Maxbias Curves of Robust Location Estimators based on Subranges

Christophe Croux

Université Libre de Bruxelles, ECARES
Av. F.D. Roosevelt 50, B-1050 Brussels, Belgium, ccroux@ulb.ac.be

Gentiane Haesbroeck

Department of Economics, University of Liège,
Boulevard du Rectorat 7, B-4000 Liège, Belgium. g.haesbroeck@ulg.ac.be

Abstract

A maxbias curve is a powerful tool to describe the robustness of an estimator. It tells us how much an estimator can change due to a given fraction of contamination. In this paper, maxbias curves are computed for some univariate location estimators based on subranges: midranges, trimmed means and the univariate Minimum Volume Ellipsoid (MVE) location estimators. These estimators are intuitively appealing and easy to calculate.

keywords: Breakdown value–Maxbias Curve–Robustness–Location Estimator

1 Introduction

The most popular location estimator is the sample average, which is however known to be extremely sensitive to outliers. For a given data set $X = \{x_1, \ldots, x_n\}$, let $x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(n)}$ be the sorted observations. The standard robust location estimator is the sample median, defined as

$$\text{median}_n = \frac{1}{2}\left(x_{(\lfloor n/2 \rfloor + 1)} + x_{(\lfloor (n+1)/2 \rfloor)}\right),$$

where $\lfloor z \rfloor$ denotes the largest integer smaller than or equal to z. The sample median has powerful robustness properties, but is sometimes considered to be not sufficiently efficient.

As an alternative to the median, one could consider the α-midrange

$$M_{n,\alpha} = \frac{1}{2}\left(x_{(\lfloor \alpha n \rfloor+1)} + x_{(n-\lfloor \alpha n \rfloor)}\right), \tag{1.1}$$

with $0 \leq \alpha \leq 0.5$. Taking $\alpha = 0.5$ returns the median again. If instead of taking the midpoint of the interval $I_\alpha = [x_{(\lfloor \alpha n \rfloor+1)}; x_{(n-\lfloor \alpha n \rfloor)}]$, we compute the average over the observations lying in I_α, then we get the α-trimmed mean

$$T_{n,\alpha} = \frac{1}{n - 2\lfloor n\alpha \rfloor} \sum_{i=\lfloor \alpha n \rfloor+1}^{n-\lfloor \alpha n \rfloor} x_{(i)}.$$

The trimming proportion equals (approximately) 2α, with $0 < \alpha < \frac{1}{2}$. The choice $\alpha = 0$ yields the usual mean.

The previous location estimators rely on the interval I_α. Another location estimator based on a subrange is the univariate version of the Minimum Volume Ellipsoid (MVE). The MVE estimator was introduced by [9] and is routinely used for estimating robust multivariate location and scale. But also in a univariate setup it might be useful. The univariate location MVE estimator is obtained by looking for the shortest interval containing at least $(1-\alpha) \times 100$ percent of the observations and then computing the midpoint of it. It is sometimes called the Shorth estimator [1].

A competitor for the MVE is the Minimum Covariance Determinant (MCD) estimator. The univariate MCD estimator is obtained by looking for the interval containing at least $(1 - \alpha) \times 100$ percent of the observations and having the minimal value of the variance computed over the observations belonging to the interval. Afterwards the average over the observations belonging to the selected interval is computed and returns the MCD location estimator. Taking $\alpha = 0.5$ yields the maximum breakdown value, and another typical value is $\alpha = 0.25$. Both estimators can be easily computed ([11], page 172) in the univariate case.

All the above estimators have an accompanying scale estimator. If we compute the range or the standard deviation over the observations belonging to the interval I_α, we obtain the α-range and the α-trimmed standard deviation. The length of the shortest interval in the definition of the MVE yields a dispersion measure and the MCD-scale equals the standard deviation over the observations in the selected interval. Maxbias curves of these scale estimators have been derived in [5]. This paper complements this study by looking at the maxbias of the location counterparts of these estimators.

In Section 2, the maxbias curve is defined and functional representations of the considered estimators are given. Section 3 contains mathematical expressions for the maxbias curves and compares them. Some conclusions are made in Section 4, where we also make a comparison with M-estimators of location.

2 Maxbias curves

There exist several measures of robustness of an estimator (cfr. [6]), but in this paper the **maxbias curve** will be used. A maxbias curve gives the maximal bias that an estimator can suffer from when a fraction ε of the data come from a contaminated distribution (while the other $(1-\varepsilon) \times 100\%$ of the data follow the model distribution). By letting ε vary between zero and the breakdown value (which is the highest fraction of contamination that an estimator can withstand before becoming degenerate) a curve is obtained. A survey on maxbias curves is given by [10]. A maxbias curve is an asymptotic concept and requires a functional representation of the estimator.

Suppose that the distribution generating the "good" observations belongs to a location-scale family

$$\{F_{\mu,\sigma}(x) := F(\frac{x-\mu}{\sigma}) | -\infty < \mu < \infty, 0 < \sigma < \infty\}.$$

Throughout the paper, the central model distribution F is supposed to satisfy the following property

(F) F has a strictly positive and continuous unimodal density f which is symmetric about the origin.

Let T denote a statistical functional representing a location estimator. All the location functionals considered in this paper are affine equivariant (meaning that $T(aX+b) = aT(X)+b$ where $T(X) \equiv T(G)$ whenever $X \sim G$) and Fisher consistent at the specified model, thus $T(F_{\mu,\sigma}) = \mu$. Therefore there will be no loss of generality to restrict attention to the central model distribution F. Define the contamination neighborhood of F

$$\mathcal{F}_\varepsilon = \{G; G = (1-\varepsilon)F + \varepsilon H;\ H \text{ any distribution}\} \tag{2.1}$$

for a given fraction of contamination ε $(0 \le \varepsilon \le 1)$. The maxbias curve is then defined by

$$B(\varepsilon; T, F) = \sup_{G \in \mathcal{F}_\varepsilon} |T(G)|. \tag{2.2}$$

The breakdown value is now obtained as

$$\varepsilon^*(T,F) = \inf\{\varepsilon > 0 | B(\varepsilon, T, F) = \infty\}. \tag{2.3}$$

The functional associated to the α-midrange is given by

$$M_\alpha(G) = \frac{1}{2}\{G^{-1}(\alpha) + G^{-1}(1-\alpha)\},$$

with $G^{-1}(\beta) = \inf\{t | G(t) > \beta\}$ for any $0 \leq \beta \leq 1$, while the functional representation of the α-trimmed mean is given by

$$T_\alpha(G) = \frac{1}{1-2\alpha}\int_\alpha^{1-\alpha} G^{-1}(t)\, dt \tag{2.4}$$

where $0 < \alpha < \frac{1}{2}$.

For the MVE_α location estimator, we first consider intervals containing $(1-\alpha)$ of the mass of a distribution G. They take the form

$$I_{G,x} = [x - H_G(x), x + H_G(x)] \tag{2.5}$$

where

$$H_G(x) = \inf\{s > 0 | P_G(|X - x| \leq s) > 1 - \alpha\}. \tag{2.6}$$

We can see $I_{G,x}$ as the smallest interval with center x covering $(1-\alpha)$ of the mass of G. Therefore we get

$$T_{\mathrm{MVE}_\alpha}(G) = \operatorname*{argmin}_x |I_{G,x}| = \operatorname*{argmin}_x H_G(x). \tag{2.7}$$

The MCD_α location functional is defined in a similar way.

The derivation of the maxbias curve of an estimator is not always an easy matter. It is often admitted that the maximal bias of a location functional is produced by taking for H in (2.1) a point mass at infinity. However, the most unfavorable or worst contaminating distribution H is not necessarily of that type and in principle it doesn't even need to be a Dirac measure.

3 Maxbias curves of the location estimators

In this section, the maxbias curves of the α-midrange, α-trimmed mean and the MVE location estimators are compared. Propositions 1 and 2 give the maxbias curves for the

α-midrange and α-trimmed mean. Both M_α and T_α belong to the class of L-estimators, for which maxbias curves were already derived in ([7], page 59) but using Levy neighborhoods instead of contamination neighborhoods. The proofs of the propositions are kept for the Appendix.

Proposition 1. The maxbias curve of the α-midrange estimator at F is given by

$$B(\varepsilon; M_\alpha, F) = \begin{cases} \frac{1}{2}\left(F^{-1}(\frac{\alpha}{1-\varepsilon}) + F^{-1}(\frac{1-\alpha}{1-\varepsilon})\right) & \text{if } \varepsilon < \alpha \\ \infty & \text{otherwise} \end{cases} \tag{3.1}$$

for $0 < \alpha < \frac{1}{2}$.

Proposition 2. The maxbias curve of the α-trimmed mean at F is given by

$$B(\varepsilon; T_\alpha, F) = \begin{cases} \frac{1-\varepsilon}{1-2\alpha} \int_{\frac{\alpha}{1-\varepsilon}}^{\frac{1-\alpha}{1-\varepsilon}} F^{-1}(t)\ dt & \text{if } \varepsilon < \alpha \\ \infty & \text{if } \varepsilon > \alpha \end{cases} \tag{3.2}$$

for $0 < \alpha < \frac{1}{2}$.

The behavior of $B(\varepsilon; T_\alpha, F)$ for $\varepsilon = \alpha$ depends on F. If the first moment of F exists, the maxbias curve of the α-trimmed mean is finite at $\varepsilon = \alpha$ (for $0 < \alpha < 0.5$). Otherwise, it is infinite. In both cases the breakdown value equals α. It is quite commonly believed that maxbias curves have an asymptote at their breakdown value, in which case we speak of "regular" explosion of the maxbias curve. But it was already noted in [5] that we don't always have regular explosion. For the trimmed mean we only have regular explosion when the first moment of F fails to exist.

The derivation of the maxbias curves of M_α and T_α is easy, since both estimators are monotone. This means that if a distribution G_1 is stochastically smaller than G_2 (notation $G_1 \overset{s}{\preceq} G_2$), then $T_\alpha(G_1) \leq T_\alpha(G_2)$ and $M_\alpha(G_1) \leq M_\alpha(G_2)$. Recall that

$$G_1 \overset{s}{\preceq} G_2 \Leftrightarrow G_1(x) \geq G_2(x) \text{ for all } x.$$

Since for every $G \in \mathcal{F}_\varepsilon$, $G \overset{s}{\preceq} (1-\varepsilon)F + \varepsilon\Delta_\infty$, where Δ_x is the Dirac measure putting all its mass on x, it follows immediately that the maximal bias will be obtained by contaminating F with a point mass at "infinity" The MVE and MCD location estimators however are not monotone, making the derivation of their maxbias curves much harder.

The next proposition gives a formal proof of the maxbias curve of the location MVE_α estimator, which has not appeared yet in the literature to our knowledge. Define for all $x \in I\!R$, $H_\varepsilon^+(x)$ as the solution s of the equation

$$F(x+s) - F(x-s) = \frac{1-\alpha}{1-\varepsilon} \tag{3.3}$$

and $H_\varepsilon^-(x)$ as the solution s of the equation

$$F(x+s) - F(x-s) = \frac{1-\alpha-\varepsilon}{1-\varepsilon} \tag{3.4}$$

for $0 \le \varepsilon \le 1-\alpha$.

Proposition 3. The maxbias $B(\varepsilon; T_{\text{MVE}_\alpha}, F)$ for $0 < \alpha \le \frac{1}{2}$, is given by the positive solution b of the equation

$$H_\varepsilon^-(b) = H_\varepsilon^+(0) \tag{3.5}$$

when $\varepsilon < \alpha$ and equals $+\infty$ otherwise. The functions H_ε^+ and H_ε^- are defined as in (3.3) and (3.4).

An interesting feature of the maxbias of the MVE location estimator is that the most unfavorable contaminating distribution is a point mass distribution $H = \Delta_z$ with z not at infinity but much closer to the center of the distribution (see proof of Proposition 3). In Figure 1, $T_{\text{MVE}_\alpha}((1-\varepsilon)F + \varepsilon\Delta_z)$ is plotted as a function of z. One sees that contamination far from the center of F leads to a zero bias, showing the redescending character of MVE. Less extreme outliers, however, can still lead to a considerable bias of T_{MVE_α}.

In Figure 2, the maxbias curves at the Normal model (so by taking for F the standard normal distribution) of the location estimators M_α, T_α and MVE_α for $\alpha = 0.25$ are represented together with the maxbias curve of the median. The α-trimmed mean has a maxbias curve close to that of the median, which has been proved to be the lower bound at any F satisfying condition (F)(see [7], page 74). The maxbias of T_α increases only slightly with ε and stays bounded right up to the breakdown point. This doesn't hold for the α-midrange nor for the MVE location estimator with 25% breakdown point.

The maxbias curve of the MCD location estimator appeared to be much harder to handle and no rigorous proof is available yet. However, for the Normal distribution and using

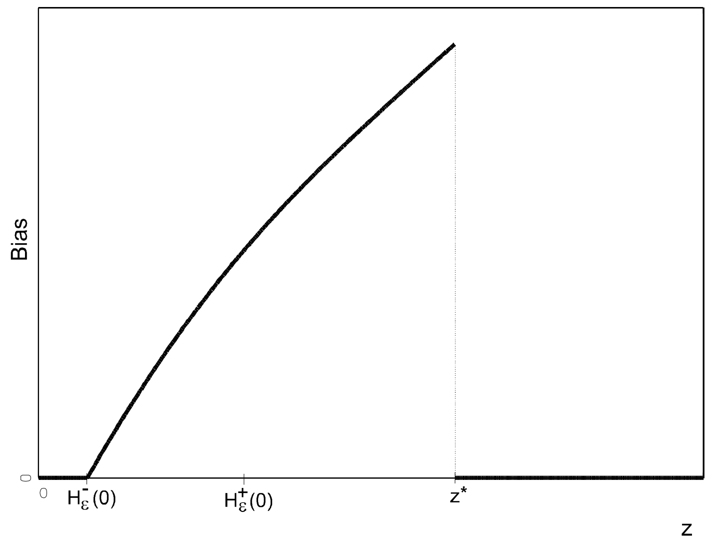

Figure 1: Bias $T_{\mathrm{MVE}_\alpha}((1-\varepsilon)F + \varepsilon\Delta_z)$ of the MVE_α location estimator corresponding to point mass contamination as a function of z for $z \geq 0$ and $\alpha = 0.5$.

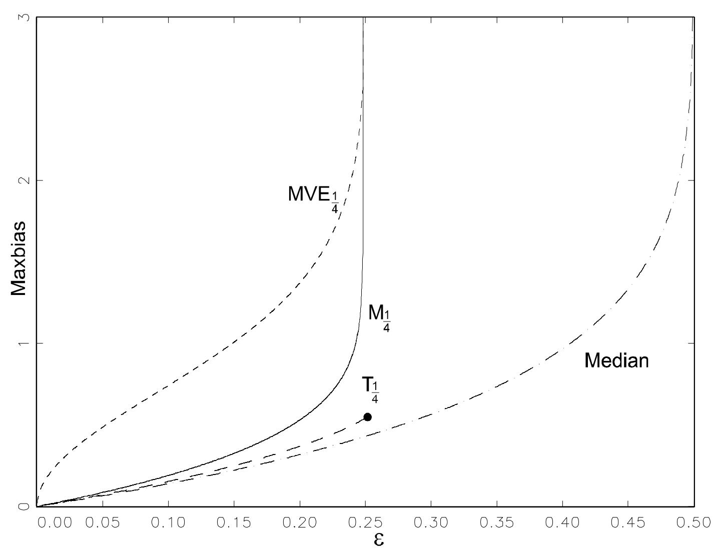

Figure 2: Maxbias curves of the location estimators M_α, T_α and MVE_α for $\alpha = 0.25$, together with the maxbias curve of the median at the normal distribution.

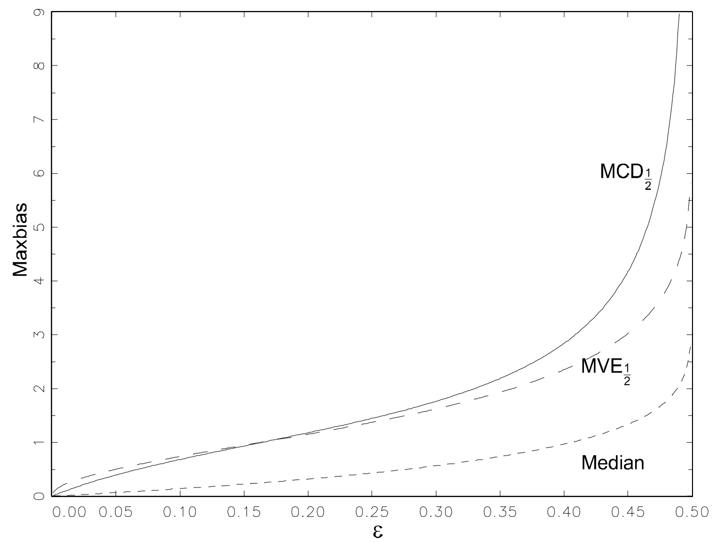

Figure 3: Maxbias curves of the location estimators MVE_α and MCD_α for $\alpha = 0.5$, together with the maxbias curve of the median at the normal distribution.

numerical verifications, an expression for the maxbias curve has been obtained. Let $x^-(z)$ be the solution of $x + H_\varepsilon^-(x) = z$. Define

$$\lambda_b(x) = \frac{1-\varepsilon}{1-\alpha} \int_{x-H_\varepsilon^-(x)}^{x+H_\varepsilon^-(x)} (y-x)^2 \, dF(y) + \frac{\varepsilon}{1-\alpha}(b-x)^2$$

for every $b \in \mathbb{R}$ and let $\tilde{x}(b) = \operatorname*{argmin}_{x \geq x^-(b)} \lambda_b(x)$. The maxbias $B(\varepsilon; T_{\mathrm{MCD}_\alpha}, F)$ with F the Normal distribution, is then given by the solution b of the equation

$$\lambda_b(\tilde{x}(b)) = \frac{1-\varepsilon}{1-\alpha} \int_{-H_\varepsilon^+(0)}^{H_\varepsilon^+(0)} y^2 dF(y)$$

for $\varepsilon \leq \alpha$.

Note that for the scale part a formally correct proof has been given in [5]. Figure 3 gives the maxbias curves of the MVE_α and MCD_α estimators with $\alpha = 0.5$ together with the lower bound. The non-differentiability at $\varepsilon = 0$ of $B(\varepsilon; T_{\mathrm{MVE}_\alpha}, F)$ makes it lie above the maxbias curve of MCD for small values of ε. In the neighborhood of the breakdown value, MCD_α has a significantly larger bias than the MVE_α location estimator for $\alpha = 0.5$.

When comparing robust estimators not only maxbias needs to be taken into account (otherwise the median would always be the optimal choice) but also efficiency, since an estimator needs to be precise when the data are generated according to the model. The asymptotic variances of M_α and T_α are easily obtained since they are L-estimators (e.g. [6,

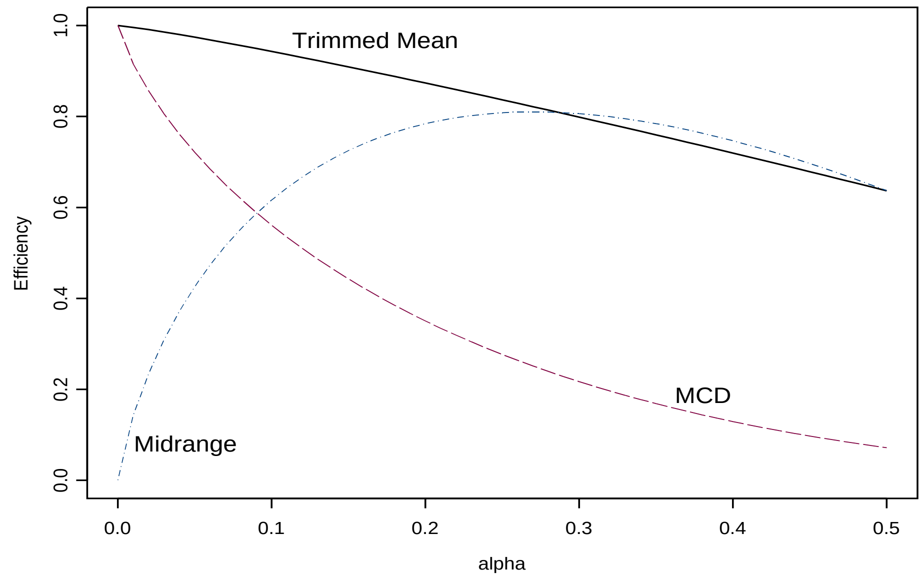

Figure 4: Gaussian efficiencies of the location estimators M_α (dashed and dotted line), T_α (solid line) and MCD_α (dashed line) as a function of α.

page 108]). For the MCD_α we used the results of [3]. The MVE_α has a slower than normal convergence [1, page 50], and therefore one could say that it has zero efficiency. In Figure 4 the Gaussian efficiencies of the location estimators M_α, T_α and MCD_α are represented as a function of α. The limit case $\alpha = 0.5$ for M_α and T_α corresponds to the efficiency of the median. The trimmed mean is always more efficient than the median, while the α-midrange beats the median for a large range of values for the trimming proportion (i.e. for $\alpha > 0.108$). The other limit case, $\alpha = 0$ corresponds with the sample average for T_α and MCD_α. The efficiency of MCD_α is rather dissapointing. We only do better than the median for for trimming proportions smaller than 8%, in which case the MCD will have a very low breakdown point. We conclude from Figure 4 that trimmed means have very good efficiency properties, while we already saw from Figure 2 that also their bias behavior is very reasonable.

4 Conclusion

The robustness of an estimator is most frequently measured by its breakdown point. However, even if two estimators have identical breakdown values, they may behave differently to given amounts of contamination. The maxbias curve helps us to compare more thoroughly the robustness of the different estimators. In this paper, maxbias curves have been computed for some simple and frequently used univariate location estimators: the α-range, α-trimmed mean and MVE_α (which all have the same breakdown value for fixed $0 < \alpha < 1/2$).

Section 3 provides support for the 25% trimmed mean as the estimator to use among the considered estimators. For realistic amounts of contamination ($\leq 25\%$), its maxbias curve is nearly optimal. Its Gaussian efficiency is fairly high (84% versus 64% for the median), it is widely known and its definition is easy to understand. More support for the use of trimmed means can be found in [4].

How do they compare to Huber's M-estimators? A location M-estimator is defined as the solution t_n of the equation

$$\frac{1}{n}\sum_{i=1}^{n}\psi\left(\frac{x_i - t_n}{s_n}\right) = 0$$

where ψ is a strictly increasing, continuous, odd function and s_n an auxiliary scale estimator. As ψ function we take Huber's $\psi(u) = \min(c, \max(u, -c))$ with the tuning constant c set to have a 95% Gaussian efficiency. It is recommended to use a maximal breakdown point scale estimator for s_n, and the $\text{MVE}_{0.5}$ is particularly well suited as auxiliary scale estimator (see [2]). This guarantees that the location estimator will also have 50% breakdown value, as long as ψ is bounded. The attained efficiency remains the same as if we would have estimated scale and location jointly, like in Huber's proposal 2 ([7, page 137]). We computed the maxbias curves of these M-estimators using results of [8]. From Figure 5 we see that the 95% efficient Huber M-estimator has a larger bias than the 25% trimmed mean for the more realistic amounts of contamination (up 25%). Taking instead an M-estimator with 84% efficiency, the same efficiency as $T_{0.25}$, gives a maxbias curve very close to that of the trimmed mean (upto 25%), with slight advantage for the Huber estimator. A drawback of an M-estimators, however, remains its implicit definition, which requires the use of an numerical algorithm to compute it.

There still remains theoretical work to be done for the maxbias of multivariate location and scatter estimators. Maxbias curves for the multivariate MVE-scale estimator under the

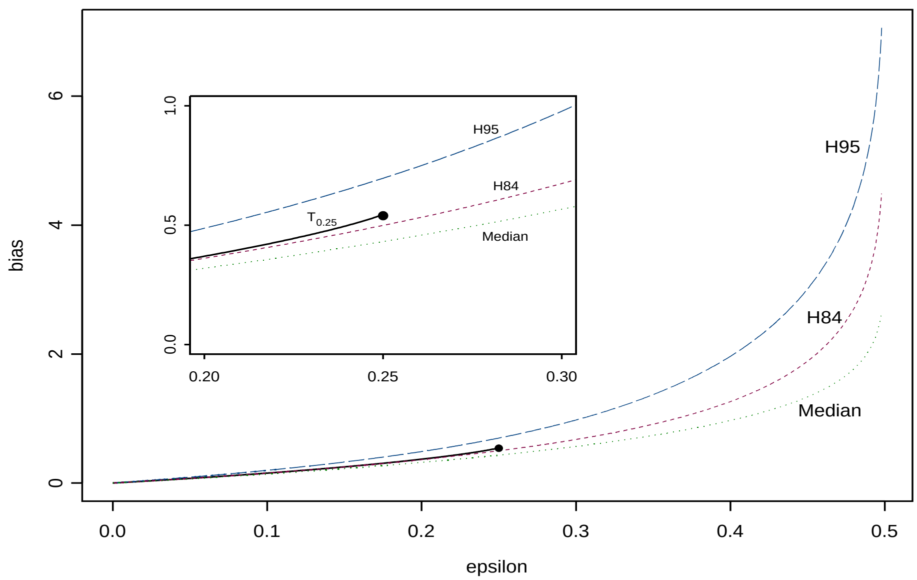

Figure 5: Maxbias curves of the 25%-trimmed mean (solid line), Huber's M-estimator with 95% efficiency (long dashes) and with 84% efficiency (short dashes), together with the maxbias curve of the median (dotted line) at the normal distribution. The inserted window shows a zoom of the graph close to the breakdown point of the 25%-trimmed mean.

assumption of a fixed location were computed by [12]. For the location MVE and MCD estimators however, no results seem to have been obtained.

Acknowledgment: The authors whish to thank Peter Rousseeuw for suggesting a comparison with the Huber M-estimator.

5 Appendix

Proof of Proposition 5: First assume that $\varepsilon < \alpha$ and let c denote the right hand side of (3.1). Let $G = (1-\varepsilon)F + \varepsilon H$, where H can be any distribution. We may assume without loss of generality that $M_\alpha(G) > 0$. We will prove that $M_\alpha(G) \leq c$. As $G^{-1}(\alpha) \leq F^{-1}(\frac{\alpha}{1-\varepsilon})$ and $G^{-1}(1-\alpha) \leq F^{-1}(\frac{1-\alpha}{1-\varepsilon})$, it follows that

$$M_\alpha(G) = \frac{1}{2}\left(G^{-1}(\alpha) + G^{-1}(1-\alpha)\right) \leq \frac{1}{2}\left(F^{-1}(\frac{\alpha}{1-\varepsilon}) + F^{-1}(\frac{1-\alpha}{1-\varepsilon})\right) = c.$$

Now consider the sequence of distribution functions $G_n = (1-\varepsilon)F + \varepsilon\Delta_{x_n}$ where $x_n \uparrow +\infty$. We claim that $\lim_{n\to+\infty} M_\alpha(G_n) = c$. By taking $x_n > F^{-1}(\frac{1-\alpha}{1-\varepsilon}), G_n^{-1}(\alpha) = F^{-1}(\frac{\alpha}{1-\varepsilon})$ and $G_n^{-1}(1-\alpha) = F^{-1}(\frac{1-\alpha}{1-\varepsilon})$. So, $\lim_{n\to+\infty} M_\alpha(G_n) = c$. Hence, $\sup_{G\in\mathcal{F}_\varepsilon} |M_\alpha(G)| = c$ for $\varepsilon < \alpha$.

If $\varepsilon \geq \alpha$, the contaminated distribution G_n leads to an infinite bias when n is sufficiently large. Taking $x_n > |F^{-1}(\frac{1-\alpha-\varepsilon}{1-\varepsilon})|$ gives $G_n^{-1}(1-\alpha) = x_n$ and $G_n^{-1}(\alpha) = F^{-1}(\frac{\alpha}{1-\varepsilon})$. So, for any $x_n > |F^{-1}(\frac{1-\alpha-\varepsilon}{1-\varepsilon})|$, the functional M_α computed at G_n is given by

$$M_\alpha(G_n) = \frac{1}{2}\left(F^{-1}(\frac{\alpha}{1-\varepsilon}) + x_n\right)$$

which goes to infinity as n increases. 2

Proof of Proposition 6: First assume that $\varepsilon < \alpha$ and let c denote the right hand side of (3.2). Let $G = (1-\varepsilon)F + \varepsilon H$, where H can be any distribution. One may assume without loss of generality that $T_\alpha(G) \geq 0$. We will prove that $T_\alpha(G) \leq c$. As $G^{-1}(t) \leq F^{-1}(\frac{t}{1-\varepsilon})$, it follows that

$$T_\alpha(G) = \frac{1}{1-2\alpha}\int_\alpha^{1-\alpha} G^{-1}(t)\ dt \leq \frac{1}{1-2\alpha}\int_\alpha^{1-\alpha} F^{-1}(\frac{t}{1-\varepsilon})\ dt = c.$$

Now consider the sequence of distribution functions $G_n = (1-\varepsilon)F + \varepsilon\Delta_{x_n}$ where $x_n \uparrow +\infty$. By taking $x_n > F^{-1}(\frac{1-\alpha}{1-\varepsilon}), G_n^{-1}(t) = F^{-1}(\frac{t}{1-\varepsilon})$ for all $0 < t \leq 1-\alpha$. So, $\lim_{n\to+\infty} T_\alpha(G_n) = \frac{1}{1-2\alpha}\int_\alpha^{1-\alpha} F^{-1}(\frac{t}{1-\varepsilon})\ dt = c$. Hence, $\sup_{G\in\mathcal{F}_\varepsilon} |T_\alpha(G)| = c$ for $\varepsilon < \alpha$.

If $\varepsilon > \alpha$, the contaminated distribution G_n leads to an infinite bias for n sufficiently large. Indeed, for any $x_n > |F^{-1}(\frac{1-\alpha-\varepsilon}{1-\varepsilon})|$, $G_n^{-1}(1-\alpha) = x_n$ and $G_n^{-1}(\alpha) = F^{-1}(\frac{\alpha}{1-\varepsilon})$. So that

$$T_\alpha(G_n) = \frac{1-\varepsilon}{1-2\alpha}\int_{F^{-1}(\frac{\alpha}{1-\varepsilon})}^{x_n} y\,dF(y) + \frac{1-\alpha-(1-\varepsilon)F(x_n)}{1-2\alpha}\,x_n,$$

which tends to infinity since the second term of the right-hand side of the above equation does. 2

Proof of Proposition 3: Let $\varepsilon < \alpha$ and consider first the maxbias over point contaminated distributions. Let $z > 0$ and denote $G_z = (1-\varepsilon)F + \varepsilon\Delta_z$. Let $x^+(z)$ be the solution of $x + H_\varepsilon^+(x) = z$ and $x^-(z)$ be the solution of $x + H_\varepsilon^-(x) = z$. Of course, $x^+(z) \le x^-(z)$. Denote the MVE location functional T_{MVE_α} by T. For computing $T(G_z)$, the MVE location estimator at G_z, definition (2.7) shows that one needs to find the minimum of $H_{G_z}(x)$. Recall that $I_{G_z,x}$ was defined by (2.5). For any x, $H_{G_z}(x)$ can be determined as follows:

(1) If $x < x^+(z)$, then the point contamination z will not belong to $I_{G_z,x}$ and therefore $H_{G_z}(x) = H_\varepsilon^+(x)$.

(2) If $x^+(z) \le x \le x^-(z)$, then z will be the right endpoint of $I_{G_z,x}$ and $H_{G_z}(x) = z - x$.

(3) If $x > x^-(z)$, the contaminating point z is inside the interval $I_{G_z,x}$ implying $H_{G_z}(x) = H_\varepsilon^-(x)$.

Note that $H_{G_z}(x)$ is a continuous function. Depending on the values of z, it is now easy to find the minimum of $H_{G_z}(x)$. Herefore, one uses the properties of $H_\varepsilon^-(x)$ and $H_\varepsilon^+(x)$ which are symmetric functions strictly increasing on the positive half-side, and the fact that $H_\varepsilon^-(x) < H_\varepsilon^+(x)$ for any x (Here we use condition (F) on the model distribution).

(a) If $z < H_\varepsilon^-(0)$, then $x^+(z) \le x^-(z) \le 0$ and clearly $T(G_z) = \text{argmin}_x H_{G_z}(x) = 0$.

(b) If $H_\varepsilon^-(0) \le z \le H_\varepsilon^+(0)$, then $x^+(z) \le 0 \le x^-(z)$ and $T(G_z) = x^-(z)$.

(c) If $z > H_\varepsilon^+(0)$, then $x^-(z) \ge x^+(z) \ge 0$. The minimum of $H_{G_z}(x)$ equals then or $H_\varepsilon^+(0)$ (attained at 0) or $H_\varepsilon^-(x^-(z))$ (attained at $x^-(z)$). Note that $H_\varepsilon^-(x^-(z))$ is strictly increasing in z for $z \ge H_\varepsilon^-(0)$. Define now z^* as the unique positive solution of

$$H_\varepsilon^-(x^-(z)) = H_\varepsilon^+(0), \tag{5.1}$$

for $z \geq H_\varepsilon^-(0)$. We have that $x^-(z^*) = b$, with b defined by (3.5). By definition of $x^-(z)$ we have $H_\varepsilon^-(x^-(H_\varepsilon^+(0))) = H_\varepsilon^+(0) - x^-(H_\varepsilon^+(0)) \leq H_\varepsilon^+(0)$ and therfore $z^* \geq H_\varepsilon^+(0)$.

So, if $H_\varepsilon^+(0) \leq z < z^*$ one has $T(G_z) = x^-(z)$, as in case (b) and for $z > z^*$, we have $T(G_z) = 0$.

Figure 1 in Section 5 summarizes the three previously considered cases by drawing $T(G_z)$. The non-zero part corresponds with the function $x^-(z)$, which is easily seen to be strictly increasing for all z. We clearly see that $\sup_z T(G_z) = x^-(z^*) = b$, implying that the maxbias for point contaminated distributions equals b.

Take now $G = (1-\varepsilon)F + \varepsilon H$, with H arbitrary. The aim is to prove that $|T(G)| \leq b$. Suppose without loss of generality that $T(G) > 0$. [If it is not the case, replace G by $\tilde{G} = (1-\varepsilon)F + \varepsilon\tilde{H}$ where $-X \sim \tilde{H}$ with $X \sim H$. Then, $T(\tilde{G}) = -T(G) > 0$.] Define $z = \sup I(G) = T(G) + H_G(T(G))$; we know that $T(G_z) \leq b$. Consider the three following cases:

Case I: $0 \leq z < H_\varepsilon^-(0)$. In this case, z is in the interior of $I(G_z)$. One has

$$\begin{aligned} P_{G_z}(I(G)) &= (1-\varepsilon)P_F(I(G)) + \varepsilon \qquad \text{since } z \in I(G) \\ &\geq (1-\varepsilon)P_F(I(G)) + \varepsilon P_H(I(G)) \\ &\geq P_G(I(G)) \geq 1-\alpha, \end{aligned}$$

and $I(G)$ is thus an interval containing mass $(1-\alpha)$ of G_z and therefore $|I(G)| \geq |I(G_z)|$ by definition of $\mathrm{MVE}_\alpha(G_z)$. But since we are in Case I and since $T(G) > 0$, we have $|I(G)| \leq 2z < |I(G_z)|$, yielding a contradiction. So Case I is excluded.

Case II: $H_\varepsilon^-(0) \leq z < z^*$. In this case z is at the right endpoint of $I(G_z)$, as can be seen from the previous enumeration. Once again $P_{G_z}(I(G)) \geq 1-\alpha$ and $|I(G_z)| \leq |I(G)|$. Since z is the right endpoint of $I(G)$ as well, one has that the midpoint of $I(G_z)$ must be bigger than the midpoint of $I(G)$, yielding $|T(G)| \leq |T(G_z)| \leq b$.

Case III: $z > z^*$. In this case, z does not belong to $I(G_z)$ and $I(G_z) = [-H_\varepsilon^+(0), H_\varepsilon^+(0)]$. Since $P_G(I(G_z)) \geq (1-\varepsilon)P_F(I(G_z)) = P_{G_z}(I(G_z)) \geq 1-\alpha$, one has $|I(G_z)| \geq |I(G)|$ by definition of $\mathrm{MVE}_\alpha(G)$. Moreover, $P_{G_z}(I(G)) = (1-\varepsilon)P_F(I(G)) + \varepsilon \geq 1-\alpha$ (since $z \in I(G)$), and therefore $|I(G)| \geq |I(G_z)|$ by definition of $\mathrm{MVE}_\alpha(G_z)$. One concludes that $|I(G)| = |I(G_z)| = 2H_\varepsilon^+(0)$ yielding $T(G) = z - H_\varepsilon^+(0)$ since z is the right endpoint of $I(G)$.

Also, $P_{G_z}(I(G)) = (1-\varepsilon)\,(F(z) - F(z-2H_\varepsilon^+(0))) + \varepsilon \geq 1-\alpha$ so that $H_\varepsilon^-(z - H_\varepsilon^+(0)) \leq H_\varepsilon^+(0)$. Or, $H_\varepsilon^-(T(G)) \leq H_\varepsilon^-(b)$ by definition (3.5) of b. Since H_ε^- is increasing on the positive numbers, it follows that $T(G) \leq b$.

If $1-\alpha \geq \varepsilon \geq \alpha$, the contaminated distribution $G_n = (1-\varepsilon)F + \varepsilon\Delta_{z_n}$ with $z_n \uparrow +\infty$ leads to an infinite bias. The point z_n must lie in $I(G_n)$, otherwise $P_{(1-\varepsilon)F+\varepsilon\Delta_{z_n}}(I(G_n)) = (1-\varepsilon)P_F(I(G_n)) < 1-\varepsilon \leq 1-\alpha$. For $z_n > H_\varepsilon^-(0)$, z_n will even be the upper bound of $I(G_n)$. Therefore,

$$T(G_n) = \frac{1}{2}(z_n + F^{-1}(F(z_n) - \frac{1-\alpha-\varepsilon}{1-\varepsilon})).$$
$$\Rightarrow \quad \lim_{n\to+\infty} T(G_n) = \frac{1}{2}(\lim_{n\to+\infty} z_n + F^{-1}(\frac{\alpha}{1-\varepsilon})) = +\infty,$$

yielding an infinite bias for $\varepsilon \geq \alpha$.

For $1-\alpha \leq \varepsilon \leq 1$, we will have that for every sequence $z_n \uparrow +\infty$, $I(G_{z_n}) = \{z_n\}$ yielding $T(G_{z_n}) \to \infty$, and an infinite bias. 2

References

[1] Andrews DF, Bickel PJ, Hampel FR, Huber PJ, Rogers WH, and Tukey JW (1972) Robust estimation of Location: Survey and Advances, Princeton University Press: Princeton NJ.

[2] Berrendero JR, Zamar RH (1999) Global robustness of location and dispersion estimates. Statistics and Probability Letters 44: 63–72.

[3] Butler RW, Davies PL, Jhun M (1993) Asymptotics for the Minimum Covariance Determinant Estimator. The Annals of Statistics 21: 1385–1400.

[4] Croux C (1996) Maximum Deviation curves for location estimators. Statistics 28: 285–305.

[5] Croux C, Haesbroeck G (2001) Maxbias Curves of Robust Scale Estimators Based on Subranges. Metrika, forthcoming.

[6] Hampel FR, Ronchetti EM, Rousseeuw PJ, Stahel WA (1986) Robust Statistics: The Approach based on Influence Functions. Wiley, New York.

[7] Huber PJ (1981) Robust Statistics. Wiley, New York.

[8] Martin RD, Zamar RH (1993) Bias Robust Estimation of Scale. The Annals of Statistics 21: 1991–1017.

[9] Rousseeuw PJ (1985) Multivariate Estimation with High Breakdown Point. In: Grossmann W, Pflug G, Vincze I, Wertz W (eds.) Mathematical Statistics and Applications (Vol. B). Reidel, Dordrecht, pp. 283–297.

[10] Rousseeuw PJ (1998) Maxbias Curves. In: Kotz S, Read C, Banks D (eds.) Encyclopedia of Statistical Sciences Update Volume 3. Wiley, New York.

[11] Rousseeuw PJ, Leroy AM (1987) Robust Regression and Outlier Detection. Wiley, New York.

[12] Yohai VJ, Maronna RA (1990) The Maximum Bias of Robust Covariances. Communications in Statistics Part A- Theory and Methods 19: 3925–3933.