The Annals of Statistics
1996, Vol. 24, No. 6, 2593–2607

ROBUST ESTIMATION IN STRUCTURED LINEAR REGRESSION

BY LAMINE MILI[1] AND CLINT W. COAKLEY

Virginia Polytechnic Institute and State University

We dedicate this paper to the memory of Fred C. Schweppe (1934–1988), who made important contributions to robust state estimation in electric power systems.

A structured linear regression model is one in which there are permanent dependencies among some p row vectors of the $n \times p$ design matrix. To study structured linear regression, we introduce a new class of robust estimators, called D-estimators, which can be regarded as a generalization of the least median of squares and least trimmed squares estimators. They minimize a dispersion function of the ordered absolute residuals up to the rank h. We investigate their breakdown point and exact fit point as a function of h in structured linear regression. It is found that the D- and S-estimators can achieve the highest possible breakdown point for h appropriately chosen. It is shown that both the maximum breakdown point and the corresponding optimal value of h, h_{op}, are sample dependent. They hinge on the design but not on the response. The relationship between the breakdown point and the design vanishes when h is strictly larger than h_{op}. However, when h is smaller than h_{op}, the breakdown point depends in a complicated way on the design as well as on the response.

1. Introduction. Since the initial proposal made by Hodges (1967) and then further developed by Hampel (1971), the concept of the breakdown point has proved to be a powerful tool for the analysis and design of robust estimators. Roughly speaking, it is defined as the smallest fraction of outliers that can ruin an estimator. Asymptotically, it cannot exceed $1/2$ when regression equivariance prevails. A finite-sample version of the breakdown point that has gained wide acceptance was formulated by Donoho and Huber (1983). Given a sample $\mathscr{Z}$, the breakdown point is defined as the minimum fraction of observations of $\mathscr{Z}$ which, when replaced by arbitrary values, can carry the estimate over all bounds. The highest possible finite-sample breakdown point under regression equivariance was derived by Rousseeuw (1984). He showed that, when there are no linear dependencies among any p row vectors $\mathbf{x}_i^t$ of the $n \times p$ design matrix $\mathbf{X}$, a condition known as general position, the breakdown point is at most equal to $([(n-p)/2]+1)/n$, where $[\cdot]$ denotes the greatest integer function. Here n is the number of observa-

Received July 1993; revised January 1996.
[1]Research supported by NSF Grants ECS-90-09099 and ECS-92-57204.
AMS 1991 *subject classifications*. Primary 62G35; secondary 62J05, 62K99, 62N99.
Key words and phrases. Robust estimation, structured regression, general position, reduced position, breakdown point, exact fit point, D-estimators.

tions in $\mathbf{y}$ and p is the number of unknown parameters in $\boldsymbol{\theta}$ in the linear model

$$\mathbf{y} = \mathbf{X}\boldsymbol{\theta} + \mathbf{e}, \tag{1.1}$$

where $\mathbf{e}$ is an $n \times 1$ vector of random errors. Again under the general position assumption, it has been shown that the upper bound of the breakdown point is attained, for example, by the least median of squares (LMS) and least trimmed squares (LTS) estimators proposed by Rousseeuw (1984), the S-estimators initiated by Rousseeuw and Yohai (1984), the MM-estimators developed by Yohai (1987), the τ-estimators suggested by Yohai and Zamar (1988) and the one-step GM-estimators of Simpson, Ruppert and Carroll (1992) and Coakley and Hettmansperger (1993).

It turns out that all these results no longer hold when there are dependencies among some p rows of the design matrix, a situation that occurs quite often in practice. This is particularly true for the large class of regression problems characterized by a structured design matrix. Sparsity of the design matrix [Mili, Phaniraj and Rousseeuw (1990, 1991) and Ruckstuhl, Stahel and Dressler (1993)], replication of the observations [Coakley and Mili (1993)] and certain designs used in response surface methodology [Myers and Montgomery (1995)] are three typical examples of structured regression problems.

This paper derives, for structured regression, the expression for the highest possible finite-sample breakdown point that any regression equivariant estimator may have. It is found that this upper bound depends on the design matrix $\mathbf{X}$, specifically on the maximum number M of row vectors of $\mathbf{X}$ that lie on a $(p-1)$-dimensional hyperplane passing through the origin. For a given number n of observations, the maximum possible breakdown point is a decreasing function of M. Asymptotically, this upper bound is smaller than $1/2$ if M/n tends to a positive limit. The paper also introduces a new class of estimators, called D-estimators to indicate that they minimize a dispersion function of the ordered absolute residuals up to the rank h. This class can be regarded as a generalization of the LMS and LTS estimators as well as of the least trimmed absolute deviations (LTAD) estimator of Bassett (1991) and Tableman (1994). The class of D-estimators also contains the trimmed weighted L_p-estimators of Müller (1995). Upper and lower bounds of the breakdown point of the D-estimators are derived as a function of the quantile index h. It is shown that there exists an optimal value of h for which the D- and S-estimators achieve the highest possible breakdown point when regression equivariance holds. An algorithm that calculates M and the optimal h has been developed. It is available from the authors upon request.

The paper is organized as follows. Section 2 defines a structured regression model and gives two practical examples of such a model. Section 3 deals with the expression for the highest possible breakdown point and exact fit point in structured regression. Section 4 introduces the class of D-estimators. Section 5 gives breakdown bounds for D- and S-estimators and outlines an algorithm that calculates the optimal quantile index. Section 6 discusses the results and summarizes some ideas for future work. Complete proofs of all the results in this paper are presented in Mili and Coakley (1993).

2. Structured linear regression. Only a few published studies of the breakdown point in regression estimation, namely Coakley and Mili (1993), Davies (1993), Mili, Cheniae and Rousseeuw (1994), Müller (1995) and Ruckstuhl (1995), have treated the situation where the data points are not in general position [Rousseeuw (1984)]. General position means that any p data points determine a unique nonvertical hyperplane in the $(p+1)$-dimensional space $\mathscr{E}$ of the n observations contained in the sample $\mathscr{Z} = \{(x_i, y_i);\ i = 1, 2, \dots, n\}$. Equivalently, we have the following definition.

DEFINITION 2.1. The $(p+1)$-dimensional observations in the sample $\mathscr{Z}$ are said to be in *general position* if any p row vectors $\mathbf{x}_i^t$ of the $n \times p$ design matrix $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_n]^t$ are linearly independent. Otherwise we say that they are in *reduced position*.

Quite simply, reduced position means that there is a linear dependence between some p vectors of $\mathbf{X}$. Equivalently, in the p-dimensional factor space $\mathscr{F}$, there exists a hyperplane passing through the origin [i.e., a $(p-1)$-dimensional vector subspace] containing more than $p-1$ vectors $\mathbf{x}_i$.

DEFINITION 2.2. A linear regression model (1.1) is said to be *structured* if there are linear dependencies among some p row vectors of the design matrix $\mathbf{X}$ that occur regardless of the values taken by the carriers.

It is clear that in a structured linear regression model the observations are in reduced position; but the converse does not necessarily hold. Indeed, linear dependencies among some p row vectors of $\mathbf{X}$ may be casual, happening at one sample but not at another one. In that case they are not intrinsic to the problem at hand.

By definition, a structured regression model has a design matrix for which some p row vectors are subject to permanent linear dependencies. Besides this constraint, all possibilities for fixed and random carriers may be envisaged. For instance, some of the carriers may be fixed and the others random, or all may be of the same type. Some may even have identical values, zero in many cases. Replication of the observations and sparsity of the design matrix are two typical examples. Usually, physical phenomena obey some laws such as Kirchhoff's and Ohm's laws in circuit analysis, conservation of mass and energy in mechanics, fluid mechanics, molecular spectroscopy and crystallography, to cite a few. These laws translate into permanent dependencies among the rows of the design matrix when considering regression models. In addition, an observation may not be related to all the unknowns (parameters or state variables) but to only a few of them, which makes the design matrix sparse. Large-scale regression models with hundreds or thousands of unknowns usually are of this type. For these reasons, structured regression models often arise in practice. They constitute a broad family of regression problems that are especially encountered in the engineering field and physical sciences.

EXAMPLE 1 (Replication). Neter, Wasserman and Kutner (1985), page 306, present an example in which a quadratic model is fit to 14 observations (x_i, y_i), with two replicates at each of seven different levels of x. If two cases with the same x value are part of a set of three observations, then there is no unique determination of $\boldsymbol{\theta}$. So these data are in reduced position due to replication.

EXAMPLE 2 (Sparsity). An electric power system consists of lines and transformers forming a network that connects electric generators to a host of consumers spread over a large geographical area. A power system is supervised by software which computes a state estimator. Its role is to estimate the system's state variables from a collection of measurements on power flows at some lines and on power injections and voltage magnitudes at some nodes. The state variables, which include the voltage magnitudes and phase angles at all the nodes of the network, are related to the measurements through a nonlinear regression model. This model can be put into the matrix form (1.1) by performing a linearization around the flat voltage profile, defined as the state where all the voltages are at their nominal values and all the phase angles are 0. The main feature of the linear model so derived is that it involves a design matrix $\mathbf{X}$ that is very sparse and possesses many groups of p row vectors that are linearly dependent.

As an example, consider the system shown in Figure 1. It has five nodes depicted as horizontal bars and nine real power measurements represented by circles. These measurements can be expressed in terms of four voltage phase angles, say those at nodes 1 through 4; the fifth one is taken as a reference and set arbitrarily to 0. Assuming that the lines are identical, we get

$$\mathbf{X}^t = [\mathbf{x}_1, \dots, \mathbf{x}_9] = \begin{bmatrix} -1 & 1 & 2 & 1 & -1 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & -1 & 1 & 2 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & 1 \end{bmatrix}.$$

We observe that the sparsity of $\mathbf{X}$ induces linear dependencies among many groups of four row vectors of $\mathbf{X}$, which violates the general position assumption. For instance, when considering the four-dimensional factor space, there

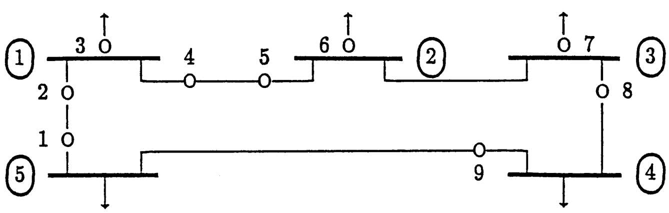

FIG. 1. *One-line diagram of the five-bus system.* ○ *denotes a real power measurement.*

are five points lying on a plane passing through the origin. These are data points 1 through 5, which satisfy $x_3 = x_4 = 0$.

3. Upper bounds for the breakdown point and exact fit point.

3.1. *Definitions.* Let us first recall the definition of the finite-sample breakdown point proposed by Donoho and Huber (1983). Consider a sample $\mathscr{Z}$ of n finite data points defined as $\mathscr{Z} = \{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_n, y_n)\}$. Replace m observations of $\mathscr{Z}$ by outliers which may take on arbitrary values, yielding the corrupted sample $\mathscr{Z}^*$. Let T be a regression estimator that maps $\mathscr{Z}$ into an estimate of $\boldsymbol{\theta}$ designated by $T(\mathscr{Z})$. Let the maximum bias be the supremum of $\|T(\mathscr{Z}^*) - T(\mathscr{Z})\|$ taken over all samples $\mathscr{Z}^*$, where $\|\cdot\|$ denotes the Euclidean norm.

DEFINITION 3.1. The breakdown point $\varepsilon_n^*(T, \mathscr{Z})$ of an estimator T at a sample $\mathscr{Z}$ is defined as

$$\varepsilon_n^*(T, \mathscr{Z}) = \min\left\{\varepsilon = m/n;\ \sup_{\mathscr{Z}^*}\|T(\mathscr{Z}^*) - T(\mathscr{Z})\| \text{ is infinite}\right\}. \tag{3.1}$$

It is the smallest fraction of outliers for which the maximum bias is unbounded or, equivalently, for which $\|T(\mathscr{Z}^*)\|$ is unbounded since $\|T(\mathscr{Z})\|$ is finite by assumption. Here, the outliers have to be placed in the least favorable way.

Another robustness concept, which is simpler than the breakdown point but closely related to it, was introduced by Donoho, Rousseeuw and Stahel [see, e.g., Donoho, Johnstone, Rousseeuw and Stahel (1985), Rousseeuw and Leroy (1987) and Ellis and Morgenthaler (1992)]. It derives from the so-called exact fit property.

DEFINITION 3.2. In linear regression, an estimator of p unknown parameters is said to have the *exact fit property* of order s/n if, whenever s observations lie exactly on a p-dimensional hyperplane, the fit yields that hyperplane.

The *exact fit point* of an estimator T is the minimum fraction $\delta_n^*(T, \mathscr{Z})$ of outliers for which the exact fit property no longer holds. Formally, we have

$$\delta_n^*(T, \mathscr{Z}) = \min\{m/n; \text{there exists } \mathscr{Z}^* \text{ such that } T(\mathscr{Z}^*) \neq \boldsymbol{\theta}\}, \tag{3.2}$$

where $\mathscr{Z}$ is the original sample containing n observations that lie on a p-dimensional hyperplane satisfying $y = \mathbf{x}^t\boldsymbol{\theta}$, and where $\mathscr{Z}^*$ is a contaminated sample obtained from $\mathscr{Z}$ by replacing any m observations.

3.2. *Upper bound on the breakdown point.* Let $\mathfrak{M}$ be the largest subset of observations $(\mathbf{x}_i, y_i)$ whose $\mathbf{x}_i$ lie on a $(p-1)$-dimensional hyperplane $\mathscr{M}$ passing through the origin. Assume that the M points in $\mathfrak{M}$ are the first numbered points, $(\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_M, y_M)$. Although the remaining $N = n - M$ observations fall outside $\mathscr{M}$, they span a vector subspace $\mathscr{N}$ whose dimension may be larger than 1. In fact, $\mathscr{N}$ may even be the whole factor space.

THEOREM 3.1. *The breakdown point* $\varepsilon_n^*(T, \mathscr{Z})$ *of any regression equivariant estimator* T *is at most equal to*

$$\varepsilon^*_{\max, n} = [(n - M + 1)/2]/n. \tag{3.3}$$

PROOF. The proof follows that of Theorem 4 of Rousseeuw and Leroy (1987), page 125. Suppose that T is a regression equivariant estimator with a breakdown point larger than $\varepsilon^*_{\max, n}$. Then $\|T(\mathscr{Z}^*)\|$ remains bounded if $m = [(n - M + 1)/2]$ points of $\mathscr{Z}^*$ take on arbitrary values. Let $\mathscr{Z}^*$ be a sample obtained by adding $\mathbf{x}_i^t\mathbf{v}$ to m data points of $\Re$, where $\mathbf{v} \in \mathbb{R}^p$ is orthogonal to the subspace $\mathscr{M}$. This means that $\mathbf{v}$ is such that $\mathbf{x}_i^t\mathbf{v} = 0$ for $i = 1, \ldots, M$ and $\mathbf{x}_i^t\mathbf{v} \neq 0$ for $i = M + 1, \ldots, n$. Let $\mathscr{Z}^{**}$ be a regression transformation of $\mathscr{Z}^*$ obtained by subtracting $\mathbf{x}_i^t\mathbf{v}$ from every point in $\mathscr{Z}^*$. Then $\mathscr{Z}^{**}$ contains $n - m - M = [(n - M)/2] \leq m$ altered points relative to $\mathscr{Z}$. From the above statements, three results can be inferred:

(i) $\|T(\mathscr{Z}^*)\|$ is bounded by hypothesis;
(ii) $\|T(\mathscr{Z}^{**})\|$ is bounded by hypothesis;
(iii) $\|T(\mathscr{Z}^{**})\| = \|T(\mathscr{Z}^*) - \mathbf{v}\|$ by regression equivariance,

yielding a contradiction because $\mathbf{v}$ can be made arbitrarily large. □

An analogous result for regression functionals may be found in Davies (1993). It is interesting to note that, when the observations are in general position, that is, when $M = p - 1$, formula (3.3) reduces to $\varepsilon^*_{\max, n} = ([(n - p)/2] + 1)/n$, which is precisely the expression for the maximum breakdown point derived by Rousseeuw (1984); see also Rousseeuw and Leroy (1987) for more details.

THEOREM 3.2. *The exact fit point* $\delta_n^*(T, \mathscr{Z})$ *of any regression equivariant estimator* T *is at most equal to*

$$\delta^*_{\max, n} = [(n - M + 1)/2]/n. \tag{3.4}$$

The proof of this theorem is analogous to that of the maximum breakdown point given in Theorem 3.1.

Concerning the asymptotic values of the breakdown and exact fit points, two cases must be considered. In the first case, M remains finite, implying that the maximum breakdown point tends to $1/2$, the same upper bound as when the data points are in general position. In the second case, M tends to ∞ with n, implying that the asymptotic maximum breakdown point may be smaller than $1/2$. To see that, let λ be the limit of M/n as n grows to ∞. From (3.3), it follows that $\varepsilon^*_{\max, n} \to (1 - \lambda)/2$, a limit that vanishes as $\lambda \to 1$. Similar conclusions hold for the asymptotic exact fit point.

4. D-estimators. Rousseeuw (1984) proposed two estimators which have an asymptotic breakdown point of 1/2 when the observations are in general position. These are the well-known LMS and LTS estimators, defined as the minimizers of the median squared residual $r^2_{(h)}$ and the sum of the h smallest squared residuals, respectively. The fact that these two estimators have different objective functions but have the same breakdown point has motivated us to define a new class of estimators, called D-estimators and first introduced in the simple regression case by Coakley and Mili (1993). One characteristic of such estimators is that, as will be seen in Section 5, their breakdown point is determined by the value of h, which we call the *quantile index*.

4.1. *Definition of D-estimators*. A D-estimator minimizes an objective function $D(\boldsymbol{\theta}, h; \mathscr{Z})$ satisfying the following properties:

P1. $D(\boldsymbol{\theta}, h; \mathscr{Z}) \geq 0$;

P2. D depends only on $|r|_{(1)}, |r|_{(2)}, \dots, |r|_{(h)}$;

P3. D is nondecreasing in each of the $|r|_{(i)}$'s, $i = 1, \dots, h-1$, and D is strictly increasing in $|r|_{(h)}$;

P4. for $|r|_{(1)}, |r|_{(2)}, \dots, |r|_{(h-1)}$ fixed, $\lim_{|r|_{(h)} \to \infty} D(\boldsymbol{\theta}, h; \mathscr{Z}) = \infty$.

For each value of h, there is a collection of D-estimators all having the same breakdown point. One example of a D-objective function is

$$D(\boldsymbol{\theta}, h; \mathscr{Z}) = \sum_{i=1}^{h} c_i |r|_{(i)}, \tag{4.1}$$

where $c_1, c_2, \dots, c_{h-1}$ are nonnegative constants and c_h is a positive constant. In the location case with $h = [n/2] + 1$, this is the LTAD estimator of Bassett (1991) and Tableman (1994). In fact, the D-estimator given by (4.1) is a generalization of the LTAD estimator to multiple linear regression. Another example is obtained by squaring the residuals instead of taking their absolute values in (4.1). Hence the LMS and LTS are D-estimators. The LMS corresponds to $c_h = 1$ and $c_i = 0$ for $i < h$; the LTS corresponds to $c_1 = c_2 = \cdots = c_h = 1$.

REMARK 4.1. (i) Our D-estimators have no known connection to the class of minimum distance estimators referred to as D-estimators by Hampel, Ronchetti, Rousseeuw and Stahel (1986), page 113.

(ii) There is very little overlap between the classes of D-estimators and S-estimators. For instance, S-estimators with continuously differentiable ρ-functions are not D-estimators. For an example of a D-estimator that is not an S-estimator, consider the D-objective function $D(\boldsymbol{\theta}, h; \mathscr{Z}) = \{\exp(-1/|r|_{(1)})\}|r|_{(h)}$.

4.2. *Equivariance properties and efficiency of D-estimators.* It is straightforward to see that any D-estimator is affine and regression equivariant because of the relationships between the transformed and untransformed residuals. What about the efficiency of D-estimators? Loosely speaking, it increases with h. For instance, Coakley (1991) showed that the Gaussian efficiency of LTS vs. LS surpasses 1/3, 1/2 and 3/4 for $h \geq 0.8n$, $0.88n$ and $0.96n$, respectively. In simple regression, Coakley, Mili and Cheniae (1994) reported on simulations which suggest that the finite-sample efficiencies of LMS and LTS vs. LS under Gaussian errors are actually higher than their asymptotic efficiencies. They also showed how the finite-sample efficiencies of LTS and LMS are related to the value of the quantile index h; the increase in efficiency is nearly linear in h and is stronger for LTS than for LMS, as one would expect.

5. Breakdown bounds of *D*- and *S*-estimators. In this section we derive the expression for the finite-sample breakdown point of D- and S-estimators when the quantile index h is larger than or equal to the optimal value, h_{op}. When h is strictly smaller than h_{op}, we give only upper and lower bounds on the breakdown point of D-estimators. We also derive the expression for their exact fit point for all h. In the sequel we will make use of the following relationships:

$$p + [q/2] = [(2p + q)/2], \tag{5.1}$$

$$p - [q/2] = [(2p - q + 1)/2] \quad \text{for } p \geq [q/2], \tag{5.2}$$

$$[p/2] - q = [(p - 2q)/2] \quad \text{for } [p/2] \geq q, \tag{5.3}$$

where p and q are two positive integers.

5.1. *Breakdown point of a D-estimator.* We first derive the expression for h_{op} and then investigate the breakdown point of a D-estimator when $h \geq h_{\mathrm{op}}$.

THEOREM 5.1. *The breakdown point of a D-estimator attains* $\varepsilon^*_{\max, n} = [(n - M + 1)/2]/n$ *for any* h_{op} *satisfying*

$$h_L \leq h_{\mathrm{op}} \leq h_U, \tag{5.4}$$

where $h_L = [(n + M + 1)/2]$ *and* $h_U = [(n + M + 2)/2]$.

PROOF. It suffices to show that the breakdown point of a D-estimator T for h_{op} given by (5.4) cannot be smaller than $\varepsilon^*_{\max, n}$ at any sample $\mathscr{Z}$. Equivalently, we have to show that $\|T(\mathscr{Z}^*)\|$ cannot be brought to ∞ for any contaminated sample $\mathscr{Z}^*$ obtained from $\mathscr{Z}$ by replacing $m = [(n - M + 1)/2] - 1 = [(n - M - 1)/2]$ data points by arbitrary values. Consider a fit with a finite $\|\boldsymbol{\theta}\|$. Because $\mathscr{Z}^*$ contains $n - m$ unaltered observations and m arbitrarily changed ones, the good fit yields $n - m$ finite residuals. Since $n - m = h_U \geq h_{\mathrm{op}}$, a D-objective function assessed at this good fit will have a finite value.

Now we will construct the least favorable bad fit $\mathscr{P}_M$ passing through m outliers. This is done as follows. Consider $p-1$ points, say $\{(\tilde{\mathbf{x}}_i, \tilde{y}_i);\ i = 1, \dots, p-1\}$, where all $\|\tilde{\mathbf{x}}_i\|$ and $\tilde{y}_i$ are finite and where the vectors $\tilde{\mathbf{x}}_i$ span the $(p-1)$-dimensional vector subspace $\mathscr{M}$. Pick any observation from the subset $\mathfrak{N}$, say $(\mathbf{x}_n, y_n)$. Since, by definition, the vector $\mathbf{x}_n$ does not lie on $\mathscr{M}$, the p points in $\{(\tilde{\mathbf{x}}_1, \tilde{y}_1), (\tilde{\mathbf{x}}_2, \tilde{y}_2), \dots, (\tilde{\mathbf{x}}_{p-1}, \tilde{y}_{p-1}), (\mathbf{x}_n, y_n)\}$ uniquely determine a p-dimensional hyperplane, namely the fit $\mathscr{P}_M$. Let $(\mathbf{x}_n, y_n)$ be made arbitrarily large under the constraint that $\mathbf{x}_n$ remains linearly independent with the $\tilde{\mathbf{x}}_i$. Pick any subset of $m-1$ data points from the subset $\mathfrak{N}$, say $\{(\mathbf{x}_{n-m+1}, y_{n-m+1}), \dots, (\mathbf{x}_{n-1}, y_{n-1})\}$. Place them on the hyperplane $\mathscr{P}_M$ by appropriately adjusting their $(\mathbf{x}_i, y_i)$ values, again under the constraint that their $\mathbf{x}_i$ do not lie on $\mathscr{M}$.

By construction, the fit $\mathscr{P}_M$ yields $M + m$ finite residuals, which is the highest possible number that any bad fit may have. Since $M + m = h_L - 1 < h_{\mathrm{op}}$, a D-objective function assessed at such a bad fit (or any other bad fit) will be unbounded. Therefore, we conclude that a D-objective function attains its minimum only at a fit with a bounded $\|\boldsymbol{\theta}\|$. □

The results given in Theorems 3.1 and 5.1 hold whether or not contamination is restricted to the responses; see Müller (1995). A D-estimator satisfying (5.4) is called optimal. If the data points are in general position, yielding $M = p - 1$, then the relationship $h_{\mathrm{op}} = [n/2] + [(M+2)/2]$, which satisfies (5.4), reduces to (4.9) of Rousseeuw and Leroy (1987), page 124, namely $[n/2] + [(p+1)/2]$. The latter can still be used in practice without severe consequences when m is slightly greater than $p-1$, as might happen when there are dependencies among few groups of p row vectors of $\mathbf{X}$. However, when M is substantially larger than $p-1$, it is advisable to find M and then to calculate h_{op} using (5.4). Otherwise, the breakdown point of the D-estimator at hand would be much smaller than the maximum.

An algorithm that calculates M has been developed. It is based on elemental sets consisting of $p-1$ rows of $\mathbf{X}$ and it is complete in the sense that it finds the exact value of M (not an estimate of M based on randomly chosen elemental subsets). However, it is computationally more efficient than an exhaustive search of all $\binom{n}{p-1}$ elemental subsets. Instead, it uses the information gained when more than $p-1$ points are found on a $(p-1)$-dimensional vector subspace of $\mathscr{F}$ in order to skip over some of the elemental subset calculations. This program is available from the authors upon request.

THEOREM 5.2. *A D-estimator T satisfying $h \ge h_U$ has a breakdown point given by*

$$\varepsilon_n^*(T, h; \mathscr{Z}) = (n - h + 1)/n. \tag{5.5}$$

PROOF. Assume $h \ge h_U$. Let us first show that $\varepsilon_n^*(T, h; \mathscr{Z}) \le (n - h + 1)/n$. Let $\mathscr{Z}^*$ be a contaminated sample obtained from $\mathscr{Z}$ by replacing $m = n - h + 1$ data points by arbitrary values. Thus, $\mathscr{Z}^*$ contains $n - m =$

$h - 1$ unaltered observations. Consider a fit with a bounded $\|\boldsymbol{\theta}\|$. By construction, this fit results in $n - m$ finite residuals. Since $n - m < h$, we conclude that a D-objective function assessed at this good fit is unbounded. Because it will be so for any good fit, it holds that the minimum of a D-objective function may be attained at a fit with unbounded $\|\boldsymbol{\theta}\|$.

Now let us prove that $\varepsilon_n^*(T, h; \mathscr{Z}) \geq (n - h + 1)/n$. Equivalently, we have to show that a D-estimator does not break down for a contaminated sample $\mathscr{Z}^*$ obtained from $\mathscr{Z}$ by replacing $m = n - h$ observations by arbitrary values. Again, consider a fit with a bounded $\|\boldsymbol{\theta}\|$. By construction, this good fit results in $n - m$ finite residuals for the sample $\mathscr{Z}^*$. Since $h = n - m$, any D-objective function assessed at this good fit is finite. Now consider the least favorable bad fit $\mathscr{P}_M$ (as described in the proof of Theorem 5.1) that passes through the m unbounded points of $\mathscr{Z}^*$. It gives $M + m$ finite residuals. Since $M + m \leq M + n - [(n + M + 2)/2] = [(n + M - 1)/2] < h$, we conclude that any D-objective function assessed at this bad fit (or any other bad fit) will be unbounded. Consequently, the minimum of a D-objective function is only attained at a fit with a bounded $\|\boldsymbol{\theta}\|$. □

5.2. *Exact fit point of a D-estimator*. We will investigate the exact fit point of a D-estimator first for $M < h \leq n$ and then for $h \leq M$.

THEOREM 5.3. *The exact fit point $\delta_n^*(T, h; \mathscr{Z})$ of a D-estimator T is given by*

$$\delta_n^*(T, h; \mathscr{Z}) = \min\{(h - M)/n, (n - h + 1)/n\} \quad \text{for } M < h \leq n. \tag{5.6}$$

PROOF. We first show that $\delta_n^*(T, h; \mathscr{Z}) \leq b$, where $b = (n - h + 1)/n$ for $h_U \leq h \leq n$ and $b = (h - M)/n$ for $M < h < h_U$. Consider a sample $\mathscr{Z}$ containing n points which lie on a p-dimensional hyperplane $\mathscr{P}_0$. By regression equivariance, we may assume without loss of generality that $\mathscr{P}_0$ is the hyperplane $y = 0$. Let $m = nb$ observations of $\mathscr{Z}$ take on arbitrary nonzero response values. For the resulting sample $\mathscr{Z}^*$, the fit $\mathscr{P}_0$ yields $n - m$ zero residuals.

Assume that $h_U \leq h \leq n$. Then $n - m = n - nb = h - 1$. It follows that a D-objective function assessed at the fit $\mathscr{P}_0$ is unbounded, implying that the minimum may be attained at a fit with an unbounded $\|\boldsymbol{\theta}\|$, which is different from $\mathscr{P}_0$. Hence, the associated D-estimator breaks down.

Assume that $M < h < h_U$. Construct a least favorable bad fit $\mathscr{P}_M$ satisfying $y = \mathbf{x}^t\boldsymbol{\theta}$, where $\boldsymbol{\theta}$ is an arbitrary vector different from $\mathbf{0}$. In addition, $\mathscr{P}_M$ passes through the M observations $\{(\mathbf{x}_i, 0);\ i = 1, \dots, M\}$ of $\mathfrak{M}$ and through one point $(\tilde{\mathbf{x}}, \tilde{y})$, where $\tilde{y} \neq 0$ and where $\tilde{\mathbf{x}}$ falls outside $\mathscr{M}$, and hence is linearly independent with a basis of $\mathscr{M}$, say $\{\mathbf{x}_1, \dots, \mathbf{x}_{p-1}\}$. Place $m = nb = h - M$ data points of the subset $\mathfrak{N}$ on $\mathscr{P}_M$ by changing their y_i values from 0 to $y_i = \mathbf{x}_i^t\boldsymbol{\theta} \neq 0$. These y_i are different from 0 because $\mathbf{x}_i = \sum_{j=1}^{p-1} a_{ij}\mathbf{x}_j + b_i\tilde{\mathbf{x}}$ with $b_i \neq 0$, which implies that $y_i = \sum_{j=1}^{p-1} a_{ij} y_j + b_i \tilde{y}$, where $y_j = 0$ and $\tilde{y} \neq 0$. Consequently, the fit $\mathscr{P}_M$ results in $M + m = h$ zero residuals, yielding the minimum value for a D-objective function, 0, by property P3. The minimum being attained for a fit different from $\mathscr{P}_0$, we conclude that a D-estimator breaks down. Similarly, it can be shown that $\delta_n^*(T, h; \mathscr{Z}) \geq b$. □

The following theorem was inspired by a suggestion from Stromberg (1992).

THEOREM 5.4. *A D-estimator T satisfying $h \leq M$ has a zero exact fit point.*

PROOF. By regression equivariance, we consider, without loss of generality, $\mathscr{Z} = \{(\mathbf{x}_i, y_i = 0),\ i = 1, \dots, n\}$. It is clear that, without modifying $\mathscr{Z}$, any fit (as steep as desired) that passes through the M observations of $\mathfrak{M}$ yields a D-objective function with zero value, which is the minimum. Therefore, $\|\boldsymbol{\theta}\|$ associated with this fit is unbounded. □

The results of Theorems 5.3 and 5.4 are depicted in Figure 2.

5.3. *Breakdown bounds of a D-estimator.* We now give upper and lower bounds for the breakdown point of a D-estimator when h is smaller than h_{op} and show that there exist samples where these bounds are attained.

THEOREM 5.5. *Let $\mathscr{Z}$ be the sample $\{(\mathbf{x}_i, y_i);\ i = 1, \dots, n\}$. Let $\mathscr{Z}'$ be a sample obtained from $\mathscr{Z}$ by changing the response values y_i so that the n data points lie exactly on a hyperplane. A D-estimator T satisfying $h < h_L$ has a breakdown point $\varepsilon_n^*(T, h; \mathscr{Z})$ given by*

$$\delta_n^*(T, h; \mathscr{Z}') \leq \varepsilon_n^*(T, h; \mathscr{Z}) \leq \min\{(h - p + 1)/n,\ \varepsilon_{\max, n}^*\}. \tag{5.7}$$

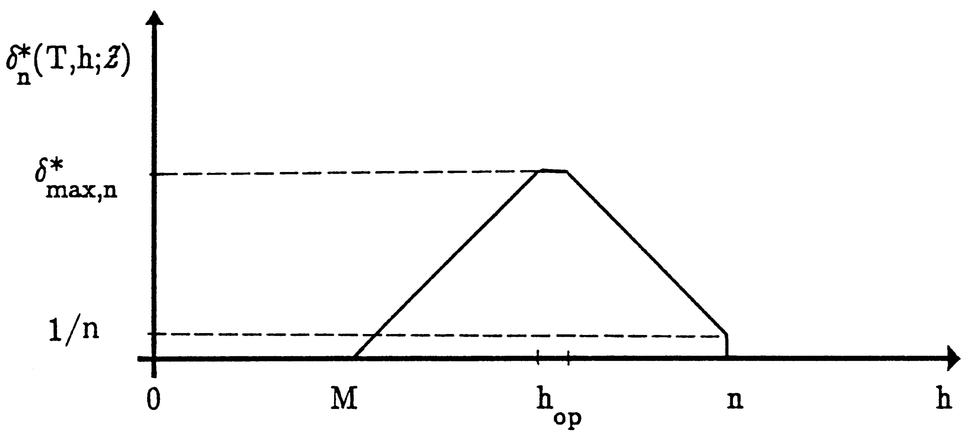

FIG. 2. *Exact fit point vs. the quantile index h of a D-estimator.*

PROOF. The proof is similar to the proofs of Theorems 5.2 and 5.3 and can be found in Mili and Coakley (1993).

REMARK 5.1. The first inequality in (5.7), namely $\delta_n^*(T, h; \mathscr{Z}') \leq \varepsilon_n^*(T, h; \mathscr{Z})$, seems to contradict the proposition shown by Rousseeuw and Leroy (1987), page 123, which states that the breakdown point cannot be larger than the exact fit point for a regression and scale equivariant estimator. However, this proposition is true only when considering the same sample $\mathscr{Z}$ for both the breakdown point and the exact fit point, viz. $\delta_n^*(T, h; \mathscr{Z}) \geq \varepsilon_n^*(T, h; \mathscr{Z})$. This does not hold in (5.7), which involves two different samples $\mathscr{Z}$ and $\mathscr{Z}'$.

It is very difficult to find a simple expression for the breakdown point of a D-estimator when $h < h_L$, because it depends in a complicated way on the responses y_i of the sample $\mathscr{Z}$, even for a fixed design. In other words, its breakdown point is sample dependent. However, the bounds given in (5.7) are sharp in the sense that there exist designs for which they are attained, as shown in Theorems 5.6 and 5.7.

THEOREM 5.6. *Consider a sample $\mathscr{Z}$ where all the M data points of the subset $\mathfrak{M}$ lie exactly on a $(p-1)$-dimensional hyperplane $\mathscr{P}$. Then a D-estimator T satisfying $M < h < h_L$ has a breakdown point $\varepsilon_n^*(T, h; \mathscr{Z})$ equal to $(h - M)/n$.*

PROOF. Assume that $h < h_L$. From Theorem 5.5, $\varepsilon_n^*(T, h; \mathscr{Z}) \geq (h - M)/n$. Let us now show that $\varepsilon_n^*(T, h; \mathscr{Z}) \leq (h - M)/n$. Equivalently, we have to prove that a D-estimator breaks down when $m = h - M$ observations of $\mathscr{Z}$ are replaced by arbitrary values. Construct the least favorable bad fit $\mathscr{P}_M$ so that it contains the $(p-1)$-dimensional hyperplane $\mathscr{P}$. In that case, $\mathscr{P}_M$ has a total number of zero residuals equal to $M + m = h$, which leads to the minimum value for the D-objective function, 0. Consequently, it may be selected as the final fit, yielding an unbounded $\|\boldsymbol{\theta}\|$. $\square$

THEOREM 5.7. *Consider a sample $\mathscr{Z}$ that satisfies the following assumptions*:

A1. The N observations of $\mathfrak{N}$ are in general position, implying that any subset with size $\min(p, N)$ *of vectors $\mathbf{x}_i$ associated with $\mathfrak{N}$ are linearly independent.*

A2. The N observations of $\mathfrak{N}$ along with μ observations of $\mathfrak{M}$, where $p - 1 \leq \mu \leq M - 1$, lie exactly on a p-dimensional hyperplane $\mathscr{P}$ in the observation space $\mathscr{E}$ of dimension $p + 1$.

A3. The remaining $M - \mu$ observations of $\mathfrak{M}$ fall outside the hyperplane $\mathscr{P}$. Under these assumptions, a D-estimator T defined with $h = n\varepsilon_{\max, n}^* + \mu$, where $\varepsilon_{\max, n}^*$ is given by (3.3), *has a breakdown point $\varepsilon_n^*(T, h; \mathscr{Z})$ equal to $\varepsilon_{\max, n}^*$.*

PROOF. Note that we have $[(n-M+1)/2]+p-1 \le h \le [(n-M+1)/2]+M-1 < h_L$. The first inequality yields $(h-p+1)/n \ge \varepsilon^*_{\max,n}$, which, by Theorem 5.5, implies that $\varepsilon^*_n(T,h;\mathscr{Z}) \le \varepsilon^*_{\max,n}$. It remains then to show that $\varepsilon^*_n(T,h;Z) \ge \varepsilon^*_{\max,n}$. In other words, we have to prove that a D-estimator with an h as previously defined remains bounded when $m = [(n-M-1)/2]$ observations of $\mathscr{Z}$ have been changed to arbitrary values. With the contaminated sample $\mathscr{Z}^*$ so obtained, the fit $\mathscr{P}$ defined under assumption A2, which has a bounded $\|\boldsymbol{\theta}\|$, leads to a number of zero residuals that is no smaller than $N+\mu-m = [(n-M+2)/2]+\mu \ge h$. Thus, a D-objective function assessed at this good fit has zero value for any set of m altered points. Now consider the least favorable fit $\mathscr{P}_M$, which yields $M+m$ finite residuals. Because $M+m = [(n-M+1)/2]+M-1 \ge h$, a D-objective function assessed at this bad fit is finite. It will be different from 0, however, because the total number of zero residuals is strictly smaller than h. Indeed, the largest number of zero residuals is achieved when $\mathscr{P}_M$ contains, in addition to the m altered points of $\mathfrak{N}$, the μ observations of $\mathfrak{M}$ that are on $\mathscr{P}$. This yields a maximum number of zero residuals equal to $m+\mu = [(n-M-1)/2]+\mu \le h-1$. Since any other bad fit leads to a number of zero residuals that is no larger than $m+\mu$, we conclude that the minimum of a D-objective function is reached at the fit $\mathscr{P}$, which has a bounded $\|\boldsymbol{\theta}\|$. □

5.4. *Breakdown point of an S-estimator*. Let ρ be a function that is (i) symmetric and continuously differentiable with $\rho(0)=0$, and (ii) strictly increasingly on $[0,c]$ and constant on $[\ c,\infty)$, where c is a positive constant. Then an S-estimator minimizes

$$s\big(r_1(\boldsymbol{\theta}),\dots,r_n(\boldsymbol{\theta})\big) \tag{5.8}$$

subject to

$$\frac{1}{n}\sum_{i=1}^{n}\rho\left(\frac{r_i}{s}\right) = k, \tag{5.9}$$

where k is a given nonnegative constant [Rousseeuw and Yohai (1984).]

THEOREM 5.8. *Consider an integer h that satisfies $h_U \le h \le n$. Then an S-estimator T defined with $k/\rho(c) = (n-h+1)/n$ has a breakdown point and an exact fit point equal to $(n-h+1)/n$.*

PROOF. Following Croux, Rousseeuw and Hössjer (1994), Theorem 3, and Rousseeuw and Leroy (1987), Lemma 4, page 136, it can easily be shown that

$$\alpha|r|_{(h)} \le s\big(r_1(\boldsymbol{\theta}),\dots,r_n(\boldsymbol{\theta})\big) \le \beta|r|_{(h)}, \tag{5.10}$$

where $\alpha = 1/c$ and $\beta = \rho^{-1}(\rho(c)/(h+1))$. Hence, the proofs of Theorems 5.1, 5.2 and 5.3 for $h \ge h_U$ hold for the S-estimators as well. □

Theorem 5.8 shows that the S-estimators do attain the maximum breakdown point for $k/\rho(c) = \varepsilon^*_{\max,n}$ and h_{op} satisfying (5.4).

6. Discussion and future work. In structured linear regression, the breakdown point of a D-estimator may hinge on the sample $\mathscr{Z}$ at hand. It depends in a very complicated way on both the response and the design when the quantile index h is strictly smaller than h_{op}, a value at which the breakdown point is maximum. When h is strictly larger than h_{op}, both dependencies vanish and the breakdown point is expressed only in terms of h and n. When h is at the optimal value, h_{op}, the dependencies on the response disappear, but those related to the design remain. Note that the maximum breakdown point and h_{op} are expressed in terms of M. In structured regression, M may be much larger than $p - 1$, hence the need to calculate M whenever such a situation occurs. Yet, this can be performed prior to experimentation in a designed regression. An algorithm that does precisely that has been developed.

Several lines of future work remain open. One is to extend these results to the generalized S-estimators of Croux, Rousseeuw and Hössjer (1994). Of great interest is the analysis of the effect of reduced position on the maximum bias curve for all these estimators [see Martin, Yohai and Zamar (1989)]. Another possibility is to investigate the local breakdown point concept [Mili, Cheniae and Rousseeuw (1994) and Ruckstuhl (1995)] in conjunction with the decomposition of the design space into subspaces.

Acknowledgments. The authors are grateful to Jeffrey D. Vest for assistance with computer programming, to a referee for helpful comments on an earlier version of the paper and to the Editors and an Associate Editor for handling the manuscript.

REFERENCES

BASSETT, G. W., JR. (1991). Equivariant, monotonic, 50% breakdown estimators. *Amer. Statist.* **45** 135–137.

COAKLEY, C. W. (1991). Advances in the study of breakdown and resistance. Ph.D. dissertation, Dept. Statistics, Pennsylvania State Univ.

COAKLEY, C. W. and HETTMANSPERGER, T. P. (1993). A bounded influence, high breakdown, efficient regression estimator. *J. Amer. Statist. Assoc.* **88** 872–880.

COAKLEY, C. W. and MILI, L. (1993). Exact fit points under simple regression with replication. *Statist. Probab. Lett.* **17** 265–271.

COAKLEY, C. W., MILI, L. and CHENIAE, M. G. (1994). Effect of leverage on the finite sample efficiencies of high breakdown estimators. *Statist. Probab. Lett.* **19** 399–408.

CROUX, C., ROUSSEEUW, P. J. and HÖSSJER, O. (1994). Generalized S-estimators. *J. Amer. Statist. Assoc.* **89** 1271–1281.

DAVIES, P. L. (1993). Aspects of robust linear regression. *Ann. Statist.* **21** 1843–1899.

DONOHO, D. L. and HUBER, P. J. (1983). The notion of breakdown point. In *A Festschrift for Erich L. Lehman* (P. J. Bickel, K. A. Doksum and J. L. Hodges, Jr., eds.) 157–184. Wadsworth, Belmont, CA.

DONOHO, D. L., JOHNSTONE, I., ROUSSEEUW, P. J. and STAHEL, W. (1985). Comment on "Projection pursuit" by P. J. Huber. *Ann. Statist.* **13** 496–500.

ELLIS, S. P. and MORGENTHALER, S. (1992). Leverage and breakdown in L_1 regression. *J. Amer. Statist. Assoc.* **87** 143–148.

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*. Wiley, New York.

HODGES, J. L., JR. (1967). Efficiency in normal samples and tolerance of extreme values for some estimates of location. *Proc. Fifth Berkeley Symp. Math. Statist. Probab.* **1** 163–168. Univ. California Press, Berkeley.

MARTIN, R. D., YOHAI, V. J. and ZAMAR, R. H. (1989). Min-max bias robust regression. *Ann. Statist.* **17** 1608–1630.

MILI, L., CHENIAE, M. G. and ROUSSEEUW, P. J. (1994). Robust state estimation of electric power systems. *IEEE Trans. Circuits and Systems* **41** 349–358.

MILI, L. and COAKLEY, C. W. (1993). Robust estimation in structured linear regression. Technical Report 93-13, Dept. Statistics, Virginia Polytechnic Institute and State Univ., Blacksburg.

MILI, L., PHANIRAJ, V. and ROUSSEEUW, P. J. (1990). Robust estimation theory for bad data diagnostics in electric power systems. In *Control and Dynamic Systems: Advances in Theory and Applications* (C. T. Leondes, ed.). *Advances in Industrial Systems* **37** 271–325. Academic Press, New York.

MILI, L., PHANIRAJ, V. and ROUSSEEUW, P. J. (1991). Least median of squares estimation in power systems. *IEEE Trans. Power Systems* **6** 511–523.

MÜLLER, CH. H. (1995). Breakdown points for designed experiments. *J. Statist. Plann. Inference* **45** 413–427.

MYERS, R. H. and MONTGOMERY, D. C. (1995). *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*. Wiley, New York.

NETER, J., WASSERMAN, W. and KUTNER, M. H. (1985). *Applied Linear Statistical Models*, 2nd ed. Irwin, Homewood, IL.

ROUSSEEUW, P. J. (1984). Least median of squares regression. *J. Amer. Statist. Assoc.* **79** 871–880.

ROUSSEEUW, P. J. and LEROY, A. M. (1987). *Robust Regression and Outlier Detection*. Wiley, New York.

ROUSSEEUW, P. J. and YOHAI, V. (1984). Robust regression by means of S-estimators. In *Robust and Nonlinear Time Series Analysis*. *Lecture Notes in Statist.* **26** 256–272. Springer, New York.

RUCKSTUHL, A. F. (1995). Analysis of the T_2 emission spectrum by robust estimation techniques. Ph.D. thesis, Seminar für Statistik, Swiss Federal Institute of Technology, Zürich.

RUCKSTUHL, A. F., STAHEL, W. A. and DRESSLER, K. (1993). Robust estimation of term values in high-resolution spectroscopy: application to the $e^3\Sigma_u^+ \to a^3\Sigma_g^+$ spectrum of T_2. *Journal of Molecular Spectroscopy* **160** 434–445.

SIMPSON, D. G., RUPPERT, D. and CARROLL, R. J. (1992). On one-step GM estimates and stability of inferences in linear regression. *J. Amer. Statist. Assoc.* **87** 439–450.

STROMBERG, A. J. (1992). Personal communication.

TABLEMAN, M. (1994). The asymptotics of the least trimmed absolute deviations (LTAD) estimator. *Statist. Probab. Lett.* **17** 387–398.

YOHAI, V. J. (1987). High breakdown point and high efficiency robust estimates for regression. *Ann. Statist.* **15** 642–656.

YOHAI, V. J. and ZAMAR, R. (1988). High breakdown-point estimates of regression by means of the minimization of an efficient scale. *J. Amer. Statist. Assoc.* **83** 406–413.

DEPARTMENT OF ELECTRICAL ENGINEERING
VIRGINIA POLYTECHNIC INSTITUTE
& STATE UNIVERSITY
BLACKSBURG, VIRGINIA 24061-0111
E-MAIL: lmili@vt.edu

DEPARTMENT OF STATISTICS
VIRGINIA POLYTECHNIC INSTITUTE
& STATE UNIVERSITY
BLACKSBURG, VIRGINIA 24061-0439
E-MAIL: coakley@vt.edu