On the global robustness of generalized S-estimators

José R. Berrendero *

Departamento de Matemáticas

Universidad Autónoma de Madrid, 28049 Madrid, Spain

Abstract

A generalized S-estimator (GS-estimator) of regression is obtained by minimizing an M-estimator of scale applied to the pairwise differences of residuals $r_i(\boldsymbol{\theta}) - r_j(\boldsymbol{\theta})$, $i < j$. In this paper, we focus on the global robustness properties of these estimators. It was pointed out by Croux *et al.* (1994) that two GS-estimators with the same breakdown point may have very different global robustness behavior. Accordingly, we supplement the information given by the breakdown point with the *explosion rate*, a summary measure of the robustness behavior of the estimator when the amount of contamination is large. We provide formulas that allow the computation of explosion rates and establish a link between the local behavior at zero of the score function on which the M-estimator is based and the robustness of the corresponding GS-estimator. Finally, we apply the explosion rate to quantify the loss of robustness when using GS-estimators instead of the simpler non-generalized S-estimators.

Running title: Global robustness of GS-estimators

Key words: Breakdown point, Explosion rate, Bias robustness, Maxbias curve, GS–estimators, Scale, Robust regression.

AMS 1991 subject classification: 62F35.

*This research has been partially supported by DGESIC (Spain) under grant number PB97–0021

1 Introduction

Consider the linear regression model with p–dimensional regressors and intercept parameter

$$y_i = \alpha_0 + \boldsymbol{\theta}_0'\mathbf{x}_i + \sigma_0 u_i,\ \ i = 1, \ldots, n, \tag{1}$$

and let $r_i = r_i(\alpha, \boldsymbol{\theta}) = y_i - \alpha - \boldsymbol{\theta}'\mathbf{x}$, $i = 1, \ldots, n$ be the residuals corresponding to the parameter set $(\alpha, \boldsymbol{\theta})$. Since classical least squares estimators of α_0 and $\boldsymbol{\theta}_0$ are extremely sensitive to departures from this model, different more stable or robust alternatives have been proposed in the statistical literature. A general approach to find robust estimators of α_0 and $\boldsymbol{\theta}_0$ is looking for the vector of parameters that minimizes a robust scale of the residuals. Following this scheme, Rousseeuw and Yohai (1984) defined regression S-estimators as the minimizers of an M-scale of the residuals. S-estimators possess very attractive robustness properties but, unfortunately, their efficiencies are quite low (see, for instance, Hössjer, 1992). To increase the efficiency, Croux, Rousseeuw and Hössjer (1994) defined the class of generalized S-estimators (GS-estimators) as the vector of parameters that minimizes an M-scale of the pairwise differences of residuals. More precisely, a GS-estimator, $\widehat{\boldsymbol{\theta}}$, is defined as $\widehat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} S_n(\boldsymbol{\theta})$, where

$$S_n(\boldsymbol{\theta}) = \inf\{s > 0 : \begin{pmatrix} n \\ 2 \end{pmatrix}^{-1} \sum_{i<j} \chi\left(\frac{r_i(\alpha, \boldsymbol{\theta}) - r_j(\alpha, \boldsymbol{\theta})}{s}\right) < b\}.$$

That is, $S_n(\boldsymbol{\theta})$ is the M-scale of the pairwise differences of the residuals based on the score function χ and the constant b. Notice that $r_i(\alpha, \boldsymbol{\theta}) - r_j(\alpha, \boldsymbol{\theta})$ does not depend on the intercept parameter α, which can be estimated later on with a robust location estimator. The asymptotic distribution of regression GS-estimators has been studied by Hössjer, Croux and Rousseeuw (1994). In this note, we address some aspects of their robustness properties.

A useful measure of the degree of robustness of an estimator T is given by its maximum bias curve or *maxbias curve*, $B_T(\epsilon)$. This function measures the maximum asymptotic bias of an estimator when the data includes up to a fraction ϵ of contamination (see more formal definitions in equations (3) and (14) below). In this paper, we are mainly interested in the global robustness of the estimators. Loosely speaking, an estimator whose maxbias is relatively small for large fractions of contamination is called *globally robust*. Usually, the breakdown point of T, defined as

$$\epsilon^* = \sup\{\epsilon > 0 : \, B_T(\epsilon) < \infty\},$$

is computed to assess the global robustness of T. For $\epsilon < \epsilon^*$, the estimator T is still informative, in the sense that its asymptotic value is not arbitrarily determined by the contaminations. However, as Croux *et al.* (1994) pointed out, it is possible (and indeed frequent) for a pair of GS-estimators to have the same breakdown point and drastically different global robustness behavior. The following example illustrates this fact:

Example. Let $\widehat{\boldsymbol{\theta}}_1$ and $\widehat{\boldsymbol{\theta}}_2$ be a pair of GS-estimators based on the score function $\chi_1(x) = I\{|x| \geq 1\}$, with constants $b_1 = 0.9375$ and $b_2 = 0.4375$ respectively ($I(A)$ stands for the indicator function on the set A). A simple reasoning shows that $\widehat{\boldsymbol{\theta}}_i$, for $i = 1, 2$, is the vector of parameters that minimizes the $(1 - b_i)$-quantile of the pairwise differences of residuals. Using Theorem 4 in Croux *et al.* (1994), it can be checked that the breakdown point of both estimators is $\epsilon^* = 0.25$. The corresponding maxbias curves, $B_1(\epsilon)$ and $B_2(\epsilon)$, are displayed in Figure 1. Notice that both maxbias curves are practically identical for $\epsilon < 0.1$. However, as $\epsilon \to \epsilon^*$, $B_1(\epsilon)$ goes to infinity much faster than $B_2(\epsilon)$. In fact, the behavior of $B_2(\epsilon)$ is only explosive for values of ϵ very close to the breakdown point. On the other hand, although the theoretical breakdown point of $\widehat{\boldsymbol{\theta}}_1$ is 0.25, we would say that its *practical* breakdown point is less than 0.25. The conclusion is that having the same breakdown point cannot be taken, without further analysis, as an indication that the maxbias curves will have similar behaviors for large values of ϵ.

FIGURE 1 ABOUT HERE

The preceding example motivates the study of the relative maxbias behavior of a pair of GS-estimators, $\widehat{\boldsymbol{\theta}}_1$ and $\widehat{\boldsymbol{\theta}}_2$, with common breakdown point ϵ^*, as $\epsilon \to \epsilon^*$. We will address this question using the *relative explosion rate* of $\widehat{\boldsymbol{\theta}}_1$ and $\widehat{\boldsymbol{\theta}}_2$, defined by Berrendero *et al.* (1998) as

$$r(\widehat{\boldsymbol{\theta}}_1, \widehat{\boldsymbol{\theta}}_2) = \lim_{\epsilon \to \epsilon^*} \frac{B_1(\epsilon)}{B_2(\epsilon)},$$

provided that the limit exists. If $0 < r(\widehat{\boldsymbol{\theta}}_1, \widehat{\boldsymbol{\theta}}_2) < \infty$, we say that $\widehat{\boldsymbol{\theta}}_1$ and $\widehat{\boldsymbol{\theta}}_2$ have the same order of global robustness. Moreover, if $r(\widehat{\boldsymbol{\theta}}_1, \widehat{\boldsymbol{\theta}}_2) = 0$, then $\widehat{\boldsymbol{\theta}}_1$ has higher order than $\widehat{\boldsymbol{\theta}}_2$. Obviously, estimators with higher order should be preferred. In Sections 2 and 3 below, we derive formulas for the relative explosion rates of GS-estimators and establish conditions under which two GS-estimators have the same order of global robustness.

It is well-known that some robustness properties of an estimator are determined by some features of its score function χ. For instance, regression S- and GS-estimators have a strictly positive breakdown point if and only if χ is also bounded. This kind of relationships are useful because they allow to determine the properties of the estimator by simple inspection of χ. By looking closely to the global robustness of the estimators, in Sections 2 and 3 we will establish a link between the global robustness of a GS-estimator and the local behavior at zero of its score function.

The paper is organized as follows: Section 2 is devoted to obtain explosion rates for dispersion GS-estimators as a previous step to the more complicated regression model, which is considered in Section 3. Some further remarks and conclusions are given in Section 4. All the proofs are deferred to the Appendix.

2 Global robustness of dispersion GS-estimators

In this section, we give some results concerning the maxbias curves of dispersion GS-estimators. Assume that we have i.i.d. observations y_i, $i = 1, \ldots, n$, which follow a location-dispersion model $y_i = \theta_0 + \sigma_0 u_i$, where the errors, u_i, have a distribution F_0, with variance 1, satisfying

A1. *F_0 has a continuous unimodal density f_0 which is symmetric about the origin.*

To allow for a fraction ϵ of outliers we suppose that the actual true distribution of the observations $y_1, \ldots, y_n$ belongs to the contamination neighborhood

$$V_\epsilon(F_0) = \{F : F(y) = (1-\epsilon)F_0[(y-\theta_0)/\sigma_0] + \epsilon H(y),\ H \text{ arbitrary distribution}\}.$$

In this setup, a GS-estimator for the dispersion parameter σ_0 is defined as $\widehat{\sigma}_n = S(F_n)$, where F_n is the empirical distribution function of the sample, and $S(F)$ is defined as

$$S(F) = \inf\{s > 0 : E_{F\times F}\chi[(X-Y)/s] < b\} \quad (2)$$

We will assume that the score function χ satisfies

A2. *χ is even, nondecreasing on $(0, \infty)$ and continuous at 0, with $\chi(0) = 0$ and at most a finite number of discontinuities. Moreover, $\chi(c) = \chi(\infty) = 1$ for some number $0 < c < \infty$, called the* tuning constant.

In order $S(F)$ be Fisher-consistent (i.e. $S(F_0) = \sigma_0$) the score function and the constant b cannot be chosen independently. Define $\mathcal{C}_a = \{\chi : E_{F_0 \times F_0} \chi(X - Y) = a\}$, for $0 < a < 1$. Then, Fisher-consistency requires that χ and b are related by the restriction $\chi \in \mathcal{C}_b$. Usually this condition is guaranteed by a suitable choice of the tuning constant so that we will assume it holds in the rest of the paper.

Fisher-consistency ensures that dispersion GS-estimators converge to σ_0 in absence of contamination. However, a fraction ϵ of contamination produces an asymptotic bias which is bounded by the maxbias curve. We must distinguish between implosion and explosion maxbias curves (see Martin and Zamar, 1993) defined by

$$B_S^+(\epsilon) = \sup_{F \in V_\epsilon(F_0)} S(F)/\sigma_0 \quad \text{and} \quad B_S^-(\epsilon) = \inf_{F \in V_\epsilon(F_0)} S(F)/\sigma_0, \tag{3}$$

which take into account the effect of outliers and inliers respectively. Accordingly, we define the explosion and the implosion breakdown points as $\epsilon^+ = \inf\{\epsilon > 0 : B_S^+(\epsilon) < \infty\}$ and $\epsilon^- = \inf\{\epsilon > 0 : B_S^-(\epsilon) > 0\}$. By equivariance considerations, we can assume w.l.o.g. that $\sigma_0 = 1$ in equation (3).

In Theorem 1 below, we give formulas to compute the maxbias curves for dispersion GS-estimators. We will need some previous notation. Define the functions

$$g(s) = E_{F_0 \times F_0} \chi[(X - Y)/s], \quad h(s) = E_{F_0} \chi(X/s), \quad \text{and} \tag{4}$$

$$m(s; \epsilon) = (1 - \epsilon)^2 g(s) + 2\epsilon(1 - \epsilon) h(s). \tag{5}$$

Define also

$$S^+(\epsilon) = g^{-1}[(b - 2\epsilon + \epsilon^2)(1 - \epsilon)^{-2}] \quad \text{and} \quad S^-(\epsilon) = m^{-1}(b; \epsilon), \tag{6}$$

where $m^{-1}(\cdot\,; \epsilon)$ is the inverse of $m(\cdot\,; \epsilon)$ for a fixed value of ϵ. We have the following result:

Theorem 1. *Let $S(F)$ be the dispersion GS-functional defined in equation (2). Under* **A1** *and* **A2**, *$B_S^+(\epsilon) = S^+(\epsilon)$ and $B_S^-(\epsilon) = S^-(\epsilon)$, where $S^+(\epsilon)$ and $S^-(\epsilon)$ are defined in equation (6).*

Notice that Theorem 1 implies that the explosion and implosion breakdown points of dispersion GS-estimators are given by $\epsilon^+ = 1-\sqrt{1-b}$ and $\epsilon^- = \sqrt{1-b}$ respectively.

In Figure 2, we display the explosion and implosion maxbias curves for several dispersion GS-estimators (as we are mainly interested in the global robustness behavior, we have only considered fractions of contamination larger than 0.25). The definition of the score functions can be found in Table 1 below. All the estimators have been tuned to have 0.5 explosion and implosion breakdown points. The implosion maxbias curves are almost identical. Therefore, the robustness properties of the estimators should be compared on the ground of their explosion maxbias curves. Regarding these curves, the most robust estimator is that based on χ_6, followed in decreasing order of robustness by the estimators based on Huber, Tukey and linear score functions.

FIGURE 2 ABOUT HERE

In Theorem 2 below, the maxbias curves of dispersion GS-estimators are explored further to detect which features of the score function are more important in determining the global robustness of the estimators. It turns out that the concept of *local order at zero* of χ will be very useful.

Definition 1. *The function χ has* local order κ at zero *if $\chi^{(j)}(0) = 0$, for all $j < \kappa$ and $\chi^{(\kappa)}(0) > 0$, where $\chi^{(j)}$ denotes the j^{th} right derivative at zero of χ.*

Using the concept of local order at zero of χ, the following theorem gives approximations for the explosion and implosion maxbias curves of dispersion GS-estimators. We use the standard notation $f(x) \sim g(x)$ as $x \to x_0$ meaning that $\lim_{x\to x_0} f(x)/g(x) = 1$.

Theorem 2. *Under the same assumptions of Theorem 1,*
(a) if χ has local order $\kappa > 0$ at zero and $\lim_{s\to\infty} s^{\kappa+1} f_0(s) = 0$, then

$$S^+(\epsilon) \sim C \left[\frac{(1-\epsilon)^2}{b - 2\epsilon + \epsilon^2}\right]^{1/\kappa} \sim \tilde{C}(\epsilon^+ - \epsilon)^{-1/\kappa}, \quad as\ \epsilon \to \epsilon^+, \tag{7}$$

where

$$C = \left[\frac{\chi^{(\kappa)}(0)}{k!} E_{F_0\times F_0}|X - Y|^\kappa\right]^{1/\kappa} \tag{8}$$

and $\tilde{C} = (\sqrt{1-b}/2)^{1/\kappa} C$.
(b) If $\chi(x)$ *is differentiable for* $x \in (0, c)$, *where* c *is the tuning constant, then*

$$S^-(\epsilon) \sim D(\epsilon^- - \epsilon), \;\; \textit{as } \epsilon \to \epsilon^-, \tag{9}$$

where

$$D = \left[\left(f_0^*(0) \frac{(1-\sqrt{1-b})^2}{\sqrt{1-b}} + 2 f_0(0)(1-\sqrt{1-b}) \right) \int_0^c (1-\chi(x))dx \right]^{-1}. \tag{10}$$

Here, f_0^* *is the density function corresponding to the distribution of* $X - Y$ *for a pair of independent random variables,* X *and* Y, *distributed as* F_0.

From (7) it follows that score functions with larger local order produce estimators with better maxbias performance. Since functions with large local order are "flat" near zero, the "flatness" of χ in a neighborhood of zero yields more robust estimators (at least for large ϵ). This accounts for the fact that χ_6 ($\kappa = 6$) produces the more robust estimator whereas the linear function ($\kappa = 1$) corresponds to the less robust estimator (see Figure 2 (a)).

It also follows from equation (7) that a pair of dispersion GS-estimators S_1 and S_2 have the same order of global robustness if the corresponding score functions χ_1 and χ_2 have the same local order at zero. In this case, from (7) and (8) we have that

$$r^+(S_1, S_2) = \lim_{\epsilon \to \epsilon^+} \frac{B^+_{S_1}(\epsilon)}{B^+_{S_2}(\epsilon)} = \left(\frac{\chi_1^{(\kappa)}(0)}{\chi_2^{(\kappa)}(0)} \right)^{1/\kappa}. \tag{11}$$

That is, the relative explosion maxbias behavior of S_1 and S_2 is determined by the curvatures at zero of χ_1 and χ_2.

From (9), the implosion maxbias curve is roughly proportional to $\epsilon^- - \epsilon$, when ϵ is close to ϵ^-, for any score function. For this reason, no great differences among the implosion bias curves of dispersion GS-estimators should be expected (see Figure 2 (b)). Furthermore, from (9) and (10),

$$r^-(S_1, S_2) = \lim_{\epsilon \to \epsilon^-} \frac{B^-_{S_1}(\epsilon)}{B^-_{S_2}(\epsilon)} = \frac{\int_0^{c_2} (1-\chi_2(y))dy}{\int_0^{c_1} (1-\chi_1(y))dy}. \tag{12}$$

In Table 1, both the explosion and the implosion rates are displayed for several dispersion GS-estimators. The tuning constants have been chosen to yield 0.5 breakdown points. The score function taken as baseline is Huber's. The corresponding efficiencies under the normal model have been computed using Lemma 3.5 in Hössjer *et al.* (1994) and are also displayed. It is surprising that the breakdown point, the implosion rate and the efficiency are almost the same for all the considered score functions. As a consequence, the local order at zero of the score function is crucial to compare the estimators. The larger the local order at zero, the better is the estimator.

TABLE 1 ABOUT HERE

REMARK 1: Due to the considerations above, the estimator based on a jump score function χ_j (that is, the limiting case when $\kappa \to \infty$) seems to be the best in the class of dispersion GS-estimators. This estimator was introduced by Rousseeuw and Croux (1993), who called it Q. This estimator is a more efficient alternative to the *median absolute deviation* (MAD) and the *shortest half* (SHORTH). It is perhaps of interest to compare the global robustness of this estimator with that of the MAD or the SHORTH. In Berrendero and Zamar (1999) it is shown that, in the Gaussian case, $r^+(\text{Q},\text{SHORTH}) = 2.12$ and $r^+(\text{Q},\text{MAD}) = 1.06$. Moreover, it holds that (see Remark 4 in the Appendix)

$$r^-(\text{Q},\text{MAD}) = r^-(\text{Q},\text{SHORTH}) = 1.11. \tag{13}$$

Therefore, there is a trade-off in this case between the gain in efficiency when using pairwise differences of observations (since the SHORTH converges at a slow rate to a non-gaussian distribution whereas Q converges to a gaussian distribution at the usual $n^{-1/2}$ rate) and the loss of robustness, which is quantified by an increment in the relative explosion rate. For large amounts of contamination the explosion maxbias of Q is roughly twice as great as the explosion maxbias of the SHORTH. However, the implosion maxbias of Q is still very close to that of the SHORTH or MAD.

3 Global robustness of regression GS-estimators

In this section we come back to the regression model (1). We will assume that the errors, u_i, and the regressors, $\mathbf{x}_i$, are independent and normally distributed. Let $\boldsymbol{\mu}$ and

Σ_0 be the mean vector and the covariance matrix of $\mathbf{x}_i$. Under these assumptions, the joint $p+1$-dimensional distribution of $(y_i, \mathbf{x}_i)$, H_0, is multivariate normal. However, as in Section 2, we assume that the actual true distribution H of $(y_i, \mathbf{x}_i)$ belongs to the contamination neighborhood

$$V_\epsilon(H_0) = \{H: \ H = (1-\epsilon)H_0 + \epsilon\tilde{H}, \ \tilde{H} \text{ arbitrary distribution}\}.$$

Let $F_{H,\alpha,\boldsymbol{\theta}}$ be the distribution of the residuals $r_i(\alpha, \boldsymbol{\theta})$. A regression GS-estimator is given by $\widehat{\boldsymbol{\theta}}_n = \mathbf{T}(H_n)$, where H_n is the empirical distribution of the data, $\mathbf{T}$ is given by $\mathbf{T}(H) = \arg\min_{\boldsymbol{\theta}} S(F_{H,\alpha,\boldsymbol{\theta}})$, and $S(F)$ was defined in equation (2) above. Notice that $S(F_{H,\alpha,\boldsymbol{\theta}})$ does not depend on the intercept α, which can be estimated in a second step. The maxbias curve of $\mathbf{T}$ is defined as

$$B_{\mathbf{T}}(\epsilon) = \sup_{H \in V_\epsilon} \left[(\mathbf{T}(H) - \boldsymbol{\theta}_0)'\Sigma_0(\mathbf{T}(H) - \boldsymbol{\theta}_0)\right]^{1/2} / \sigma_0. \quad (14)$$

Since $B_{\mathbf{T}}(\epsilon)$ is defined to be a regression and affine invariant quantity and GS-estimators are regression and affine equivariant, we can assume w.l.o.g. that $\boldsymbol{\mu} = \mathbf{0}$, $\Sigma_0 = I$, $\boldsymbol{\theta}_0 = \mathbf{0}$ and $\sigma_0 = 1$.

To obtain a formula for the maxbias curve of GS-estimators under the Gaussian model we follow the indications given by Croux *et al.* (1994) (see remarks under Theorem 4 in that paper). We must compute $s_1 = g^{-1}[(b - 2\epsilon + \epsilon^2)(1-\epsilon)^{-2}] = S^+(\epsilon)$ and find the solution of the following equation in γ:

$$(1-\epsilon)^2 g\left(\frac{S^+(\epsilon)}{\sqrt{1+\gamma^2}}\right) + 2\epsilon(1-\epsilon)h\left(\frac{S^+(\epsilon)}{\sqrt{1+\gamma^2}}\right) = b.$$

With our notation, see (5), the last equation amounts to $m[S^+(\epsilon)/\sqrt{1+\gamma^2}; \epsilon] = b$. It follows that $S^+(\epsilon)/\sqrt{1+\gamma^2} = m^{-1}(b;\epsilon) = S^-(\epsilon)$ and, therefore,

$$\gamma = B_{\mathbf{T}}(\epsilon) = \left[\left(\frac{S^+(\epsilon)}{S^-(\epsilon)}\right)^2 - 1\right]^{1/2}, \quad \text{for} \ \ \epsilon < \min\{\epsilon^-, \epsilon^+\}, \quad (15)$$

and $B_{\mathbf{T}}(\epsilon) = \infty$, for $\epsilon > \min\{\epsilon^-, \epsilon^+\}$. Here, $S^+(\epsilon)$ and $S^-(\epsilon)$ are the explosion and the implosion maxbias curves for the dispersion GS-estimator based on the same score function as $\mathbf{T}$ (see equation (6) and Theorem 1 above). Equation (15) shows a nice relationship between the maxbias curves of dispersion and regression GS-estimators.

This relationship allows to extend the conclusions obtained in Section 2 to the Gaussian regression model.

Notice that the breakdown point of $\mathbf{T}$ is $\epsilon^* = \min\{\epsilon^-, \epsilon^+\} = \min\{\sqrt{1-b}, 1-\sqrt{1-b}\}$. Using this expression it is easy to check that $\epsilon^* = 0.25$ for the estimators of the example in the Introduction. In general, two score functions $\chi_1 \in \mathcal{C}_{b_1}$ and $\chi_2 \in \mathcal{C}_{b_2}$ such that $\sqrt{1-b_1} + \sqrt{1-b_2} = 1$ will produce two regression GS-estimators with the same breakdown point. However if both score functions have the same local order at zero $\kappa > 1$, the estimator corresponding to the lower constant is always preferable. This fact is shown in the following theorem.

Theorem 3. *Let $\mathbf{T}_1$ and $\mathbf{T}_2$ be two regression GS-estimators based on a pair of score functions $\chi_1 \in \mathcal{C}_{b_1}$ and $\chi_2 \in \mathcal{C}_{b_2}$ satisfying* **A2** *with the same local order at zero $\kappa > 1$. Assume that $\sqrt{1-b_1} + \sqrt{1-b_2} = 1$. If $b_1 > b_2$, then $r(\mathbf{T}_1, \mathbf{T}_2) = \infty$.*

Theorem 3 explains the behavior of the estimators in the example of the Introduction. What Theorem 3 shows is that the estimator $\widehat{\boldsymbol{\theta}}_2$ has higher order of global robustness than $\widehat{\boldsymbol{\theta}}_1$ and it is therefore more globally robust.

Given $b \in (0,1)$, and a pair of score functions $\chi_1 \in C_b$ and $\chi_2 \in C_b$ with the same local order at zero, the corresponding regression GS-estimators share a common breakdown point and also a common order of global robustness. In this case, from (15), the relative explosion rate is given by

$$r(\mathbf{T}_1, \mathbf{T}_2) = \frac{r^+(S_1, S_2)}{r^-(S_1, S_2)},$$

where S_1 and S_2 are the dispersion GS-estimators based on χ_1 and χ_2. We have applied this fact together with (11) and (12) to obtain the relative explosion rate (with respect to the estimator based on Huber's function) of several regression GS-estimators with breakdown point $\epsilon^* = 0.5$. The results are displayed in Table 2. The efficiencies of the estimators are also reported. Note that these efficiencies are remarkably similar for all the considered estimators. Again, as in the dispersion setup, the explosion rate and, therefore, the local order at zero of the score function, is of concern to rank the estimators. Similarly to the dispersion case, the estimator based on a jump function (which coincides with the vector of parameters that gives the *least quartile of differences* (LQD) of the residuals) appears to be the best one among the considered estimators.

Table 2 about here

To conclude this section we compare the global robustness of a regression GS-estimator, $\mathbf{T}^*$, with that of a simpler non-generalized S-estimator, $\mathbf{T}$, when both are based on the same score function. Our aim is to quantify the trade-off between the gain in efficiency and the loss of robustness when considering pairwise differences of residuals instead of the residuals themselves. Obviously, for each score function we must consider two different tuning constants c^* and c, in order to match the breakdown point of the two estimators. For the sake of simplicity we will only consider the case when the common breakdown point is $\epsilon^* = 0.5$ although similar expressions could be obtained for $\epsilon^* < 0.5$. We show in Theorem 4 below that although both estimators have the same breakdown point and the same order of global robustness, using pairwise differences produces a loss of robustness since the relative explosion rate is usually greater than one.

Theorem 4. *Let $\mathbf{T}^*$ be a regression GS-estimator based on χ_{c^*} and let $\mathbf{T}$ be a regression S-estimator based on χ_c. Assume that the score functions satisfy* **A2** *and have local order κ at zero. Finally, assume that the tuning constants are chosen so that both estimators have a common breakdown point $\epsilon^* = 0.5$. Then,*

$$r(\mathbf{T}^*, \mathbf{T}) = 2^{\frac{\kappa-2}{2\kappa}} \frac{2\sqrt{2}}{2\sqrt{2}+1} \left(\frac{\int_0^{c^*} [1 - \chi_{c^*}(y)]\, dy}{\int_0^{c} [1 - \chi_{c}(y)]\, dy} \right) \left(\frac{\chi_{c^*}^{(\kappa)}(0)}{\chi_{c}^{(\kappa)}(0)} \right)^{1/\kappa}. \tag{16}$$

Remark 2: When the score functions χ_c and χ_{c^*} only differ in the tuning constant (that is, $\chi_c(x) = \chi(x/c)$ and $\chi_{c^*}(x) = \chi(x/c^*)$ for some score function χ with tuning constant equal to 1) then the last two factors of (16) cancel and (16) reduces to

$$r(\mathbf{T}^*, \mathbf{T}) = 2^{\frac{\kappa-2}{2\kappa}} \frac{2\sqrt{2}}{2\sqrt{2}+1} \tag{17}$$

and so the relative explosion rate only depends on χ through its local order at zero, κ.

Remark 3: Note that Theorem 4 above does not cover the important case when χ is a jump function, that is, when $\mathbf{T}^*$ is the LQD-estimator and $\mathbf{T}$ is the LMS-estimator (see Rousseeuw, 1984). However, in this case we have that (see Remark 1 above)

$$r(\text{LQD,LMS}) = \frac{r^+(\text{Q,SHORTH})}{r^-(\text{Q,SHORTH})} = \frac{2.12}{1.11} = 1.91.$$

This result coincides with the corresponding computation of Croux, *et al.* (1994), equation (16).

We have applied formula (17) to several score functions and collected the results in Table 3. As shown by the proof of Theorem 4, the relative rate for the regression estimators is the quotient between the explosion and implosion rates of the corresponding dispersion estimators, which are also included. We see in Table 3 that, as the local order at zero of the score function increases, the loss of robustness also increases. For the truncated linear score function ($\kappa = 1$), the maxbias curves of $\mathbf{T}$ and $\mathbf{T}^*$ are very similar for large amount of contamination; however, for the jump score function ($\kappa = \infty$) the maxbias curve of $\mathbf{T}^*$ for large amount of contamination is roughly twice as great as that of $\mathbf{T}$. It is somewhat surprising that dispersion generalized S-estimators handle inliers slightly better than non-generalized ones as the implosion rates of the table are all slightly greater than one.

Table 3 about here

4 Conclusions

In this paper we have applied the concept of relative explosion rate to study the global robustness properties of dispersion and regression generalized S-estimators. First, we have observed that two estimators with the same breakdown point may have very different maxbias curves. This fact motivates a further study of the factors that govern the relative maxbias performance of two estimators with common breakdown point. Regarding this question, we have showed that the local behavior at zero of the score functions on which the estimators are based is crucial. In fact, the relative rate between two explosion maxbias curves is determined by the rate between the curvatures at zero of the corresponding score functions. As a consequence, score functions which are “flat” near zero produce more robust estimators.

The choice of the score function is not very important for the efficiency of the estimators. In the dispersion model, all the considered estimators have an efficiency of approximately 82% under the Gaussian central model whereas in the regression setup, the efficiency of all the estimators decreases to roughly 68%. Therefore, it seems reasonable to rank the estimators attending to robustness criteria. From the results of this paper, jump score functions, which are totally flat near zero, should be selected to compute regression and dispersion generalized S-estimators.

Finally, we have also applied the explosion rate to quantify the loss of robustness when using regression GS-estimators instead of non-generalized S-estimators. Given a score function, the corresponding GS- and S-estimators have the same order of global robustness (finite explosion rate). However, there is a cost in terms of robustness since the explosion rate ranges from approximately 1 (linear score function) to roughly 2 (jump score function). The results obtained in this paper suggest that the robustness loss caused by using pairwise differences of residuals instead of the residuals themselves increases with the *flatness* at zero of the score function.

Acknowledgements: The author wish to thank two referees for their helpful comments and suggestions.

Appendix. Proofs.

Proof of Theorem 1: First, we will show that, if F is a continuous distribution, the function $f(s) = E_F\chi(X/s)$ is continuous and decreasing for all $s > 0$. Let s_n be a sequence such that $s_n \to \overline{s} > 0$. Since χ has at most a finite number of discontinuities, $\chi(u/s_n) \to \chi(u/\overline{s})$ a.s. with respect to Lebesgue measure. Moreover, $|\chi(u/s_n)| \leq 1$, and we can apply Dominated Convergence Theorem to obtain $f(s_n) \to f(\overline{s})$. It is straightforward to show that f is decreasing.
Now, define $s_n = S[(1-\epsilon)F_0+\epsilon V_n]$, where V_n is the uniform distribution on the interval $(n, 2n)$. Assume that $\sup_n s_n < \infty$ (if this is not the case, $B_S^+(\epsilon) = \infty$). Then, there exists a convergent subsequence, denoted also by s_n. Let $\overline{s} = \lim_{n\to\infty} s_n$. We will show that $\overline{s} = S^+(\epsilon)$ and that $\overline{s} \geq S[(1-\epsilon)F_0 + \epsilon H]$ for any arbitrary distribution H. Since F_0 and V_n are continuous, s_n satisfies

$$(1-\epsilon)^2 E_{F_0\times F_0}\chi\left(\frac{X-Y}{s_n}\right) + 2\epsilon(1-\epsilon)E_{F_0\times V_n}\chi\left(\frac{X-Y}{s_n}\right) + \epsilon^2 E_{V_n\times V_n}\chi\left(\frac{X-Y}{s_n}\right) = b. \tag{18}$$

For large enough n and $u > 0$, we have that $V_n^*(u) \equiv \mathrm{P}_{V_n \times V_n}\{X - Y \le u\} = 1/2 + \int_0^u (n-x)/n^2 dx = 1/2 + u/n[1 - u/(2n)]$. Then, $V_n^*(u) \to 1/2$ as $n \to \infty$ for all $u > 0$. Let $k > 0$ be any (large enough) real number such that $\chi(k/s_n)$ is continuous for all n. Then,

$$\begin{aligned} E_{V_n \times V_n}\chi\left(\frac{X-Y}{s_n}\right) &= 2\int_0^\infty \chi(u/s_n)dV_n^*(u) \\ &\ge 2\chi(k/s_n)[1 - V_n^*(k)] \to \chi(k/\overline{s}), \quad \text{as} \quad n \to \infty. \end{aligned}$$

Letting $k \to \infty$, it follows that $\lim_{n\to\infty} E_{V_n \times V_n}\chi[(X-Y)/s_n] = 1$. In a similar way, it can also be proved that $\lim_{n\to\infty} E_{F_0 \times V_n}\chi[(X-Y)/s_n] = 1$. Therefore, taking limits in (18) as $n \to \infty$, yields

$$(1-\epsilon)^2 E_{F_0 \times F_0}\chi\left(\frac{X-Y}{\overline{s}}\right) + 2\epsilon(1-\epsilon) + \epsilon^2 = b. \tag{19}$$

This proves that $\overline{s} = S^+(\epsilon)$. Now let H be an arbitrary distribution, denote $s_H = S[(1-\epsilon)F_0 + \epsilon H]$ and assume to find a contradiction that

$$(1-\epsilon)^2 E_{F_0 \times F_0}\chi\left(\frac{X-Y}{s_H}\right) + 2\epsilon(1-\epsilon) + \epsilon^2 < b.$$

By continuity, there exists $s^* < s_H$ such that

$$\begin{aligned} (1-\epsilon)^2 E_{F_0 \times F_0}\chi\left(\frac{X-Y}{s^*}\right) &+ 2\epsilon(1-\epsilon)E_{F_0 \times H}\chi\left(\frac{X-Y}{s^*}\right) + \epsilon^2 E_{H \times H}\chi\left(\frac{X-Y}{s^*}\right) \le \\ (1-\epsilon)^2 E_{F_0 \times F_0}\chi\left(\frac{X-Y}{s^*}\right) &+ 2\epsilon(1-\epsilon) + \epsilon^2 < b, \end{aligned}$$

but this contradicts the definition of s_H. Therefore,

$$(1-\epsilon)^2 E_{F_0 \times F_0}\chi\left(\frac{X-Y}{s_H}\right) + 2\epsilon(1-\epsilon) + \epsilon^2 \ge b.$$

Comparing the last inequality with (19), it follows that $\overline{s} \ge s_H$ and therefore $\overline{s} = B_S^+(\epsilon)$. To find the implosion maxbias, define $s_0 = S^-(\epsilon) = m^{-1}(b, \epsilon)$. Therefore, $m(s_0, \epsilon) = b$. Since $m(\cdot, \epsilon)$ is decreasing, it is enough to show that $m(s_H, \epsilon) \le b$. To find a contradiction, assume that $m(s_H, \epsilon) > b$. By continuity, there exists $s^* > s_H$ such that $m(s^*, \epsilon) > b$ and therefore

$$\begin{aligned} (1-\epsilon)^2 E_{F_0 \times F_0}\chi\left(\frac{X-Y}{s^*}\right) &+ 2\epsilon(1-\epsilon)E_{F_0 \times H}\chi\left(\frac{X-Y}{s^*}\right) + \\ \epsilon^2 E_{H \times H}\chi\left(\frac{X-Y}{s^*}\right) &\ge m(s^*, \epsilon) > b. \end{aligned} \tag{20}$$

The first inequality follows since F_0 symmetrical and unimodal implies that $E_{F_0 \times H}\chi[(X-Y)/s] \geq E_{F_0}\chi(X/s)$, for all $s > 0$. Notice that (20) contradicts the definition of s_H.

Proof of Theorem 2: (a) For $\delta > 0$ and $s > 0$, define the sets $A_{\delta,s} = \{(x,y) \in \mathbb{R}^2 : 0 < x - y < \delta s\}$ and $B_{\delta,s} = \{(x,y) \in \mathbb{R}^2 : \delta s < x - y\}$. Since κ is the local order at zero of χ, there exists $\delta > 0$ such that the following expansion is valid for all $s > 0$ and $(x,y) \in A_{\delta,s}$

$$\chi\left(\frac{x-y}{s}\right) = \frac{\chi^{(\kappa)}(0)}{k!}\left(\frac{x-y}{s}\right)^{\kappa} + o\left(\left(\frac{x-y}{s}\right)^{\kappa}\right).$$

Then,

$$\begin{aligned} g(s) &= 2\iint_{\{x>y\}} \chi\left(\frac{x-y}{s}\right) f_0(x) f_0(y) dx dy = 2\iint_{A_{\delta,s}} \frac{\chi^{(\kappa)}(0)}{k!}\left(\frac{x-y}{s}\right)^{\kappa} f_0(x) f_0(y) dx dy \\ &+ 2\iint_{A_{\delta,s}} o\left(\left(\frac{x-y}{s}\right)^{\kappa}\right) f_0(x) f_0(y) dx dy + 2\iint_{B_{\delta,s}} \chi\left(\frac{x-y}{s}\right) f_0(x) f_0(y) dx dy \\ &\equiv 2[I_1(\delta,s) + I_2(\delta,s) + I_3(\delta,s)]. \end{aligned}$$

Therefore, $s^{\kappa} g(s) = 2s^{\kappa}[I_1(\delta,s) + I_2(\delta,s) + I_3(\delta,s)]$. Since $\lim_{s\to\infty} A_{\delta,s} = \{(x,y) \in \mathbb{R}^2 : x > y\}$, we have that $\lim_{s\to\infty} 2s^{\kappa} I_1(\delta,s) = C$, where

$$C \equiv \left[\frac{\chi^{(\kappa)}(0)}{k!} E_{F_0 \times F_0}|X-Y|^{\kappa}\right]^{1/\kappa}. \tag{21}$$

Moreover, by Dominated Convergence Theorem $\lim_{s\to\infty} s^{\kappa} I_2(\delta,s) = 0$. Finally,

$$\lim_{s\to\infty} s^{\kappa} I_3(\delta,s) \leq \lim_{s\to\infty} s^{\kappa} \mathrm{P}_{F_0 \times F_0}\{X - Y > \delta s\} = \lim_{s\to\infty} \int_{-\infty}^{\infty} s^{\kappa}[1 - F_0(y+\delta s)] f_0(y) dy$$

The assumption $\lim_{s\to\infty} s^{\kappa+1} f_0(s) = 0$ implies that $\lim_{s\to\infty} s^{\kappa}[1 - F_0(y+\delta s)] = 0$. Then, by Dominated Convergence Theorem $\lim_{s\to\infty} s^{\kappa} I_3(\delta,s) = 0$. Therefore, we have shown that $\lim_{s\to\infty} s^{\kappa} g(s) = C$. We obtain the first approximation of (7) by letting $s = S^{+}(\epsilon)$ in this expression. The second part of (7) follows from

$$C\left(\frac{(1-\epsilon)^2}{b - 2\epsilon + \epsilon^2}\right)^{1/\kappa} = C\left(\frac{(1-\epsilon)^2}{\epsilon - (1+\sqrt{1-b})}\right)^{1/\kappa} (\epsilon^{+} - \epsilon)^{-1/\kappa} \sim \tilde{C}(\epsilon^{+} - \epsilon)^{-1/\kappa}.$$

(b) It remains to show the approximation (9). Since $\lim_{\epsilon\to\sqrt{1-b}} S^-(\epsilon) = 0$, by L'Hôpital's Rule it is enough to compute $\lim_{\epsilon\to\sqrt{1-b}} \frac{dS^-(\epsilon)}{d\epsilon} \equiv -D$. By definition of $S^-(\epsilon)$, we have

$$(1-\epsilon)^2 g[S^-(\epsilon)] + 2\epsilon(1-\epsilon)h[S^-(\epsilon)] = b.$$

Differentiating in this equation with respect to ϵ and taking limits as $\epsilon \to \sqrt{1-b}$, we have

$$\lim_{\epsilon\to\sqrt{1-b}} \frac{dS^-(\epsilon)}{d\epsilon} = \frac{2\sqrt{1-b}}{(1-\sqrt{1-b})^2 g'(0) + 2\sqrt{1-b}(1-\sqrt{1-b})h'(0)}. \tag{22}$$

Now, notice that $h(s) = 2\left[\int_0^{cs} \chi(y/s) f_0(y)dy + 1 - F_0(cs)\right]$. In order to compute $h'(s)$ we apply Leibniz's formula and, after that, use the change of variable $x = y/s$. We obtain $h'(s) = -2\int_0^c \chi'(x)xf_0(sx)dx$. Evaluating at $s = 0$ and integrating by parts, $h'(0) = -2f_0(0)\int_0^c[1-\chi(y)]dy$. Analogously, it can be proved that $g'(0) = -2f_0^*(0)\int_0^c[1-\chi(y)]dy$. Replacing these expressions in the right hand side of (22) we obtain $\lim_{\epsilon\to\sqrt{1-b}} \frac{dS^-(\epsilon)}{d\epsilon} = -D$, where

$$D \equiv \left[\left(f_0^*(0)\frac{(1-\sqrt{1-b})^2}{\sqrt{1-b}} + 2f_0(0)(1-\sqrt{1-b})\right)\int_0^c (1-\chi(x))dx\right]^{-1} \tag{23}$$

Proof of Theorem 3: Notice that $\mathbf{T}_1$ and $\mathbf{T}_2$ have the same breakdown point $\epsilon^* = \sqrt{1-b_1} = 1-\sqrt{1-b_2}$. Let S_1 and S_2 be the dispersion GS-estimators based on the same score functions as $\mathbf{T}_1$ and $\mathbf{T}_2$. From Theorem 2 and Equation (15),

$$[1+B^2_{\mathbf{T}_1}(\epsilon)]^{1/2} \sim -\frac{B^+_{S_1}(\sqrt{1-b_1})}{D_1\cdot(\epsilon-\sqrt{1-b_1})}, \quad \text{as} \quad \epsilon\to\epsilon^*,$$

and

$$[1+B^2_{\mathbf{T}_2}(\epsilon)]^{1/2} \sim \frac{C_2\cdot[(1-\epsilon)^2(b_2-2\epsilon+\epsilon^2)^{-1}]^{1/\kappa}}{B^-_{S_2}(1-\sqrt{1-b_2})}, \quad \text{as} \quad \epsilon\to\epsilon^*.$$

Since $B^+_{S_1}(\sqrt{1-b_1}) < \infty$ and $B^-_{S_2}(1-\sqrt{1-b_2}) > 0$, to show that

$$r(\mathbf{T}_1,\mathbf{T}_2) = \lim_{\epsilon\to\epsilon^*} \frac{[1+B^2_{T_1}(\epsilon)]^{1/2}}{[1+B^2_{T_2}(\epsilon)]^{1/2}} = \infty,$$

it is enough to prove that

$$\lim_{\epsilon\to\epsilon^*}(\epsilon-\sqrt{1-b_1})\left[\left(\frac{b_2-2\epsilon+\epsilon^2}{(1-\epsilon)^2}\right)\right]^{-1/\kappa}=0,$$

but this fact can be checked easily applying L'Hôpital's Rule.

Proof of Theorem 4: Let S^* and S be the dispersion GS- and S-estimators based on the same score functions as $\mathbf{T}^*$ and $\mathbf{T}$ respectively. By (15), it holds that $r(\mathbf{T}^*,\mathbf{T}) = r^+(S^*,S)/r^-(S^*,S)$, so that we will compute $r^+(S^*,S)$ and $r^-(S^*,S)$ separately. From (7),

$$B^+_{S^*}(\epsilon)\sim C^*\left(\frac{(1-\epsilon)^2}{3/4-2\epsilon+\epsilon^2}\right)^{1/\kappa}, \qquad \text{as } \epsilon\to 1/2,$$

where C^* is given by (8). Moreover, Theorem 1(a) in Berrendero *et al.* (1998) implies that

$$B^+_{S}(\epsilon)\sim C\left(\frac{1-\epsilon}{1/2-\epsilon}\right)^{1/\kappa}, \qquad \text{as } \epsilon\to 1/2,$$

where $C=[\chi^{(\kappa)}(0)E_{F_0}|X|^\kappa]^{1/\kappa}/\kappa!$. Then,

$$r^+(S^*,S)=\frac{C^*}{C}\lim_{\epsilon\to 1/2}\frac{\left(\frac{(1-\epsilon)^2}{3/4-2\epsilon+\epsilon^2}\right)^{1/\kappa}}{\left(\frac{(1-\epsilon)}{1/2-\epsilon}\right)^{1/\kappa}}=2^{-1/\kappa}\frac{C^*}{C}=2^{\frac{\kappa-2}{2\kappa}}\left(\frac{\chi^{(\kappa)}_{c^*}(0)}{\chi^{(\kappa)}_{c}(0)}\right)^{1/\kappa}, \tag{24}$$

since, under the Gaussian model, it is easy to check that $C^*/C=\sqrt{2}\left(\chi^{(\kappa)}_{c^*}(0)/\chi^{(\kappa)}_{c}(0)\right)^{1/\kappa}$. On the other hand, from (9),

$$B^-_{S^*}(\epsilon)\sim D^*(1/2-\epsilon), \qquad \text{as } \epsilon\to 1/2, \tag{25}$$

where the constant D^* is given by (10). Under the Gaussian model, (10) reduces to

$$D^*=\frac{1}{\varphi(0)}\left[\left(\frac{2\sqrt{2}+1}{2\sqrt{2}}\right)\int_0^{c^*}[1-\chi_{c^*}(x)]\,dx\right]^{-1}.$$

Furthermore, Theorem 4(b) in Martin and Zamar (1993) implies that the implosion maxbias curve of S satisfies $(1-\epsilon)f[B^-_S(\epsilon)]=1/2$, where $f(s)=E_{F_0}\chi_c(X/s)$. Differentiating with respect to ϵ, evaluating at $\epsilon=1/2$ and rearranging terms, we get

$$\left[\frac{dB^-_S(\epsilon)}{d\epsilon}\right]_{\epsilon=1/2}=-\frac{1}{\varphi(0)}\left[\int_0^{c}[1-\chi_c(x)]\,dx\right]^{-1}\equiv -D$$

and, therefore,

$$B_S^-(\epsilon) \sim D(1/2-\epsilon), \qquad \text{as } \epsilon \to 1/2. \tag{26}$$

From (25) and (26),

$$r^-(S^*,S) = \frac{D^*}{D} = \frac{2\sqrt{2}}{2\sqrt{2}+1}\frac{\int_0^c [1-\chi_c(x)]\,dx}{\int_0^{c^*}[1-\chi_{c^*}(x)]\,dx}. \tag{27}$$

Finally, the result follows from (24) and (27).

Remark 4: Note that equation (13) in Section 2 is only a particular case (for jump score functions) of (27).

References

Berrendero, J.R., Mazzi, S., Romo, J. and Zamar, R., 1998. On the explosion rate of maximum bias functions. *Canad. J. Statist.*, **26**, 333-351.

Berrendero, J.R. and Zamar, R., 1999. Global robustness of location and dispersion estimates. *Statist. Probab. Lett.*, **44**, 63–72.

Croux, C., Rousseeuw, P.J. and Hössjer, O., 1994. Generalized S-estimators. *J. Amer. Statist. Assoc.* **89**, 1271–1281.

Hössjer, O., 1992. On the optimality of S–estimators. *Statist. Probab. Lett.*, **14**, 413–419.

Hössjer, O. ,Croux, C. and Rousseeuw, P.J., 1994. Asymptotics of generalized S-estimators. *J. Multivariate Anal.*, **51**, 148–177.

Martin, R.D. and Zamar, R., 1993. Bias robust estimation of scale. *Ann. Statist.* **21**, 991–1017.

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

Rousseeuw, P.J. and Croux, C., 1993. Alternatives to the median absolute deviation. *J. Amer. Statist. Assoc.*, **88**, 1273–1283.

Rousseeuw, P.J. and Yohai, V.J., 1984. Robust regression by means of S-estimators. In *Robust and Nonlinear Time Series Analysis*, J. Franke, W. Härdle and R.D. Martin (Eds.), Lecture Notes in Statistics 26, Springer, New-York, 256–272.

Score function	Tuning constant	$r^+(\chi, \chi_h)$	$r^-(\chi, \chi_h)$	Eff
$\chi_1(y) = \min\{\lvert y/c\rvert, 1\}$	0.92	∞	0.989	83.4 %
$\chi_t(y) = \min\{3(y/c)^2 - 3(y/c)^4 + (y/c)^6, 1\}$	0.99	1.182	0.996	82.9 %
$\chi_h(y) = \min\{(y/c)^2, 1\}$	0.68	1	1	82.6 %
$\chi_6(y) = \min\{(y/c)^6, 1\}$	0.53	0	1.005	82.2 %
$\chi_j(y) = I\{\lvert y\rvert \geq 0\}$	0.45	0	1.006	82.2 %

Table 1: Explosion and implosion rates for several dispersion GS-estimators.

Score function	Tuning constant	Explosion rate	Eff
$\chi_1(y) = \min\{\lvert y/c\rvert, 1\}$	0.92	∞	69.26 %
$\chi_t(y) = \min\{3(y/c)^2 - 3(y/c)^4 + (y/c)^6, 1\}$	0.99	1.187	68.38 %
$\chi_h(y) = \min\{(y/c)^2, 1\}$	0.68	1	67.88 %
$\chi_6(y) = \min\{(y/c)^6, 1\}$	0.53	0	67.26 %
$\chi_j(y) = I\{\lvert y\rvert \geq 0\}$	0.45	0	67.14 %

Table 2: Explosion rates for several regression GS-estimators.

Score function	c	c^*	$r^+(S^*, S)$	$r^-(S^*, S)$	$r(\mathbf{T}^*, \mathbf{T})$
$\chi_1(y) = \min\{\lvert y/c\rvert, 1\}$	1.47	0.92	1.13	1.18	0.96
$\chi_t(y) = \min\{3(y/c)^2 - 3(y/c)^4 + (y/c)^6, 1\}$	1.55	0.99	1.57	1.16	1.35
$\chi_h(y) = \min\{(y/c)^2, 1\}$	1.04	0.68	1.53	1.13	1.35
$\chi_6(y) = \min\{(y/c)^6, 1\}$	0.79	0.53	1.88	1.10	1.71
$\chi_j(y) = I\{\lvert y\rvert \geq 0\}$	0.67	0.45	2.12	1.11	1.91

Table 3: Comparison between generalized and non-generalized S-estimators based on the same score function.

Figure 1: Maxbias curves of $\widehat{\boldsymbol{\theta}}_1$ (dashed line) and $\widehat{\boldsymbol{\theta}}_2$ (solid line). For both estimators, the breakdown point is $\epsilon^* = 0.25$.

Figure 2: Explosion and implosion maxbias curves for four dispersion GS-estimators with the following score functions: χ_6 (solid line), Huber function χ_h (dashed line), Tukey function χ_t (dashed-dotted line) and truncated linear function χ_1 (dotted line).