-\phi^2)\}\sqrt{(1-\phi^2)}} \, \mathrm{d}t, \tag{15}$$

with $\mathbf{1}_A$ denoting the indicator function of the set A. This follows directly from equation (14) by noting that

$$Y_i / Y_{i-1} \stackrel{\mathrm{d}}{=} \phi + \sqrt{(1-\phi^2)}\varepsilon_i / \varepsilon_{i-1}$$

follows a Cauchy distribution with location ϕ and scale $(1-\phi^2)^{1/2}$. The distribution in expression (15) has a jump at 0 and at 1 for $p > 0$. Denote $F(0-)$ and $F(1-)$ as the lower left-hand limits of $F(x)$ at 0 and 1 respectively. For $p = \frac{1}{2}$ and all $|\phi| < 1$, we have $F(1) > \frac{1}{2}$ and $F(1-) < \frac{1}{2}$, such that the badness set collapses to a singleton for $p = \frac{1}{2}$. For $p < \frac{1}{2}$, we show that the badness $|\hat{\phi} - \phi|$ can still take a continuum of values such that the breakdown point of the deepest regression estimator in the AR(1) model is indeed $\frac{1}{2}$. Let $p = \frac{1}{2} - \eta$ for some $\eta > 0$, possibly arbitrarily small. Then

$$F(x) - \frac{1}{2} = \left[\frac{\tan^{-1}\{(x-\phi)/\sqrt{(1-\phi^2)}\}}{\pi} - \frac{1}{2} \right] \left(\frac{1}{2} + \eta \right)^2 + \eta,$$

for $x = 0$ and $x = 1-$. Using simple manipulations, rewrite $F(0) < \frac{1}{2} < F(1-)$ as

$$\frac{-\phi}{\sqrt{(1-\phi^2)}} < \tan\left\{ \frac{\pi}{2} - \frac{\eta\pi}{(\frac{1}{2} + \eta)^2} \right\} < \frac{1-\phi}{\sqrt{(1-\phi^2)}}.$$

There clearly is a continuum of values for ϕ sufficiently close, but not too close, to -1 such that these inequalities are satisfied. This proves the claim that the deepest regression estimator has a breakdown point of $\frac{1}{2}$ in this setting. The finite sample breakdown point of the deepest regression estimator converges to the asymptotic breakdown point as $(\lfloor n/2 \rfloor - 1)/n \to \frac{1}{2}$ for $n \to \infty$.

5.3. *Least median of squares estimator*

We conclude with a more involved example: the least median of squares estimator for time series models (Rousseeuw and Leroy, 1987; Lucas, 1997). Consider the AR(1) model where the innovations e_i form Gaussian white noise. We assume that badness is measured by $|\hat{\phi}(Y' + Z_p^{\zeta}) - \hat{\phi}(Y)|$ or $|\hat{\phi}(Y' + Z_p^{\zeta}) - \phi|$ for $Y' \in \mathcal{Y}$, where $\mathcal{Y}$ is the set of strictly stationary Gaussian AR(1) processes. We can characterize the elements Y' of $\mathcal{Y}$ by the value of the autoregressive parameter ϕ'.

Under the contaminating process (13), the least median of squares estimator $\hat{\phi}_{\mathrm{LMS}}$ of ϕ minimizes

$$\mathrm{median}\{e_i + Z_i^{\zeta} - \hat{\phi}_{\mathrm{LMS}} Z_{i-1}^{\zeta} + (\phi - \hat{\phi}_{\mathrm{LMS}}) Y_{i-1}\}^2. \tag{16}$$

With $\tilde{e}_i = \tilde{e}_i(\hat{\phi}_{\mathrm{LMS}}) = e_i + (\phi - \hat{\phi}_{\mathrm{LMS}})Y_{i-1}$, expression (16) solves

$$\begin{aligned}\tfrac{1}{2} = (1-p)^2\, P(\tilde{e}_i^2 < c) + \tfrac{1}{4}\sum_{i=1}^{2}(2p(1-p)\, P\{(\tilde{e}_i + w_i\zeta)^2 < c\} \\ + 2p(1-p)\, P\{(\tilde{e}_i - w_i\hat{\phi}_{\mathrm{LMS}}\zeta)^2 < c\} + p^2\, P[\{\tilde{e}_i + w_i(1-\hat{\phi}_{\mathrm{LMS}})\zeta\}^2 < c] \\ + p^2\, P[\{\tilde{e}_i + w_i(1+\hat{\phi}_{\mathrm{LMS}})\zeta\}^2 < c]) \qquad (17)\end{aligned}$$

with respect to c, where $w_1 = -w_2 = 1$. If we let $\zeta = \infty$, we see that there are only four interesting values of $\hat{\phi}_{\mathrm{LMS}}$, namely $\hat{\phi}_{\mathrm{LMS}}$ equal to ϕ, ± 1 and 0. We can now rewrite equation (17) as

$$c = \begin{cases} P^{-1}_{(e+\phi Y)^2}\{0.5(1-p)^{-1}\} & \text{for } \hat{\phi}_{\mathrm{LMS}} = 0, \\ P^{-1}_{e^2}\{0.5(1-p)^{-2}\} & \text{for } \hat{\phi}_{\mathrm{LMS}} = \phi, \\ P^{-1}_{\{e+(\phi-1)Y\}^2}\{0.5(1-2p+1.5p^2)^{-1}\} & \text{for } \hat{\phi}_{\mathrm{LMS}} = 1, \\ P^{-1}_{\{e+(\phi+1)Y\}^2}\{0.5(1-2p+1.5p^2)^{-1}\} & \text{for } \hat{\phi}_{\mathrm{LMS}} = -1, \end{cases}$$

where $P_X^{-1}(\cdot)$ is the inverse cumulative distribution function corresponding to the random variable X. It is clear from these formulae that, for p sufficiently small, $\hat{\phi}_{\mathrm{LMS}} = \phi$ gives the lower value of c. For ϕ near 0 or ± 1 and p sufficiently large, however, one of the other branches may dominate. Assume that $\mathcal{Y}$ is the set of AR(1) processes characterized by $\phi \in (-1, 1)$. Fig. 1 plots the value of p as a function of ϕ for which the median under $\hat{\phi}_{\mathrm{LMS}}$ equal to -1, 0 or 1 equals that under $\hat{\phi}_{\mathrm{LMS}} = \phi$. The minimum of these curves is depicted in bold. The maximum of the bold curve, 22.1% at $\phi = \pm 0.734$, is the breakdown point according to our new definition. If the fraction of contamination exceeds this breakdown point, the set of badness values under the postulated extreme additive outlier process equals $\{|-1-\phi|, |\phi|, |1-\phi|\}$ for $\mathcal{Y}$ (or $\phi' \in (-1, 1)$), which for every given Y (or ϕ) is a finite set. The breakdown point, therefore, lies below the familiar $\frac{1}{2}$ for the least median of squares in the regression setting. If we note that one additive outlier contaminates two couples (Y_{i-1}, Y_i) and (Y_i, Y_{i+1}), a breakdown point of $0.5/(1+p)$, where $p = 1$ is the order of the autoregressive process, might be deemed more appropriate in the autoregressive time series setting. Fig. 1, however, shows that the actual breakdown point is even lower. It is interesting that the breakdown points of the least median of squares and the deepest regression are different in the time series setting, whereas they are the same ($\frac{1}{2}$) in the setting of simple regression. This distinction does not emerge when earlier definitions of breakdown points are used.

6. Concluding remarks

We have introduced a new unified concept of breakdown which is applicable to settings with independent and dependent observations. Our new definition is more involved than traditional definitions, since in the setting of dependent observations outlier contaminations can exhibit many different patterns. Using examples from linear and non-linear regression, as well as time series, we showed that our new definition covers most of the familiar notions of breakdown. In some cases, however, our definition differs from the traditional definitions. This is especially relevant if allowable outlier configurations are subject to constraints, as in the qualitative response example in Section 4. Of particular interest is the setting of simple autoregressive time series, where our breakdown point correctly indicates that the estimator has been broken, whereas traditional definitions still do not indicate breakdown. A completely different example

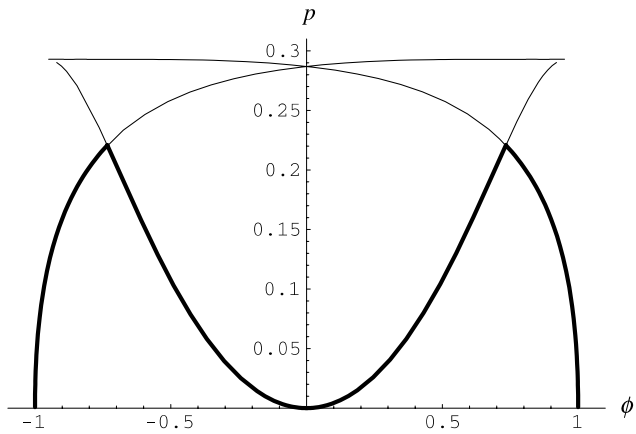

Fig. 1. Breakdown curve (in bold) of the least median of squares estimator in the Gaussian AR(1) model as a function of ϕ

from the non-linear regression setting revealed similar patterns. Our definition removes artificial sample dependence in breakdown points as presented in earlier definitions. An important consequence of our new definition is that different robust estimators of regression with the same high breakdown point can have surprisingly different breakdown points in the time series context. We believe that this will prove useful to distinguish between various robust estimators in time series and other situations with dependent observations, for instance spatial or spatiotemporal statistics.

Acknowledgements

The authors thank three referees, the Associate Editor and the Joint Editor for valuable comments that led to a marked improvement of the paper, as well as the participants at the Antwerp Seminar and International Conference on Robust Statistics 2001 for helpful remarks. The work of Genton was supported in part by National Science Foundation grant DMS-0204297.

References

Boente, G., Fraiman, R. and Yohai, V. J. (1987) Qualitative robustness for stochastic processes. *Ann. Statist.*, **15**, 1293–1312.

Donoho, D. L. and Huber, P. J. (1983) The notion of breakdown point. In *A Festschrift for Erich L. Lehmann* (eds P. J. Bickel, K. A. Doksum and J. L. Hodges, Jr), pp. 157–184. Belmont: Wadsworth.

Genton, M. G. (1998) Spatial breakdown point of variogram estimators. *Math. Geol.*, **30**, 853–871.

Hampel, F. R. (1971) A general qualitative definition of robustness. *Ann. Math. Statist.*, **42**, 1887–1896.

Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J. and Stahel, W. A. (1986) *Robust Statistics: the Approach based on Influence Functions*. New York: Wiley.

He, X. (1991) A local breakdown property of robust tests in linear regression. *J. Multiv. Anal.*, **38**, 294–305.

He, X. and Simpson, D. G. (1992) Robust direction estimation. *Ann. Statist.*, **20**, 351–369.

He, X. and Simpson, D. G. (1993) Lower bounds for contamination bias: globally minimax versus locally linear estimation. *Ann. Statist.*, **21**, 314–337.

He, X., Simpson, D. G. and Portnoy, S. L. (1990) Breakdown robustness of tests. *J. Am. Statist. Ass.*, **85**, 446–452.

Lucas, A. (1997) Asymptotic robustness of least median of squares for autoregressions with additive outliers. *Communs Statist. Theory Meth.*, **26**, 2363–2380.

Ma, Y. and Genton, M. G. (2000) Highly robust estimation of the autocovariance function. *J. Time Ser. Anal.*, **21**, 663–684.

Martin, R. D. (1980) Robust estimation of autoregressive models. In *Directions in Time Series*, pp. 228–262. Hayward: Institute of Mathematical Statistics.

Martin, R. D. and De Jong, J. (1977) Asymptotic properties of robust generalized M-estimates for the first-order autoregressive parameter. *Technical Memorandum*. Bell Laboratories, Murray Hill.
Rousseeuw, P. J. and Leroy, A. M. (1987) *Robust Regression and Outlier Detection*. New York: Wiley.
Sakata, S. and White, H. (1995) An alternative definition of finite sample breakdown point with applications to regression model estimators. *J. Am. Statist. Ass.*, **90**, 1099–1106.
Sakata, S. and White, H. (1998) High breakdown point conditional dispersion estimation with application to S&P 500 daily returns volatility. *Econometrica*, **66**, 529–567.
Stefanski, L. A. (2000) Measurement error models. *J. Am. Statist. Ass.*, **95**, 1353–1358.
Stromberg, A. J. and Ruppert, D. (1992) Breakdown in nonlinear regression. *J. Am. Statist. Ass.*, **87**, 991–997.
Van Aelst, S. and Rousseeuw, P. J. (2000) Robustness of deepest regression. *J. Multiv. Anal.*, **73**, 82–106.
Zhang, J. and Li, G. (1998) Breakdown properties of location *M*-estimators. *Ann. Statist.*, **26**, 1170–1189.