NANZAN-TR-2001-01

A characterization of the neighborhoods defined by certain special capacities and their applications to bias-robustness of estimates

Masakazu Ando and Miyoshi Kimura

January 2002

Technical Report of the Nanzan Academic Society
Mathematical Sciences and Information Engineering

A characterization of the neighborhoods defined by certain special capacities and their applications to bias-robustness of estimates

Masakazu Ando and Miyoshi Kimura

Nanzan University

Abstract

A new type of neighborhood (called (c, γ)-neighborhood) is defined by a certain special capacity. As special cases, the neighborhood includes ε-contamination, total variation and Rieder's neighborhoods. A characterization theorem of the neighborhood and a fundamental theorem of the stochastically smallest distribution of the absolute difference of two i.i.d random variables are proved. It is shown that the median has minimax-bias among all location equivariant estimates with respect to (c, γ)-neighborhoods. The implosion biases of five scale estimates including MAD, S and Q over (c, γ)-neighborhoods are derived to be compared. A lower bound on the maximum asymptotic bias of an estimate of θ over (c, γ)-neighborhoods in a general parametric family $\{F_\theta\}$ is derived. The lower bound, which is an extension of He and Simpson's lower bound, depends on a parametric family $\{(F_0 - W)_\theta\}$ of improper distributions with some measure $W \leq F_0$. In the location parametric case, the accuracy of the lower bound is investigated by using the median and the best W is proposed. Some tables and figures of the implosion bias and the lower bound are also given in the case that the model distribution is normal.

AMS 1991 Subject classifications: *Primary 62F35*
Secondary 62G35, 62F10, 28A12, 62E10

Keywords: *Characterization of neighborhoods, Special capacity, Minimax bias, Implosion bias, Median, Scale estimate, Location estimate, Rieder neighborhood, Discrepancy, Lower bounds for the maximum bias, Variation gauge, Bias-robustness*

1. Introduction

In robust statistical inference, the degree of departure from an assumed model distribution of a sample is usually expressed by some suitably chosen neighborhood of the model distribution. In order to describe the departure, various types of neighborhoods have been used to date. Among them, the neighborhoods in terms of ε-contamination and total variation have been most frequently employed in the literatures. As a generalization of such ε-contamination and total variation neighborhoods, Rieder (1977) introduced a neighborhood defined by a special capacity to use it in his works (1978,1981a, 1981b etc) of robust estimation and testing. This special capacity has all properties of 2-alternating Choquet-capacity except the continuity property of Huber and Strasen (1973). A comprehensive study of special capacities is given by Bednarski (1981).

In robust estimation theory there have been proposed various measures of robustness of an estimate such as influence function, gross error sensitivity, breakdown point and maximum asymptotic bias and so on. In particular, the maximum asymptotic bias is the most informative global robustness measure of an estimate, which shows the whole performance of the estimate between the model distribution and the breakdown point. Huber (1964,1981) established that in robust estimation of location the median minimizes the maximum bias among all location equivariant estimates (i.e., the median has minimax bias.) with respect to ε-contamination and Lévy neighborhoods. Chen (1998) showed that the minimax bias property of the median also holds for the neighborhoods in terms of Kolmogorov distance, Kuiper distance and total variation distance. In robust estimation of scale the median absolute deviation (MAD) has been commonly used. However, MAD strongly depends on the symmetry of distributions and it has low Caussian efficiency. As alternatives to MAD, Rousseeuw and Croux (1993) proposed two new scale estimates S and Q whose efficiencies are higher than that of MAD, obtaining their implosion and explosion biases over ε-contamination neighborhoods. On the other hands, He and Simpson (1993) gave a lower bound on the maximum asymptotic bias of an estimate of a general parameter over ε- contamination neighborhoods and considered the accuracy of the lower bound in the case of location.

The purpose of this paper is (1) to introduce a certain type of neighborhood (called (c,γ)-neighborhood) which generalizes Rieder's neighborhood, (2) to prove a characterization theorem of (c,γ)-neighborhoods and a fundamental theorem of the stochastically minimum distribution of the absolute difference of two independent and identically distributed random variables over (c,γ)-neighborhoods, (3) to derive the maximum asymptotic bias of the median and the implosion bias of five scale estimates including MAD, S and Q over (c,γ)-neighborhoods, (4) to obtain a lower bound on the maximum asymptotic bias of an estimate of the general parameter as well as the location parameter over (c,γ)-neighborhoods, and (5) to give some tables and figures of the implosion bias and the lower bound in the case that the model distribution is normal.

In Section 2 we define a (c,γ)-neighborhood $\mathcal{P}_{F_0}(c,\gamma)$ of the model distribution F_0 by a certain special capacity, which is a superposition $g(F_0)$ of F_0 and a concave function $g(x) = \min\{cx+\gamma, 1\}$, where c and γ are some real numbers such that $c \geq 1-\gamma$ and $0 \leq \gamma < 1$. The neighborhood $\mathcal{P}_{F_0}(c,\gamma)$ reduces to Rieder's neighborhood in the case of $1-\gamma \leq c \leq 1$. We prove a characterization theorem of $\mathcal{P}_{F_0}(c,\gamma)$. This characterization theorem is very interesting in its own right as well as for its broad application, and it makes the structure of $\mathcal{P}_{F_0}(c,\gamma)$ clear. A characterization of Rieder's neighborhoods immediately follows from this theorem as a special case. We also verify a useful fundamental theorem that gives us the stochastically smallest one among all distributions of the absolute difference of two independant random variables with common $F \in \mathcal{P}_{F_0}(c,\ \gamma)$.

Let $\{F_\theta\}$ be a parametric family of distributions where the parameter θ is to be estimated. For an estimate T of θ, the maximum asymptotic bias of T over $\mathcal{P}_{F_\theta}(c,\ \gamma)$ is defined by

$$B_T(c,\gamma;F_\theta) \quad = \quad \sup\{|T(G)-\theta| \ : \ G \in \mathcal{P}_{F_\theta}(c,\gamma)\},$$

where T is assumed to be Fisher consistent, $T(F_\theta) = \theta$. In Section 3 we consider the case that θ is the location parameter. We derive the maximum asymptotic bias of the median and show that the median has minimax bias among all location equivariant estimates. This is an important result which should be added to the well known minimax-bias results of the median due to the Huber (1964,1981) and Chen (1998). In Section 4 we treat the case that θ is the scale parameter, and derive the implosion bias of five robust scale estimates including MAD, S and Q. In the case of ε-contamination neighborhoods, the implosion bias reduces to that in Rousseeuw and Croux (1993).

In Section 5 we derive a lower bound on the maximum asymptotic bias $b_T(c,\gamma;F_\theta)$ in the general parametric family. To this end, using some suitable real valued function $\varphi(c,\gamma)$ we define a discrepancy $d_\varphi(G,F)$ of G from F based on (c,γ)- neighborhoods. This discrepancy d_φ is a generalization of the Huber discrepancy based on ε-contamination neighborhoods. We define the neighborhood $P_{F_\theta}(a)$ of F_θ with discrepancy a, and we derive a lower bound on the maximum asymptotic bias of T over $P^{\varphi}_{F_\theta}(a)$ by making use of a parametric family $\{(F_0-W)_\theta\}$ of improper distributions, where W is some measure with mass $(c+\gamma-1)/c$ such as $W \le F_0$. The obtained lower bound is an extension of He and Simpson's (1993) lower bound. The neighborhood $\mathcal{P}^{\varphi}_{F_\theta}(a)$ reduces to some $\mathcal{P}_{F_\theta}(c,\gamma)$ under a special φ, and hence we can obtain a lower bound on $b_T(c,\gamma;F_\theta)$ as a special case of this lower bound. Since the lower bound depends on W, we need to choose a suitable W such that the lower bound is as tight as possible.

In Section 6 we are concerned with a lower bound on $b_T(c,\gamma;F_0)$ in the location parametric family. We propose $W\,(=W_1)$ which yields the best lower bound among all W such that (F_0-W) have even and unimodal densities. In Section 7, we consider the case of $F_0\equiv\Phi$, the standard normal distribution. We give some tables and figures of the implosion bias of MAD, S and Q. We also present a table of the lower bound with respect to W_1 for a location estimate T and by using the median we investigate how the accuracy of the lower bound is. In last Section 8 we collect the proofs of lemmas and theorems.

2. The neighborhoods and their characterization

Let $\mathcal{X}$ be a polish space (i.e., a complete, separable and metrizable space), $\mathcal{B}$ the Borel σ-algebra of subsets of $\mathcal{X}$ and $\mathcal{M}$ the set of all probability measures on $\mathcal{B}$. For some specified $F_0\in\mathcal{M}$ we consider the following type of neighborhood of F_0:

(2.1) $\mathcal{P}_{F_0}(c,\gamma)=\{F\in\mathcal{M}\ :\ c\,F_0(B)-(c+\gamma-1)\le F(B)\le c\,F_0(B)+\gamma,\ {}^{\forall}B\in\mathcal{B}\}$,

where $0\le\gamma<1$ and $1-\gamma\le c<\infty$. This neighborhood is a generalization of that introduced by Rieder (1977). In fact, when $c\le 1$, letting $c=1-\varepsilon$ and $\gamma=\varepsilon+\delta$, we have Rieder neighborhood $\mathcal{P}_{F_0}(1-\varepsilon,\varepsilon+\delta)$, where $\varepsilon\ge0$, $\delta\ge0$ and $\varepsilon+\delta<1$. We also have ε-contamination neighborhood $\mathcal{P}_{F_0}(1-\varepsilon,\varepsilon)$ for $c=1-\varepsilon$ and $\gamma=\varepsilon$, and total variation neighborhood $\mathcal{P}_{F_0}(1,\delta)$ for $c=1$ and $\gamma=\delta$. We note that $\mathcal{P}_{F_0}(c,\gamma)$ is increasing in c and γ.

The neighborhood $\mathcal{P}_{F_0}(c,\gamma)$ can be also defined by a special capacity as follows: Let

$$g(x)=\min(c\,x+\gamma,\ 1),\quad 0\le x\le 1,$$

and let

$$v(B)\ =\ \begin{cases} g(F_0(B)), & for\ \ B\ne\phi,\ B\in\mathcal{B},\\ 0, & for\ \ B=\phi.\end{cases}$$

Then, by Lemma 3.1 of Bednarski(1981), v is a special capacity, which satisfies all the conditions of Choquet's 2-alternating capacity except the condition (4) in Huber and Strassen (1973).

As easily seen, we have

$$\mathcal{P}_{F_0}(c,\ \gamma)\ =\ \{F\in\mathcal{M}\mid F(B)\le v(B),\ \ {}^{\forall}B\in\mathcal{B}\}. \tag{2.2}$$

The following theorem which gives a characterization of $\mathcal{P}_{F_0}(c,\ \gamma)$ is essential and important.

Theorem 2.1 *For $0\le\gamma<1$ and $1-\gamma\le c<\infty$ it holds that*

$$\mathcal{P}_{F_0}(c,\ \gamma)\ =\ \{F=c\,(F_0-W)+\gamma\,K\ :\ W\in\mathcal{W}_{F_0,\lambda}\ ,K\in\mathcal{M}\}, \tag{2.3}$$

where $\mathcal{W}_{F_0,\lambda}$ is the set of all measures W on $\mathcal{B}$ such that $W(B)\le F_0(B)$ for ${}^{\forall}B\in\mathcal{B}$ and $W(\mathcal{X})=\lambda=(c+\gamma-1)/c$.

Proof. First we show that for any $F \in \mathcal{P}_{F_0}(c,\ \gamma)$ and for $\gamma \neq 0$ F is expressed in the form of (2.3). Let f_0 and f be the densities of F_0 and F with respect to a σ-finite measure μ (e.g., $\mu = F_0 + F$), respectively, and let

$$A = \{x \in \mathcal{X} \mid f(x) \leq c\, f_0(x)\}.$$

Then, by (2.1) we have

$$\begin{aligned} 0 \leq \int_A \left(f_0(x) - \frac{f(x)}{c} \right) d\mu &= F_0(A) - \frac{1}{c} F(A) \\ &\leq F_0(A) - \frac{1}{c}(c\, F_0(A) - (c + \gamma - 1)) \\ &= \frac{c + \gamma - 1}{c}. \end{aligned}$$

Hence, as easily seen, we can take two functions ψ_1 and ψ_2 defined on A and A^c, respectively, such that $f_0 - \frac{f}{c} \leq \psi_1 \leq f_0$, $0 \leq \psi_2 \leq f_0$ and

$$\begin{array}{lll} \displaystyle\int_A \psi_1(x) d\mu = \frac{c + \gamma - 1}{c}, \quad \psi_2 \equiv 0, & & if \ \ F_0(A) \geq \dfrac{c + \gamma - 1}{c}, \\ \psi_1 \equiv f_0, \quad \displaystyle\int_{A^c} \psi_2(x) d\mu = \frac{c + \gamma - 1}{c} - F_0(A), & & if \ \ F_0(A) < \dfrac{c + \gamma - 1}{c}. \end{array} \tag{2.4}$$

By using ψ_1 and ψ_2, we define a function ψ on $\mathcal{X}$ by

$$\psi(x) = \begin{cases} \psi_1(x), & x \in A, \\ \psi_2(x), & x \in A^c. \end{cases}$$

Then it is clear that $0 \leq \psi \leq f_0$ and

$$\int \psi(x) d\mu = \frac{c + \gamma - 1}{c}.$$

Letting

$$k(x) = \frac{1}{\gamma}\{f(x) - c(f_0(x) - \psi(x))\}, \qquad x \in \mathcal{X}, \tag{2.5}$$

we can see $k \geq 0$ and

$$\int k(x) d\mu = 1.$$

From (2.5) it follows that

$$f(x) = c(f_0(x) - \psi(x)) + \gamma\, k(x), \qquad x \in \mathcal{X}.$$

This implies

$$F = c(F_0 - W) + \gamma\, K,$$

where $W \in \mathcal{W}_{F_0,\lambda}$ and $K \in \mathcal{M}$ are the measures with the densities ψ and k with respect to μ, respectively.

Secondly, we consider the case of $\gamma = 0$. In this case, for any $F \in \mathcal{P}_{F_0}(c,\ 0)$ we have $A = \mathcal{X}$. Hence, letting $\psi = f_0 - \frac{f}{c}$ on $\mathcal{X}$, we obtain

$$\int \psi(x) d\mu = \frac{c - 1}{c},$$

and

$$f(x) = c\,(f_0(x) - \psi(x)), \quad x \in \mathcal{X}.$$

This implies

$$F = c\,(F_0 - W),$$

where $W \in \mathcal{W}_{F_0,\lambda}$ is the measure with the density ψ with respect to μ.

Conversely, let F be any probability measure expressed in the form (2.3). Then it is easy to see

$$c\,F_0(B) - (c + \gamma - 1) \leq F(B) \leq c\,F_0(B) + \gamma, \qquad {}^{\forall}B \in \mathcal{B}.$$

This implies $F \in \mathcal{P}_{F_0}(c,\ \gamma)$. □

As an important special case of Theorem 2.1, we obtain the following characterization of Rieder neighborhood.

Corollary 2.1 *For* $\varepsilon \geq 0,\ \delta \geq 0$ *and* $\varepsilon + \delta < 1$, *it holds that*

$$(2.6) \quad \mathcal{P}_{F_0}(1-\varepsilon,\ \varepsilon+\delta) \ = \ \{F = (1-\varepsilon)(F_0 - W) + (\varepsilon+\delta)K \ : \ W \in \mathcal{W}_{F_0,\lambda}, K \in \mathcal{M}\},$$

where $\mathcal{W}_{F_0,\lambda}$ *is the set of all measures* W *on* $\mathcal{B}$ *such that* $W(B) \leq F_0(B)$ *for* ${}^{\forall}B \in \mathcal{B}$ *and* $W(\mathcal{X}) = \lambda = \delta/(1-\varepsilon)$.

Remark 2.1

(i) The role of W in the characterization (2.3) and (2.6) is essentially important. When $W = \lambda F_0$, we see $c\,(F_0 - W) + \gamma K = (1-\gamma)F_0 + \gamma K$.

(ii) The first inequality in the definition (2.1) of $\mathcal{P}_{F_0}(c, \gamma)$ is not necessary, i.e.,

$$\mathcal{P}_{F_0}(c, \gamma) = \{F \in \mathcal{M} \ : \ F(B) \leq cF_0(B) + \gamma, \ \ {}^{\forall}B \in \mathcal{B}\}.$$

Hereafter, we consider the case of $\mathcal{X} = R$, the real line. Let X and Y be independent and identically distributed random variables with a common F. We are interested in finding $F \in \mathcal{P}_{F_0}(c,\ \gamma)$ such that the distribution of $|X - Y|$ under F is stochastically smallest in $\mathcal{P}_{F_0}(c,\ \gamma)$. To this end, we need a fundamental result.

Let f be a nonnegative real valued measurable function such that

$$0 < \int_{-\infty}^{\infty} f(x)dx = M < \infty,$$

where M is a constant. For some positive constant $m (0 < m < M)$ let a be the positive number satisfying

$$\int_{-a}^{a} f(x)dx = m,$$

and let

$$\hat{g}(x) \ = \ \begin{cases} f(x), & -a \leq x \leq a, \\ 0, & \textit{otherwise}. \end{cases}$$

Furthermore let

$$\mathcal{F}_0 = \{g \in \mathcal{F} \mid 0 \le g \le f, \;\; 0 \le \int_{-\infty}^{\infty} g(x)dx \le m\},$$
$$G(x) = \int_{-\infty}^{x} g(t)dt \qquad \text{and} \qquad \hat{G}(x) = \int_{-\infty}^{x} \hat{g}(t)dt,$$

where $\mathcal{F}$ is the set of all measurable functions defined on R. Note that $\hat{g} \in \mathcal{F}$. We obtain the following result which is used to derive Theorem 2.2 below

Lemma 2.1 *Let f be even and unimodal. Then*

(i) $$\sup_{g \in \mathcal{F}_0} \int_{-\infty}^{\infty} \{G(x+t) - G(x)\} g(x)dx = \int_{-\infty}^{\infty} \{\hat{G}(x+t) - \hat{G}(x)\} \hat{g}(x)dx, \qquad 0 \le {}^{\forall}t < \infty,$$

(ii) $$\sup_{g \in \mathcal{F}_0} \int_{-\infty}^{\infty} G(x+t) g(x)dx = \int_{-\infty}^{\infty} \hat{G}(x+t) \hat{g}(x)dx, \qquad 0 \le {}^{\forall}t < \infty.$$

Let F_0 be a probability measure on $(R, \mathcal{B})$ with a density f_0 which is even and unimodal. Let a be the upper $\frac{100(c+\gamma-1)}{2c}\%$ percent point of F_0 and let $\hat{W}$ be the measure defined by

$$\hat{W}(B) = F_0(B \cap [-a, \; a]^c), \qquad {}^{\forall}B \in \mathcal{B}.$$

Further let

$$\hat{F} = c\,(F_0 - \hat{W}) + \gamma\,\Delta_0. \tag{2.7}$$

where Δ_0 denotes the probability measure which puts mass 1 at the origin 0. We note $\hat{F} \in \mathcal{P}_{F_0}(c, \; \gamma)$. The following fundamental result is obtained.

Theorem 2.2 *Let X and Y be independent and identically distributed random variables with a common $F \in \mathcal{P}_{F_0}(c, \; \gamma)$. Then the distribution of $|X - Y|$ is stochastically smallest under $\hat{F}$, i.e.,*

$$\sup_{F \in \mathcal{P}_{F_0}(c, \; \gamma)} P_{F \times F}(|X - Y| \le t) = P_{\hat{F} \times \hat{F}}(|X - Y| \le t), \quad 0 \le {}^{\forall}t < \infty.$$

Proof. By Theorem 2.1 we have

$$\mathcal{P}_{F_0}(c, \; \gamma) = \{F = c\,(F_0 - W) + \gamma\,K \;\; : \;\; W \in \mathcal{W}_{F_0,\lambda}, K \in \mathcal{M}\}.$$

Hence for ${}^{\forall}F \in \mathcal{P}_{F_0}(c, \; \gamma)$ and for $0 \le {}^{\forall}t < \infty$,

$$\begin{aligned} P_{F \times F}(|X - Y| \le t) &= c^2((F_0 - W) \times (F_0 - W))(|X - Y| \le t) \\ &\quad + 2\,c\,\gamma((F_0 - W) \times K)(|X - Y| \le t) \\ &\quad + \gamma^2(K \times K)(|X - Y| \le t), \end{aligned} \tag{2.8}$$

where for two measures H_1 and H_2 the notation $(H_1 \times H_2)(|X - Y| \leq t)$ denotes the measure of the set $\{(x,\ y) : |x - y| \leq t\}$ under the product measure $H_1 \times H_2$. From Lemma 2.1 and the fact that the distribution of $X - Y$ is symmetric about the origin, it follows that

$$
\begin{aligned}
&((F_0 - W) \times (F_0 - W))(|X - Y| \leq t) \\
&\quad = 2((F_0 - W) \times (F_0 - W))(0 \leq X - Y \leq t) \\
&\quad = 2\int_{-\infty}^{\infty} (F_0 - W)(y \leq X \leq y + t)(F_0 - W)(dy) \\
&\quad = 2\int_{-\infty}^{\infty} \{(F_0 - W)(y + t) - (F_0 - W)(y)\}(F_0 - W)(dy) \\
&\quad \leq 2\int_{-\infty}^{\infty} \{(F_0 - \hat{W})(y + t) - (F_0 - \hat{W})(y)\}(F_0 - \hat{W})(dy) \\
&\quad = ((F_0 - \hat{W}) \times (F_0 - \hat{W}))(|X - Y| \leq t),
\end{aligned}
\tag{2.9}
$$

where the notation $H(r \leq X \leq s)$ denotes the measure of the interval $[r,\ s]$ under H. Also, it follows that

$$
\begin{aligned}
((F_0 - W) \times K)(|X - Y| \leq t) &= \int_{-\infty}^{\infty} (F_0 - W)(y - t \leq X \leq y + t)K(dy) \\
&\leq (F_0 - \hat{W})(-t \leq X \leq t) \\
&= ((F_0 - \hat{W}) \times \Delta_0)(|X - Y| \leq t),
\end{aligned}
\tag{2.10}
$$

and that

$$
(K \times K)(|X - Y| \leq t) \leq 1 = (\Delta_0 \times \Delta_0)(|X - Y| \leq t). \tag{2.11}
$$

Substituting (2.9), (2.10) and (2.11) into (2.8), we obtain

$$
P_{F \times F}(|X - Y| \leq t) \leq P_{\hat{F} \times \hat{F}}(|X - Y| \leq t).
$$

This completes the proof of the theorem. □

3. The minimax bias property of the median

Let F_0 be a symmetric distribution about the origin and let $F_\theta(x) = F_0(x - \theta)$, where the location parameter θ is to be estimated. Let $X_1, \ldots, X_n$ be independent random variables distributed with a common F. We assume that F belongs to the neighborhood

$$
\mathcal{P}_{F_\theta}(c,\ \gamma) = \{F : F(x) = c\,(F_0 - W)(x - \theta) + \gamma\, K(x), x \in R, W \in \mathcal{W}_{F_0,\lambda}, K \in \mathcal{M}\}, \tag{3.1}
$$

where $\mathcal{W}_{F_0,\lambda}$ and $\mathcal{M}$ are given in (2.3). Let T be an estimating functional (estimate) defined on $\mathcal{M}$. We assume that T is Fisher consistent. Since we consider only location equivariant estimates, we can assume $\theta = 0$ without loss of generality. In this case, the maximum asymptotic bias of T over $\mathcal{P}_{F_0}(c, \gamma)$ is defined by

$$
B_T(c,\ \gamma) = \sup\{|T(F)| : F \in \mathcal{P}_{F_0}(c,\ \gamma)\}. \tag{3.2}
$$

Let T_{M_n} be the sample median of $X_1, \ldots, X_n$, i.e.,

$$
T_{M_n} = \underset{j}{med}\, X_j,
$$

which is the middle order statistic when n is odd, and the average of the order statistics with ranks $\frac{n}{2}$ and $\frac{n}{2}+1$ when n is even. The asymptotic version of T_{M_n} is the median of F, i.e.,

$$T_M(F) \quad = \quad F^{-1}\left(\frac{1}{2}\right).$$

where $F^{-1}(u) = \inf\{x \mid F(x) \geq u\}, 0 \leq u \leq 1$. Huber (1964,1981) shows that when F_0 has an even and unimodal density, T_M is a minimax-bias functional among all location equivariant functionals with respect to ε-contamination and Lévy neighborhoods. By using the results of Donoho and Liu (1988) and He and Simpson (1993), Chen (1998) obtained the same result as Huber's mentioned above with respect to the Kolmogorov, the Kuiper and total variation neighborhoods.

The following theorem states that the Huber's result of the minimax-bias property of T_M also holds with respect to our neighborhood $\mathcal{P}_{F_0}(c,\ \gamma)$.

Theorem 3.1 *Let F_0 have an even and unimodal density f_0. Then, for $0 \leq \gamma < \frac{1}{2}$ the median T_M has minimax-bias in the class $\mathcal{T}$ of all location equivariant estimates, i.e.,*

$$\inf\{B_T(c,\ \gamma) : T \in \mathcal{T}\} \quad = \quad B_{T_M}(c,\ \gamma),$$

where

$$B_{T_M}(c, \gamma) \quad = \quad F_0^{-1}\left(\frac{2c+2\gamma-1}{2c}\right).$$

Proof. First we note that the maximum absolute bias of the median over $\mathcal{P}_{F_0}(c,\ \gamma)$ is attained when $F = c\,(F_0 - \hat{W}_L) + \gamma\,\Delta_{x_M}$, where

$$\hat{W}_L(x) \quad = \quad \min\left\{F_0(x), \frac{c+\gamma-1}{c}\right\}, \qquad -\infty < x < \infty,$$

and Δ_{x_M} denotes the probability measure with mass 1 at x_M (sufficiently large). Hence, letting x_0 be the solution of

$$c\,(F_0 - \hat{W}_L)(x_0) \quad = \quad \frac{1}{2}, \qquad i.e., \qquad x_0 = F_0^{-1}\left(\frac{2c+2\gamma-1}{2c}\right),$$

we have

$$\sup\{|T_M(F)| \ : \ F \in \mathcal{P}_{F_0}(c,\ \gamma)\} \quad = \quad x_0. \tag{3.3}$$

Let D_1 and D_2 be the regions enclosed by $y = f_0(x)$ and the x-axis, and by $y = f_0(x - 2x_0)$ and the x-axis, respectively, and let $D = D_1 - D_2$. Then the area of D is $\frac{c+2\gamma-1}{c}$. By making use of this fact, we construct two distributions $F_+, F_- \in \mathcal{P}_{F_0}(c,\ \gamma)$ which are symmetric about x_0 and $-x_0$, respectively, and which are translates of each other. We define the densities f_+ and f_- of F_+ and F_-, respectively, as follows (see Figure 3.1):

$$f_+(x) \quad = \quad \begin{cases} c\left[f_0(x) - \left(\frac{c+\gamma-1}{c+2\gamma-1}\right)(f_0(x) - f_0(x-2x_0))\right], & for \ \ x < x_0, \\ c\left[f_0(x-2x_0) - \left(\frac{c+\gamma-1}{c+2\gamma-1}\right)(f_0(x-2x_0) - f_0(x))\right], & for \ \ x \geq x_0, \end{cases}$$

$$f_-(x) \quad = \quad f_+(x+2x_0).$$

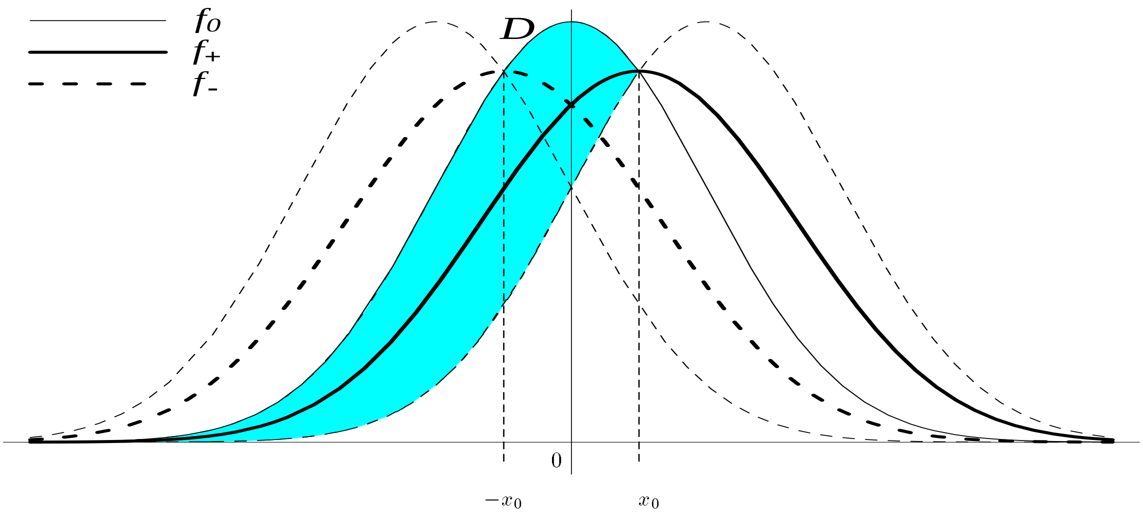

Figure 3.1: The density functions f_+, f_-, f_0

It is easy to check that F_+ and F_- belong to $\mathcal{P}_{F_0}(c,\ \gamma)$. Since $F_-(x) = F_+(x + 2x_0)$, it follows that for any $T \in \mathcal{T}$

$$T(F_+) - T(F_-) = 2x_0.$$

This implies

$$\max(|T(F_+)|,\ |T(F_-)|) \geq x_0,$$

and hence

$$\sup\{|T(F)|\ :\ F \in \mathcal{P}_{F_0}(c,\ \gamma)\} \ \geq \ x_0. \tag{3.4}$$

Therefore the theorem follows from (3.3) and (3.4). □

4. The implosion bias of scale estimates

Let F_0 be a specified distribution function with an even and unimodal density f_0. Let $X_1, \ldots, X_n$ be independent and identically distributed with F. We assume that F belongs to the neighborhood

$$\begin{aligned}\mathcal{P}_{F_{\mu,s}}(c,\gamma) \ = \ \Big\{F\ :\ F(x) = c\,(F_0 - W)\left(\frac{x-\mu}{s}\right) + \gamma K(x),\\ x \in R,\ W \in \mathcal{W}_{F_0,\lambda}\ , K \in \mathcal{M}\Big\}\,,\end{aligned} \tag{4.1}$$

where μ is the unknown location parameter and $s > 0$ is the unknown scale parameter to be estimated. Among robust estimates of scale proposed to date, we especially consider the following five estimates:

$$\begin{aligned}\mathrm{MAD}_n \ &= \ a_1 \operatorname*{med}_i\{|X_i - \operatorname*{med}_j X_j|\},\\ \mathrm{S}_n \ &= \ a_2 \operatorname*{med}_i\{\operatorname*{med}_j |X_i - X_j|\},\end{aligned}$$

$$
\begin{aligned}
\mathrm{Q}_n &= a_3 \{|X_i - X_j| : i < j\}_{(k)}, \\
\mathrm{LMS}_n &= a_4 \min_i |X_{(i+h-1)} - X_{(i)}|, \\
\mathrm{L}_n &= a_5 \min_i \{\operatorname{med}_j |X_i - X_j|\},
\end{aligned}
$$

where $a_i,\ i = 1, \cdots, 5$ are some constants, $k = \binom{h}{2}$ and $h = [\frac{n}{2}] + 1$. The MAD_n (the median absolute deviation about the median) with $a_1 = 1.4826$ is well known and used commonly. The S_n and Q_n, which were proposed as alternatives to the MAD_n and investigated by Rousseeuw and Croux (1993), have 50 % breakdown points and higher efficiency than MAD_n. The LMS_n, which was first used in Rousseeuw (1984), has a 50% breakdown point and the same influence function as that of the MAD (Rousseeow and Leroy, 1987). Its efficiency equals that of the MAD (Grübel,1988). The L_n which is obtained from the p-subst algorithm of Rousseeuw and Leroy (1987), is asymptotically equivalent to LMS_n, and also has a 50% breakdown point.

The above five estimates are location invariant and scale equivariant. We derive the implosion bias of these estimates over $\mathcal{P}_{F_0}(c, \gamma)$ (the case of $F_{\mu,s} = F_0$ with $\mu = 0$ and $s = 1$). The implosion bias of a scale estimate T over $\mathcal{P}_{F_0}(c, \gamma)$ is defined by

$$
B_T^-(c,\ \gamma) \ = \ \inf\{T(F) \ : \ F \in \mathcal{P}_{F_0}(c,\ \gamma)\}. \tag{4.2}
$$

In what follows, we let X and Y be independent random variables.

The asymptotic version of MAD_n is given by

$$
\mathrm{MAD}(F) = a_1 \operatorname{med}_F \{|X - \operatorname{med}_F Y|\}.
$$

Theorem 4.1 *Let F_0 have an even and unimodal density f_0. Then*

$$
B_{\mathrm{MAD}}^-(c,\ \gamma) \ = \ \begin{cases} a_1\, F_0^{-1}\left(\frac{2c-2\gamma+1}{4c}\right), & if \ \ 0 \le \gamma < \frac{1}{2}, \\ 0, & if \ \ \gamma \ge \frac{1}{2}. \end{cases} \tag{4.3}
$$

The asymptotic version of S_n is given by

$$
\mathrm{S}(F) = a_2 \operatorname{med}_F g_F(X),
$$

where

$$
g_F(x) = \operatorname{med}_F |x - Y|.
$$

We note that if F_n is the empirical distribution, then $\mathrm{S}(F_n) = \mathrm{S}_n$.

Theorem 4.2 *Let F_0 have an even and unimodal density f_0. Then*

$$
B_{\mathrm{S}}^-(c,\ \gamma) \ = \ \begin{cases} a_2\, g^-\left(F_0^{-1}\left(\frac{2c-2\gamma+1}{4c}\right)\right), & if \ \ 0 \le \gamma < \frac{1}{2}, \\ 0, & if \ \ \gamma \ge \frac{1}{2}, \end{cases} \tag{4.4}
$$

where g^- is defined implicitly by

$$
F_0(x + g^-(x)) - F_0(x - g^-(x)) \ = \ \frac{1 - 2\gamma}{2c}. \tag{4.5}
$$

The asymptotic version of Q_n is given by

$$\mathrm{Q}(F) = a_3\, H_F^{-1}\left(\frac{1}{4}\right) = a_3\, K_F^{-1}\left(\frac{5}{8}\right),$$

where H_F and K_F denote the distributions of $|X-Y|$ and $X-Y$ under F, respectively. We note that K_F is symmetric about the origin.

Theorem 4.3 *Let F_0 have an even and unimodal density f_0. Then*

$$B_{\mathrm{Q}}^{-}(c,\ \gamma) \;=\; \begin{cases} \mathrm{Q}(\hat{F}), & if\ \ 0 \le \gamma < \frac{1}{2}, \\ 0, & if\ \ \gamma \ge \frac{1}{2}, \end{cases} \tag{4.6}$$

and $\mathrm{Q}(\hat{F})$ satisfies the equation

$$c^2(F_0-\hat{W})^{*2}(a_3^{-1}\,\mathrm{Q}(\hat{F})) + 2c\gamma(F_0-\hat{W})(a_3^{-1}\,\mathrm{Q}(\hat{F})) + \gamma^2 = \frac{5}{8}, \tag{4.7}$$

where $\hat{F}$ is given by (2.7) and $(F_0-\hat{W})^{*2}$ denotes the convolution $(F_0-\hat{W})*(F_0-\hat{W})$.

The asymptotic version of LMS_n is given by

$$\mathrm{LMS}(F) = a_4 \inf_{t\in[0,\ \frac{1}{2}]}\left|F^{-1}\left(t+\frac{1}{2}\right) - F^{-1}(t)\right|.$$

Theorem 4.4 *Let F_0 have an even and unimodal density f_0. Then*

$$B_{\mathrm{LMS}}^{-}(c,\ \gamma) \;=\; \begin{cases} 2a_4F_0^{-1}\left(\dfrac{2c-2\gamma+1}{4c}\right), & if\ \ 0 \le \gamma < \dfrac{1}{2}, \\ 0, & if\ \ \gamma \ge \dfrac{1}{2}. \end{cases} \tag{4.8}$$

The asymptotic version of L_n is

$$\mathrm{L}(F) = a_5 \inf_x g_F(x),$$

where

$$g_F(x) = \operatorname*{med}_F |x-Y|.$$

Theorem 4.5 *Let F_0 have an even and unimodal density f_0. Then*

$$B_{\mathrm{L}}^{-}(c,\ \gamma) \;=\; \begin{cases} a_5F_0^{-1}\left(\dfrac{2c-2\gamma+1}{4c}\right), & if\ \ 0 \le \gamma < \frac{1}{2}, \\ 0, & if\ \ \gamma \ge \frac{1}{2}. \end{cases} \tag{4.9}$$

Remark 4.1 When $c = 1-\varepsilon$ and $\gamma = \varepsilon$, $B_S^-(c,\gamma)$ and $B_Q^-(c,\gamma)$ in Theorem 4.2 and 4.3 reduce to those in Theorems 4 and 7 of Rousseeuw and Croux (1993).

5. The derivation of lower bounds on the maximum bias

Let $\mathcal{X}$ be a polish space and let $\{F_\theta\}$ be a parametric family indexed by a real-valued parameter $\theta \in \Theta$ where θ is to be estimated. Let

$$(5.1) \qquad \Omega \;=\; \left\{(c,\ \gamma)\ :\ 1-\gamma \le c < \infty,\ 0 \le \gamma < \frac{1}{2}\right\},$$

and let $\varphi(c,\ \gamma)$ be a nonnegative continuous real valued function defined on Ω with $\varphi(1,\ 0) = 0$. We assume that φ is nondecreasing in c and γ. For any two probability measures $F, G \in \mathcal{M}$ we define a discrepancy $d_\varphi(G, F)$ as follows:

$$(5.2) \qquad d_\varphi(G, F) \;=\; \inf\{\varphi(c,\ \gamma)\ :\ (c,\ \gamma) \in \Omega_{G,F}\},$$

where

$$(5.3) \qquad \Omega_{G,F} \;=\; \{(c,\ \gamma) \in \Omega\ :\ G(B) \le c\,F(B) + \gamma,\ \ {}^{\forall}B \in \mathcal{B}\}.$$

When $c = 1-\gamma$, the discrepancy d_φ reduces to the Huber discrepancy based on ε-contamination neighborhoods. We can see that $\Omega_{G,F}$ is convex and closed, and hence there exists a point $(c_0,\ \gamma_0) \in \Omega_{G,F}$ such that $d_\varphi(G, F) = \varphi(c_0,\ \gamma_0)$. By using d_φ we define a neighborhood of F with discrepancy a as

$$(5.4) \qquad \mathcal{P}_F^\varphi(a) \;=\; \{G \in \mathcal{M} \mid d_\varphi(G, F) \le a\}.$$

Note that $\mathcal{P}_F^\varphi(a)$ is nondecreasing in a. As easily seen, we have

$$(5.5) \qquad \mathcal{P}_F^\varphi(a) \;=\; \bigcup_{\varphi(c,\ \gamma)\le a} \mathcal{P}_F(c,\ \gamma),$$

where $\mathcal{P}_F(c,\ \gamma)$ is given by (2.3). For any Fisher consistent estimate T, the maximum asymptotic bias of T over $\mathcal{P}_{F_\theta}^\varphi(a)$ is defined as

$$(5.6) \qquad b_T^\varphi(a,\ F_\theta) \;=\; \sup\{\rho(T(G),\ \theta)\ :\ G \in \mathcal{P}_{F_\theta}^\varphi(a)\},$$

where ρ is a distance defined on Θ. By Theorem 2.1, for any $(c,\ \gamma)$ and any $W \in \mathcal{W}_{F_0,\lambda}$ we consider a parametric family $\{\tilde{F}_{\theta,W}\}$ of improper distributions ($\tilde{F}_{\theta,W}(\mathcal{X}) = 1-\gamma$), where $\tilde{F}_{\theta,W} = (F_0 - W)_\theta$. The variation distance $\tilde{d}_v(\tilde{F}_{\theta,W}, \tilde{F}_{\eta,W})$ between $\tilde{F}_{\theta,W}$ and $\tilde{F}_{\eta,W}$ is defined as

$$(5.7) \qquad \tilde{d}_v(\tilde{F}_{\theta,W}, \tilde{F}_{\eta,W}) \;=\; \sup\{|\tilde{F}_{\theta,W}(B) - \tilde{F}_{\eta,W}(B)|\ :\ B \in \mathcal{B}\}.$$

Let $\tilde{f}_{\theta,W}$ and $\tilde{f}_{\eta,W}$ be the densities of $\tilde{F}_{\theta,W}$ and $\tilde{F}_{\eta,W}$ with respect to a σ-finite measure μ. Then it is clear that

$$(5.8) \qquad \tilde{d}_v(\tilde{F}_{\theta,W}, \tilde{F}_{\eta,W}) \;=\; \frac{1}{2}\int |\tilde{f}_{\theta,W} - \tilde{f}_{\eta,W}|d\mu = \int (\tilde{f}_{\theta,W} - \tilde{f}_{\eta,W})_+ d\mu = \int (\tilde{f}_{\theta,W} - \tilde{f}_{\eta,W})_- d\mu,$$

where $f_+ = \max(0,\ f)$ and $f_- = \max(0,\ -f)$. Note that $0 \le \tilde{d}_v(\tilde{F}_{\theta,W}, \tilde{F}_{\eta,W}) \le \dfrac{1-\gamma}{c}$. As in Donoho and Liu (1988), we define a variation gauge $\tilde{b}_{v,W}$ (depending on W) by

$$(5.9) \qquad \tilde{b}_{v,W}(a,\ F_\theta) \;=\; \sup\{\rho(\theta,\ \eta)\ :\ \eta \text{ such that } \tilde{d}_v(\tilde{F}_{\theta,W}, \tilde{F}_{\eta,W}) \le a\}.$$

We establish the following result which generalizes Theorem 2.1 of He and Simpson (1993).

Theorem 5.1 *Suppose that $\{F_\theta\}$ is dominated by a σ-finite measure μ and let $(c_0,\ \gamma_0)$ be a given point in Ω. If T is an estimating functional of θ, then for each $W \in \mathcal{W}_{F_0,\lambda}$ it holds that*

$$\sup_{\eta:\rho(\theta,\eta)\le \tilde{b}_{v,W}((1-\lambda)\frac{a}{1+a},F_\theta)} b_T^{\varphi}(J_\lambda(a),\ F_\eta) \ \ge\ \frac{1}{2}\tilde{b}_{v,W}\left((1-\lambda)\frac{a}{1+a},\ F_\theta\right),\quad a\ge 0, \tag{5.10}$$

where

$$\begin{aligned} J_\lambda(a) &= \varphi(c^*(a),\ \gamma^*(a)),\quad \lambda=\frac{c_0+\gamma_0-1}{c_0},\\ c^*(a) &= \frac{1+a}{(1-\lambda)(1+2a)},\quad \gamma^*(a)=\frac{a}{1+2a}. \end{aligned}$$

Proof. We fix $\theta\in\Theta$. For each $\eta\in\Theta$ we set

$$\xi = \frac{\tilde{d}_v(\tilde{F}_{\theta,W},\tilde{F}_{\eta,W})}{(1-\lambda)-\tilde{d}_v(\tilde{F}_{\theta,W},\tilde{F}_{\eta,W})}, \tag{5.11}$$

where $\tilde{d}_v$ is given in (5.7). This implies

$$\tilde{d}_v(\tilde{F}_{\theta,W},\tilde{F}_{\eta,W}) = \frac{(1-\lambda)\xi}{1+\xi}. \tag{5.12}$$

Note that $0\le \tilde{d}_v \le 1-\lambda$ and $0\le\xi<\infty$. Let

$$g=\left(\frac{1+\xi}{(1-\lambda)\xi}\right)\left(\tilde{f}_{\eta,W}-\tilde{f}_{\theta,W}\right)_+ \quad\text{and}\quad h=\left(\frac{1+\xi}{(1-\lambda)\xi}\right)\left(\tilde{f}_{\eta,W}-\tilde{f}_{\theta,W}\right)_-$$

Then, by (5.8) and (5.12) we have

$$\int g\,d\mu=\int h\,d\mu=1.$$

Thus g and h are probability density functions. Since

$$(\tilde{f}_{\eta,W}-\tilde{f}_{\theta,W})_+ = (\tilde{f}_{\eta,W}-\tilde{f}_{\theta,W})+(\tilde{f}_{\eta,W}-\tilde{f}_{\theta,W})_-$$

it follows that

$$(1+\xi)\tilde{f}_{\theta,W}+(1-\lambda)\xi\, g=(1+\xi)\tilde{f}_{\eta,W}+(1-\lambda)\xi\, h. \tag{5.13}$$

Hence, letting

$$c^*(\xi)=\frac{1+\xi}{(1-\lambda)(1+2\xi)}\quad\text{and}\quad \gamma^*(\xi)=\frac{\xi}{1+2\xi}, \tag{5.14}$$

we have

$$c^*(\xi)\tilde{f}_{\theta,W}+\gamma^*(\xi)g = c^*(\xi)\tilde{f}_{\eta,W}+\gamma^*(\xi)h. \tag{5.15}$$

Note that $(c^*,\gamma^*)\in\Omega$ and $\dfrac{c^*+\gamma^*-1}{c^*}=\dfrac{c_0+\gamma_0-1}{c_0}=\lambda$. We can also see that $c^*(\xi)$ and $\gamma^*(\xi)$ are decreasing and increasing in ξ, respectively. Let

$$F^* = c^*\tilde{F}_{\theta,W}+\gamma^* G, \tag{5.16}$$

where G is the probability measure with the density g. Then it follows from (5.2), (5.15) and Theorem 2.1 that

$$d_\varphi(F^*, F_\theta) \le \varphi(c^*,\ \gamma^*) \quad \text{and} \quad d_\varphi(F^*, F_\eta) \le \varphi(c^*,\ \gamma^*).$$

Hence

$$\begin{aligned} \rho(\theta,\ \eta) &\le \rho(\theta,\ T(F^*)) + \rho(\eta,\ T(F^*)) \\ &\le \sup_{d_\varphi(F,F_\theta)\le\varphi(c^*,\gamma^*)} \rho(\theta,\ T(F)) + \sup_{d_\varphi(F,F_\eta)\le\varphi(c^*,\gamma^*)} \rho(\eta,\ T(F)) \\ &= b_T^\varphi(\varphi(c^*,\gamma^*),\ F_\theta) + b_T^\varphi(\varphi(c^*,\gamma^*),\ F_\eta) \\ &= b_T^\varphi(J_\lambda(\xi),\ F_\theta) + b_T^\varphi(J_\lambda(\xi),\ F_\eta), \end{aligned}$$

where

(5.17)
$$J_\lambda(\xi) \;=\; \varphi(c^*(\xi),\ \gamma^*(\xi)).$$

We assume that $J_\lambda(\xi)$ is increasing in ξ. Since

$$\tilde{d}_v(\tilde{F}_{\theta,W}, \tilde{F}_{\eta,W}) \le \frac{(1-\lambda)a}{1+a} \quad \text{if and only if} \quad \xi \le a,$$

it follows that

$$\begin{aligned} \tilde{b}_{v,W}\left(\frac{(1-\lambda)a}{1+a},\ F_\theta\right) &= \sup_{\eta:\tilde{d}_v(\tilde{F}_{\theta,W},\tilde{F}_{\eta,W})\le\frac{(1-\lambda)a}{1+a}} \rho(\theta,\ \eta) \\ &\le \sup_{\eta:\xi(\eta)\le a} \rho(\theta,\ \eta) \\ &\le \sup_{\eta:\xi(\eta)\le a} \{b_T^\varphi(J_\lambda(\xi),\ F_\theta) + b_T^\varphi(J_\lambda(\xi),\ F_\eta)\} \\ &\le 2 \sup_{\eta:\xi(\eta)\le a} b_T^\varphi(J_\lambda(a),\ F_\eta). \end{aligned}$$

The last inequality follows from the facts that $J_\lambda(\xi)$ is increasing in ξ and that $\xi(\theta) = 0$. This completes the proof of the theorem. □

Let us consider the case of $c + \gamma = 1$, that is, the ε-contamination case. In this case we see $\lambda = 0$, $W \equiv 0$, $\tilde{d}_v = d_v$ and $\tilde{b}_v = b_v$, where

(5.18)
$$\begin{aligned} d_v(F_\theta, F_\eta) &= \sup\{|F_\theta(B) - F_\eta(B)| \ :\ B \in \mathcal{B}\}, \\ b_v(\varepsilon; F_\theta) &= \sup\{\rho(\theta,\eta) \ :\ \eta \text{ such that } d_v(F_\theta, F_\eta) \le \varepsilon\}. \end{aligned}$$

By taking $\varphi(c,\gamma) = \gamma$ and $a = \frac{\varepsilon}{1-2\varepsilon}$ we also have $J_\lambda(a) = \varepsilon$, $\frac{a}{1+a} = \frac{\varepsilon}{1-\varepsilon}$ and $b_T^\varphi = b_T$, where

$$b_T(\varepsilon; F_\theta) \;=\; \sup\{\rho(T(G),\theta) \ :\ G \in \mathcal{P}_{F_\theta}(1-\varepsilon,\varepsilon)\}.$$

Therefore as a special case of Theorem 5.1 we obtain the following result.

Corollary 5.1 (Theorem 2.1 of He and Simpson, 1993) *Suppose $\{F_\theta\}$ is dominated by a σ-finite measure. If T is a functional mapping distributions to parameter values, then its contamination bias satisfies*

$$\sup_{\eta:\rho(\theta,\eta)\le b_v(\varepsilon/(1-\varepsilon);F_\theta)} b_T(\varepsilon; F_\eta) \ge \frac{1}{2} b_v\left(\frac{\varepsilon}{1-\varepsilon}; F_\theta\right).$$

Remark 5.1 Although the definition of ξ in (5.11) is different from that of $\delta = \dfrac{d_v(F_\theta, F_\eta)}{1 + d_v(F_\theta, F_\eta)}$ in (7.1) of He and Simpson (1993), both of the definitions yield the same results.

We are now interested in deriving a lower bound on the maximum bias $B_T(c, \gamma; F)$ of T over $\mathcal{P}_F(c, \gamma)$. To do this we consider two cases of $\frac{1}{2} < c \leq 1$ and $c \geq 1$ separately. First we treat the case $\frac{1}{2} \leq c \leq 1$ and restrict Ω to its subset Ω_1 defined as

$$\Omega_1 = \left\{ (c, \gamma) \ : \ 1 - \gamma \leq c \leq 1, \ 0 \leq \gamma < \frac{1}{2} \right\}.$$

In this case, we note that $\mathcal{P}_F(c, \gamma)$ reduces to the neighborhood introduced by Rieder (1977) (Take $c = 1 - \varepsilon$ and $\gamma = \varepsilon + \delta$). Let

$$\varphi_1(c, \gamma) = \varphi_{k,\lambda}^{(1)}(c, \gamma) = \max(1 - c, \ k(c + \gamma - 1)), \tag{5.19}$$

where k is a given positive real number. Then we have

$$J_1(\xi) = J_{k,\lambda}^{(1)}(\xi) = \varphi_{k,\lambda}^{(1)}(c^*(\xi), \gamma^*(\xi)) = \max(1 - c^*(\xi), \ k(c^*(\xi) + \gamma^*(\xi) - 1)),$$

where $c^*(\xi)$ and $\gamma^*(\xi)$ are given in (5.14). We assume $c^*(\xi) \leq 1$, i.e., $\lambda \leq \frac{\xi}{2\xi+1}$. Since $1 - c^*(\xi)$ and $c^*(\xi) + \gamma^*(\xi) - 1$ are increasing and decreasing in ξ, it follows that $J_1(\xi)$ is increasing at ξ if and only if $1 - c^*(\xi) \geq k(c^*(\xi) + \gamma^*(\xi) - 1)$, i.e.,

$$\xi \ \geq \ \frac{(k+1)\lambda}{1 - (k+2)\lambda}, \quad 0 < k \leq \frac{1 - 2\lambda}{\lambda}, \quad \lambda = \frac{c_1 + \gamma_1 - 1}{c_1}. \tag{5.20}$$

Noting

$$\mathcal{P}_F^{\varphi_1}(a) = \bigcup_{\varphi_1(c,\gamma) \leq a} \mathcal{P}_F(c, \gamma) = \mathcal{P}_F\left(1 - a, \frac{(k+1)a}{k}\right),$$

we have

$$b_T^{\varphi_1}(J_1(a), F_\eta) = B_T\left(1 - J_1(a), \left(\frac{k+1}{k}\right) J_1(a); F_\eta\right).$$

Hence, by Theorem 5.1 we obtain the following important result which gives a lower bound on $B_T(c, \gamma; F)$ for $\frac{1}{2} < c < 1$.

Theorem 5.2 *Let (c_1, γ_1) be a given point in Ω_1. If T is an estimating functional of θ, then for each $W_\lambda \in \mathcal{W}_{F_0,\lambda}$ it holds that*

$$\sup_{\eta: \rho(\theta,\eta) \leq \tilde{b}_{v,W_\lambda}((1-\lambda)\frac{a}{1+a}, F_\theta)} B_T\left(1 - J_1(a), \left(\frac{k+1}{k}\right) J_1(a); F_\eta\right) \geq \frac{1}{2} \tilde{b}_{v,W_\lambda}\left((1-\lambda)\frac{a}{1+a}, F_\theta\right),$$

where a and k satisfy (5.20) with ξ replaced by a.

Next, in order to obtain a lower bound on $B_T(c,\gamma;F)$ for $c \geq 1$ we treat the case $c \geq 1$ and restrict Ω to its subset Ω_2 defined as

$$\Omega_2 = \left\{(c,\gamma)\ :\ 1 \leq c < \infty,\ 0 \leq \gamma < \frac{1}{2}\right\}.$$

Let

$$\varphi_2(c,\gamma) \quad = \quad \varphi^{(2)}_{k,\lambda}(c,\gamma) = \max(c-1, k\gamma - 1), \tag{5.21}$$

where k is a given positive real number. Then we have

$$J_2(\xi) = J^{(2)}_{k,\lambda}(\xi) = \varphi^{(2)}_{k,\lambda}(c^*(\xi), \gamma^*(\xi)) = \max(c^*(\xi) - 1, k\gamma^*(\xi) - 1). \tag{5.22}$$

We assume $c^*(\xi) \geq 1$, i.e., $\lambda \geq \frac{\xi}{2\xi+1}$. Since $c^*(\xi)$ and $\gamma^*(\xi)$ are decreasing and increasing in ξ, respectively, it follows that $J_2(\xi)$ is increasing in ξ if and only if $c^*(\xi) - 1 \leq k\gamma^*(\xi) - 1$, i.e.,

$$\xi \geq \frac{1}{(1-\lambda)k - 1}, \quad 0 < k \leq \frac{1}{1-\lambda}, \quad \lambda = \frac{c_2 + \gamma_2 - 1}{c_2}. \tag{5.23}$$

We also see

$$\mathcal{P}^{\varphi_2}_F(a) = \bigcup_{\varphi_2(c,\gamma) \leq a} \mathcal{P}_F(c,\gamma) = \mathcal{P}_F\left(a+1,\ \frac{a+1}{k}\right).$$

Hence, from the condition (5.23) with ξ replaced by a it follows that

$$b^{\varphi_2}_T(J_2(a),\ F_\eta) \quad = \quad B_T\left(J_2(a)+1,\ \frac{1}{k}(J_2(a)+1);\ F_\eta\right). \tag{5.24}$$

Thus, by Theorem 5.1 we obtain a lower bound on $B_T(c,\gamma;F)$ for $c \geq 1$.

Theorem 5.3 *Let (c_2,γ_2) be a given point in Ω_2. If T is an estimating functional of θ, then for each $W_\lambda \in \mathcal{W}_{F_0,\lambda}$ it holds that*

$$\sup_{\eta\ :\ \rho(\theta,\eta) \leq \tilde{b}_{v,W_\lambda}\left((1-\lambda)\frac{a}{1+a},\ F_\theta\right)} B_T\left(J_2(a)+1,\ \frac{1}{k}(J_2(a)+1);\ F_\eta\right) \quad \geq \quad \frac{1}{2}\tilde{b}_{v,W_\lambda}\left((1-\lambda)\frac{a}{1+a},\ F_\theta\right), \tag{5.25}$$

where a and k satisfies the inequality (5.23) with ξ replaced by a.

6. Lower bounds on the maximum bias in the location parameter case

6.1 Lower bounds

Let $\mathcal{X}$ be the real line R and let $F_\theta(x) = F_0(x-\theta)$, where F_0 is a distribution with a density f_0 symmetric about the origin. We consider the following neighborhood of F_θ given in (3.1).

$$\mathcal{P}_{F_\theta}(c,\gamma) \quad = \quad \{G\ :\ G(x) = c\,(F_0 - W)(x-\theta) + \gamma\, K(x), x \in \mathcal{X}, W \in \mathcal{W}_{F_0,\lambda}, K \in \mathcal{M}\}. \tag{6.1}$$

We let $\rho(\theta,\eta) = |\theta - \eta|$. An estimate T is said to be location equivariant if it satisfies

$$T(G_\theta) = T(G) + \theta,\ \ ^\forall\theta \in \Theta,\ ^\forall G \in \mathcal{M},$$

where $G_\theta(x) = G(x-\theta)$. For a location equivariant estimate T, we have

$$\begin{aligned} b_T^\varphi(J_\lambda(a),\ F_\theta) &= b_T^\varphi(J_\lambda(a),\ F_0), \quad {}^\forall\theta \in \Theta, \\ \tilde{b}_{v,W}(a,\ F_\theta) &= \tilde{b}_{v,W}(a,\ F_0), \quad {}^\forall\theta \in \Theta. \end{aligned}$$

Therefore, in this case, Theorems 5.1, 5.2 and 5.3 are expressed as follows.

Corollary 6.1 *Suppose that* $\{F_\theta(x) = F_0(x-\theta)\}$ *is a location parametric family and let* $(c_0,\ \gamma_0)$ *be a given point in* Ω. *If* T *is an estimating functional of* θ, *then for each* $W_\lambda \in W_{F_0,\lambda}$,

$$b_T^\varphi(J_\lambda(a),\ F_0) \ \geq\ \frac{1}{2}\tilde{b}_{v,W_\lambda}\left((1-\lambda)\frac{a}{1+a},\ F_0\right), \tag{6.2}$$

where λ *is given in (5.10).*

Corollary 6.2 *Suppose that* $\{F_\theta(x) = F_0(x-\theta)\}$ *is a location parametric family and let* (c_1, γ_1) *be a given point in* Ω_1. *If* T *is an estimating functional of* θ, *then for each* $W_\lambda \in \mathcal{W}_{F_0,\lambda}$ *it holds that*

$$B_T\left(1 - J_1(a), \left(\frac{k+1}{k}\right) J_1(a); F_0\right) \geq \frac{1}{2}\tilde{b}_{v,W_\lambda}\left((1-\lambda)\frac{a}{1+a}, F_0\right), \tag{6.3}$$

where a *and* k *satisfy (5.20) with* ξ *replaced by* a.

Corollary 6.3 *Suppose that* $\{F_\theta(x) = F_0(x-\theta)\}$ *is a location parametric family and let* (c_2, γ_2) *be a given point in* Ω_2. *If* T *is an estimating functional of* θ, *then for each* $W_\lambda \in \mathcal{W}_{F_0,\lambda}$ *it holds that*

$$B_T\left(J_2(a)+1,\ \frac{1}{k}(J_2(a)+1);\ F_0\right) \ \geq\ \frac{1}{2}\tilde{b}_{v,W_\lambda}\left((1-\lambda)\frac{a}{1+a},\ F_0\right), \tag{6.4}$$

where a *and* k *satisfies (5.23) with* ξ *replaced by* a.

6.2 The choice of W and λ

Let us investigate $\tilde{b}_{v,W_\lambda}$ in the lower bounds in (6.3) and (6.4). Since $\tilde{b}_{v,W_\lambda}$ depends on W_λ, we need to use W_λ which makes the lower bounds as large as possible. To this end, we propose the following $\tilde{F}_{0,W_{1\lambda}} = F_0 - W_{1\lambda}$:

$$\begin{aligned} W_{1\lambda}(B) &= F_0(B \cap [-z_\lambda,\ z_\lambda]) - f_0(z_\lambda)\mu(B \cap [-z_\lambda, z_\lambda]), \quad B \in \mathcal{B}, \\ \tilde{F}_{0,W_{1\lambda}}(x) &= \begin{cases} F_0(x), & x \leq -z_\lambda \\ F_0(-z_\lambda) + (x+z_\lambda)f_0(z_\lambda), & -z_\lambda < x \leq z_\lambda, \\ F_0(x) - \lambda, & x > z_\lambda, \end{cases} \end{aligned} \tag{6.5}$$

where z_λ is the constant satisfying

$$F_0(z_\lambda) - z_\lambda f_0(z_\lambda) \ =\ \frac{1+\lambda}{2}, \quad \lambda = \frac{c+\gamma-1}{c}. \tag{6.6}$$

Let $\tilde{\mathcal{F}}_\lambda$ be the set of all $\tilde{F}_{0,W_\lambda}$ that have even and unimodal densities. The following theorem shows that $W_{1\lambda}$ is the best in $\tilde{\mathcal{F}}_\lambda$.

Theorem 6.1 *Let (c_0, γ_0) be a given point in Ω. Then*

$$(6.7)\qquad \tilde{b}_{v,W_{1\lambda}}(t,\ F_0) = \sup_{\tilde{F}_{0,W_\lambda}\in\tilde{\mathcal{F}}_\lambda} \tilde{b}_{v,W_\lambda}(t, F_0), \quad \lambda \le {}^{\forall}t \le 1-\lambda,$$

where

$$(6.8)\qquad \tilde{b}_{v,W_{1\lambda}}(t,\ F_0) = \begin{cases} \dfrac{t}{f_0(z_\lambda)}, & 0 \le t \le 2\,F_0(z_\lambda) - (1+\lambda), \\ 2\,F_0^{-1}\left(\dfrac{1+t+\lambda}{2}\right), & 2\,F_0(z_\lambda) - (1+\lambda) \le t \le 1-\lambda. \end{cases}$$

As easily seen from (5.21) and (5.25), we have

$$(6.9)\qquad \begin{aligned} \lambda &\le \frac{a}{k+1+(k+2)a} \qquad for\ \ \frac{1}{2} < c < 1, \\ \lambda &\le \frac{(k-1)a-1}{ka} \qquad for\ \ c \ge 1. \end{aligned}$$

Hence the following lemma implies that $\lambda(=\lambda^*)$ satisfying the equality in (6.9) is the best for the lower bound with respect to $W_{1\lambda}$.

Lemma 6.1 *For given a it holds that*

$$\tilde{b}_{v,W_{1\lambda}}\left((1-\lambda)\frac{a}{1+a}; F_0\right) \ \ is\ increasing\ in\ \lambda.$$

In order to investigate the accuracy of the lower bounds in (6.3) and (6.4) we consider the median T_M. By Theorem 3.1 we see

$$(6.10)\qquad \begin{aligned} B_{T_M}\left(1-J_1, \left(\frac{k+1}{k}\right)J_1; F_0\right) &= F_0^{-1}\left(\frac{k+2J_1}{2k(1-J_1)}\right), \\ B_{T_M}\left(J_2+1, \frac{1}{k}(J_2+1); F_0\right) &= F_0^{-1}\left(\frac{2(k+1)(J_2+1)-k}{2k(J_2+1)}\right). \end{aligned}$$

Since J_1 and J_2 are increasing in a, we have

$$(6.11)\qquad \begin{aligned} J_1 &= J_{k,\lambda}^{(1)}(a) = 1-c^*(a) = \frac{a-\lambda(2a+1)}{(1-\lambda)(2a+1)}, \quad a = \frac{\lambda+(1-\lambda)J_1}{1-2\lambda-2(1-\lambda)J_1}. \\ J_2 &= J_{k,\lambda}^{(2)}(a) = k\gamma^*(a) - 1 = \frac{(k-2)a-1}{1+2a}, \quad a = \frac{J_2+1}{k-2(J_2+1)}, \end{aligned}$$

Therefore from Corollaries 6.2, 6.3, Lemma 6.1, (6.9) and (6.11) we obtain the following results.

Theorem 6.2 *Suppose that $\{F_\theta(x) = F_0(x-\theta)\}$ is the location parametric family and let (c_1, γ_1) be a given point in Ω_1. Then*

$$B_{T_M}\left(1-J_1, \left(\frac{k+1}{k}\right)J_1; F_0\right) \ge \frac{1}{2}\tilde{b}_{v,W_{1\lambda^*}}\left(\frac{(k+1)J_1}{k(1-J_1)}; F_0\right), \quad \lambda^* = \frac{J_1}{k(1-J_1)}.$$

Theorem 6.3 *Suppose that $\{F_\theta(x) = F_0(x-\theta)\}$ is the location parametric family and let (c_2, γ_2) be a given point in Ω_2. Then*

$$B_{T_M}\left(J_2+1, \frac{1}{k}(J_2+1); F_0\right) \geq \frac{1}{2}\tilde{b}_{v,W_{1\lambda^*}}\left(\frac{1}{k}; F_0\right), \quad \lambda^* = \frac{(k+1)J_2+1}{k(J_2+1)}.$$

Remark 6.1 As an example of φ different from (5.19) and (5.24) we can consider $\varphi(c,\gamma) = \frac{c+2\gamma-1}{2c}$ corresponding to the maximum bias of the median T_M. In this case, we have

$$\mathcal{P}_F^\varphi(a) = \bigcup_{\gamma=-(\frac{1}{2}-a)c+\frac{1}{2}} \mathcal{P}_F(c,\gamma) \quad \text{and} \quad b_{T_M}^\varphi(a) = B_{T_M}(1,a), \quad 0 \leq a < \frac{1}{2}.$$

7. The normal distribution model case

In this section we consider the case that the central model distribution F_0 is the standard normal distribution Φ and present some tables and figures of the implosion bias of scale estimates and the lower bounds in Theorem 6.2 and 6.3 for the median together with comments.

First we consider the scale estimates discussed in Section 4. In order to make their estimates consistent at the model Φ we take $a_1 = 2a_4 = a_5 = 1.4826$, $a_2 = 1.1926$ and $a_3 = 2.2219$. Then we can see $B^-_{\mathrm{MAD}}(c,\gamma) = B^-_{\mathrm{LMS}}(c,\gamma) = B^-_{\mathrm{L}}(c,\gamma)$. Therefore we are concerned with MAD, S and Q. Tables 7.1, 7.2 and 7.3 exhibit $B^-_{\mathrm{MAD}}(c,\gamma), B^-_{\mathrm{S}}(c,\gamma), B^-_{\mathrm{Q}}(c,\gamma)$ for selected c and γ. For clarity we denote the maximum and the minimum among the three values for the same (c,γ) by the boldface and the italics, respectively. Roughly speaking, from these tables we can observe the following features:

(i) $B^-_{\mathrm{Q}}(c,\gamma) > B^-_{\mathrm{MAD}}(c,\gamma) > B^-_{\mathrm{S}}(c,\gamma)$ for $c \leq 0.90$.

(ii) $B^-_{\mathrm{MAD}}(c,\gamma) > B^-_{\mathrm{Q}}(c,\gamma) > B^-_{\mathrm{S}}(c,\gamma)$ for $c \geq 1$ and $\gamma \geq 0.03$.

(iii) $B^-_{\mathrm{MAD}}(c,\gamma) > B^-_{\mathrm{Q}}(c,\gamma) \doteqdot B^-_{\mathrm{S}}(c,\gamma)$ for $c \geq 1$ and $\gamma < 0.03$.

Figure 7.1 shows the graphs of the implosion biase for $\gamma = 0,\ 0.2$ and 0.3 when c varies, and Figure 7.2 for $c = 0.9,\ 1.0,\ 1.5$ and 3.0 when γ varies. We can observe that the implosion biase curves are convex in c and nearly linear in γ.

Next we consider the lower bounds in Theorem 6.2 and 6.3. Table 7.4 gives $\frac{1}{2}\tilde{b}_{v,W_{1\lambda^*}}\left(\frac{\gamma}{c};\Phi\right)$ for selected c and γ. We denote $c = 1-J_1$, $\gamma = \left(\frac{k+1}{k}\right)J_1$ for $c<1$ and $c = J_2+1$, $\gamma = \frac{1}{k}(J_2+1)$ for $c \geq 1$. Note that $\frac{(k+1)J_1}{k(1-J_1)} = \frac{\gamma}{c}$ for $c<1$ and $\frac{1}{k} = \frac{\gamma}{c}$ for $c \geq 1$. Table 7.6 exhibits $B_{T_M}(c,\gamma;\Phi)$ for the same c and γ. We can observe that the larger γ is, the better the lower bound is, and that $\frac{1}{2}\tilde{b}_{v,W_{1\lambda^*}}\left(\frac{\gamma}{c};\Phi\right) = B_{T_M}(c,\gamma;\Phi)$ holds for $c+\gamma = 1$.

For comparison let us consider $W_{2\lambda} = \lambda F_0$, $\tilde{F}_{0,W_{2\lambda}} = (1-\lambda)F_0 \in \tilde{\mathcal{F}}_\lambda$ (i.e., ε-contamination case). In this case, since

$$\tilde{d}_v(\tilde{F}_{\theta,W_{2\lambda}}, \tilde{F}_{\eta,W_{2\lambda}}) = (1-\lambda)d_v(F_\theta, F_\eta),$$

Table 7.1 Implosion bias $B^-_{MAD}(c,\ \gamma)$

$c \setminus \gamma$	0.00	0.01	0.03	0.05	0.07	0.10	0.15	0.20	0.25	0.30	0.35	0.40
0.70	-	-	-	-	-	-	-	-	-	0.543	0.403	0.267
0.80	-	-	-	-	-	-	-	0.725	0.596	0.472	0.352	0.233
0.85	-	-	-	-	-	-	0.803	0.679	0.560	0.444	0.331	0.219
0.90	-	-	-	-	-	0.874	0.754	0.639	0.527	0.418	0.312	0.207
0.95	-	-	-	0.939	0.892	0.823	**0.711**	**0.603**	**0.498**	0.396	0.295	0.196
0.99	-	**0.988**	**0.942**	**0.896**	**0.852**	**0.786**	**0.680**	**0.577**	**0.477**	**0.379**	0.283	0.188
1.00	1.000	**0.977**	**0.931**	**0.886**	**0.842**	**0.777**	**0.673**	**0.571**	**0.472**	**0.376**	0.280	0.186
2.00	**0.472**	**0.463**	**0.443**	**0.424**	**0.404**	**0.376**	**0.328**	**0.280**	**0.233**	**0.186**	**0.140**	**0.093**
5.00	**0.186**	**0.183**	**0.175**	**0.168**	**0.160**	**0.149**	**0.130**	**0.112**	**0.093**	**0.074**	**0.056**	**0.037**
10.00	**0.093**	**0.091**	**0.087**	**0.084**	**0.080**	**0.074**	**0.065**	**0.056**	**0.046**	**0.037**	**0.028**	**0.019**
100.00	**0.009**	**0.009**	**0.009**	**0.008**	**0.008**	**0.007**	**0.007**	**0.006**	**0.005**	**0.004**	**0.003**	**0.002**

Table 7.2 Implosion bias $B^-_{S}(c,\ \gamma)$

$c \setminus \gamma$	0.00	0.01	0.03	0.05	0.07	0.10	0.15	0.20	0.25	0.30	0.35	0.40
0.70	-	-	-	-	-	-	-	-	-	*0.467*	*0.336*	*0.218*
0.80	-	-	-	-	-	-	-	*0.656*	*0.520*	*0.400*	*0.291*	*0.190*
0.85	-	-	-	-	-	-	*0.746*	*0.606*	*0.483*	*0.373*	*0.273*	*0.178*
0.90	-	-	-	-	-	*0.833*	*0.689*	*0.563*	*0.452*	*0.350*	*0.257*	*0.168*
0.95	-	-	-	*0.918*	*0.856*	*0.770*	*0.641*	*0.527*	*0.424*	*0.330*	*0.242*	*0.159*
0.99	-	*0.984*	*0.921*	*0.861*	*0.805*	*0.726*	*0.607*	*0.501*	*0.404*	*0.315*	*0.232*	*0.153*
1.00	1.000	*0.968*	*0.907*	*0.848*	*0.793*	*0.716*	*0.599*	*0.495*	*0.400*	*0.312*	*0.230*	*0.151*
2.00	0.400	0.391	*0.373*	*0.355*	*0.338*	*0.312*	*0.270*	*0.230*	*0.190*	*0.151*	*0.113*	*0.075*
5.00	0.151	0.148	*0.142*	*0.136*	*0.130*	*0.120*	*0.105*	*0.090*	*0.075*	*0.060*	*0.045*	*0.030*
10.00	0.075	*0.073*	*0.070*	*0.067*	*0.064*	*0.060*	*0.052*	*0.045*	*0.037*	*0.030*	*0.022*	*0.015*
100.00	*0.007*	*0.007*	*0.007*	*0.007*	*0.006*	*0.006*	*0.005*	*0.004*	*0.004*	*0.003*	*0.002*	*0.001*

Table 7.3 Implosion bias $B^-_{Q}(c,\ \gamma)$

$c \setminus \gamma$	0.00	0.01	0.03	0.05	0.07	0.10	0.15	0.20	0.25	0.30	0.35	0.40
0.70	-	-	-	-	-	-	-	-	-	**0.587**	**0.429**	**0.282**
0.80	-	-	-	-	-	-	-	**0.768**	**0.622**	**0.489**	**0.363**	**0.241**
0.85	-	-	-	-	-	-	**0.840**	**0.700**	**0.573**	**0.454**	**0.338**	**0.225**
0.90	-	-	-	-	-	**0.903**	**0.767**	**0.646**	**0.533**	**0.423**	**0.317**	**0.211**
0.95	-	-	-	**0.957**	**0.901**	**0.824**	0.708	0.601	**0.498**	**0.397**	**0.298**	**0.199**
0.99	-	0.993	0.938	0.888	0.841	0.773	0.669	0.570	0.474	0.379	**0.285**	**0.190**
1.00	**1.000**	0.973	0.921	0.873	0.827	0.762	0.660	0.563	0.468	0.374	**0.282**	**0.188**
2.00	*0.393*	*0.388*	0.376	0.364	0.352	0.332	0.297	0.259	0.219	0.178	0.135	0.091
5.00	*0.150*	*0.148*	0.144	0.140	0.135	0.128	0.115	0.101	0.086	0.070	0.054	0.036
10.00	*0.074*	0.074	0.072	0.070	0.068	0.064	0.058	0.051	0.043	0.035	0.027	0.018
100.00	0.008	0.008	0.007	0.007	0.007	0.006	0.006	0.005	0.004	0.004	0.003	0.002

the definition (5.11) of ξ becomes

$$\xi = \frac{d_v(F_\theta, F_\eta)}{1 - d_v(F_\theta, F_\eta)}, \tag{7.1}$$

where d_v is the total variation distance given in (5.18). That is to say, the use of $W_{2\lambda}$ corresponds to that of d_v. Since $W_{2\lambda}$ satisfies Lemma 6.1, Theorem 6.1 states that $W_{2\lambda}$ is inferior to $W_{1\lambda}$. These facts show that $\tilde{d}_v$ based improper distributions $\tilde{F}_{\theta,W_\lambda}$ is useful. Table 7.5 presents $\frac{1}{2}\tilde{b}_{v,W_{2\lambda}}\left(\frac{\gamma}{c};\Phi\right) = \frac{1}{2}b_v\left(\frac{\gamma}{1-\gamma}\right)$, which depends on only γ.

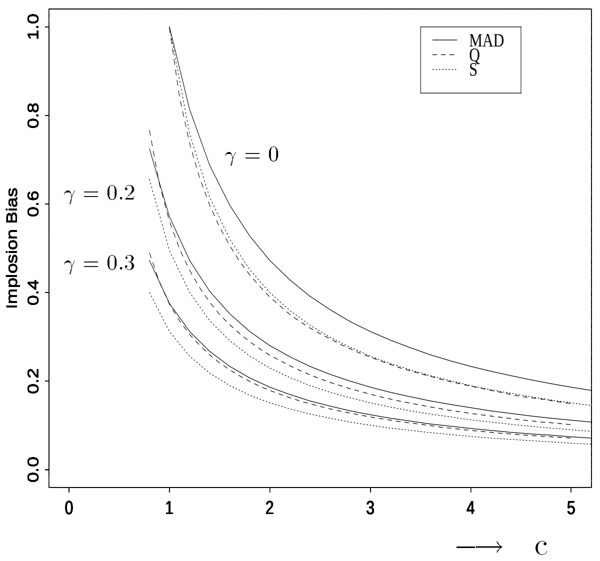

Figure7.1

Implosion bias curves $B_*^-(c, 0.0)$, $B_*^-(c, 0.2)$, $B_*^-(c, 0.3)(* = \mathrm{MAD, S, Q})$

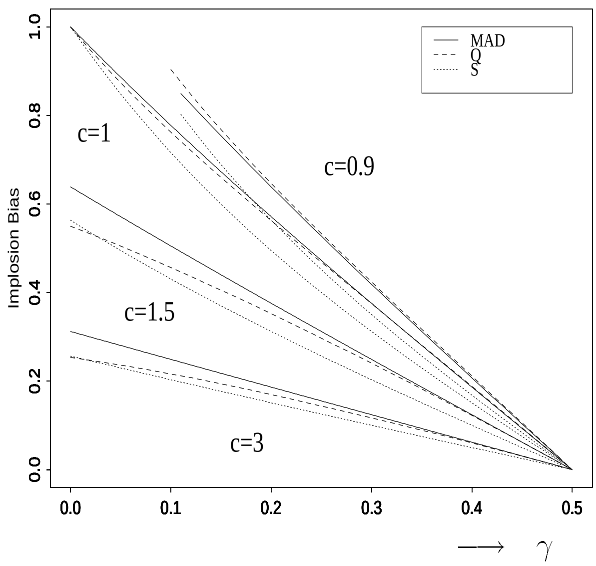

Figure7.2

Implosion bias curves $B_*^-(0.9, \gamma)$, $B_*^-(1, \gamma)$, $B_*^-(1.5, \gamma)$, $B_*^-(3, \gamma)$ $(* = MAD, S, Q)$

Table 7.4 $\tilde{b}_{v,W_{1\lambda^*}}(\gamma/c; \Phi)/2, \quad \Phi = N(0,1)$

$c \setminus \gamma$	0.00	0.01	0.03	0.05	0.07	0.10	0.15	0.20	0.25	0.30	0.35	0.40
0.70	-	-	-	-	-	-	-	-	-	0.566	0.792	1.068
0.80	-	-	-	-	-	-	-	0.319	0.482	0.664	0.883	1.150
0.85	-	-	-	-	-	-	0.223	0.359	0.513	0.696	0.918	1.187
0.90	-	-	-	-	-	0.140	0.253	0.382	0.537	0.723	0.947	1.220
0.95	-	-	-	0.066	0.102	0.158	0.268	0.400	0.557	0.746	0.973	1.250
0.99	-	0.013	0.042	0.074	0.108	0.166	0.278	0.412	0.571	0.762	0.992	1.271
1.00	0.000	0.013	0.043	0.075	0.110	0.168	0.280	0.414	0.575	0.766	0.997	1.276
2.00	0.000	0.021	0.064	0.109	0.158	0.235	0.379	0.547	0.742	0.971	1.242	1.566
5.00	0.000	0.026	0.079	0.136	0.195	0.289	0.464	0.664	0.895	1.163	1.478	1.853
10.00	0.000	0.029	0.089	0.151	0.217	0.322	0.515	0.736	0.989	1.283	1.627	2.035
100.00	0.000	0.037	0.113	0.193	0.276	0.408	0.652	0.927	1.243	1.607	2.030	2.528

Table 7.5 $\tilde{b}_{v,W_{2\lambda^*}}(\gamma/(1-\gamma); \Phi)/2, \quad \Phi = N(0,1)$

γ	0.00	0.01	0.03	0.05	0.07	0.10	0.15	0.20	0.25	0.30	0.35	0.40
bias	0.000	0.013	0.039	0.066	0.094	0.140	0.223	0.319	0.431	0.566	0.736	0.967

Table 7.6 $B_{T_M}(c, \gamma; \Phi), \quad \Phi = N(0,1)$

$c \setminus \gamma$	0.00	0.01	0.03	0.05	0.07	0.1	0.15	0.2	0.25	0.3	0.35	0.4
0.70	-	-	-	-	-	-	-	-	-	0.566	0.792	1.068
0.80	-	-	-	-	-	-	-	0.319	0.489	0.674	0.887	1.150
0.85	-	-	-	-	-	-	0.223	0.377	0.541	0.722	0.929	1.187
0.90	-	-	-	-	-	0.140	0.282	0.431	0.589	0.765	0.967	1.221
0.95	-	-	-	0.066	0.119	0.199	0.336	0.480	0.634	0.805	1.003	1.252
0.99	-	0.013	0.063	0.114	0.165	0.243	0.376	0.516	0.667	0.834	1.030	1.276
1.00	0.000	0.025	0.075	0.126	0.176	0.253	0.385	0.524	0.674	0.842	1.036	1.282
2.00	0.674	0.690	0.722	0.755	0.789	0.842	0.935	1.036	1.150	1.282	1.440	1.645
5.00	1.282	1.293	1.317	1.341	1.366	1.405	1.476	1.555	1.645	1.751	1.881	2.054
10.00	1.645	1.655	1.675	1.695	1.717	1.751	1.812	1.881	1.960	2.054	2.170	2.326
100.00	2.576	2.583	2.597	2.612	2.628	2.652	2.697	2.748	2.807	2.878	2.968	3.09

8. Proofs

Proof of Lemma 2.1. The assertion (ii) follows from the assertion (i) and the fact that

$$\int_{-\infty}^{\infty} G(x)g(x)dx \quad \leq \quad \int_{-\infty}^{\infty} \hat{G}(x)\hat{g}(x)dx = \frac{1}{2}m^2, \quad {}^{\forall}g \in \mathcal{F}_0.$$

We prove the assertion (i). Let $g \in \mathcal{F}_0$. First assume $\int_{-\infty}^{\infty} g(x)dx < m$. Then it is clear that there exists $g_1 \in \mathcal{F}_0$ such that $g \leq g_1$ and $\int_{-\infty}^{\infty} g_1(x)dx = m$. In this case, it readily follows that

$$\int_{-\infty}^{\infty}\{G(x+t) - G(x)\}g(x)dx \leq \int_{-\infty}^{\infty}\{G_1(x+t) - G_1(x)\}g_1(x)dx, \qquad for \ {}^{\forall}t \geq 0,$$

because of $G_1(x+t) - G_1(x) \geq 0$ for ${}^{\forall}x \in R$. Hence, we assume $\int_{-\infty}^{\infty} g(x)dx = m$. Let

$$h_{g,t}(x) \quad = \quad G(x+t) - G(x).$$

For simplicity we hereafter omit the subscript t of $h_{g,t}$. Since

$$\int_{-\infty}^{\infty} h_g(x)g(x)dx \quad = \quad \int_{0}^{\infty} G(h_g^{-1}[u, \ \infty))du,$$

the inequality

$$G(h_g^{-1}[u, \ \infty)) \leq \hat{G}(h_{\hat{g}}^{-1}[u, \ \infty)), \qquad for \ \ 0 \leq {}^{\forall}u < \infty, \tag{8.1}$$

is sufficient for proving the assertion (i). To show (8.1) we consider three cases (1) $0 \leq t < a$, (2) $a \leq t < 2a$ and (3) $t \geq 2a$.

The proofs in (2) and (3) are similar to that in (1). Hence we give only the proof in the case of (1). Let

$$L_g(t) \quad = \quad \int_{-\infty}^{\infty} h_{g,t}(x)dx.$$

Differentiating $L_g(t)$ with respect to t, we have

$$L_g'(t) \quad = \quad m.$$

Since $L_g(0) = 0$, it follows that

$$L_g(t) = m\,t.$$

Thus the area enclosed by the graph of $y = h_g(x)$ and the x-axis does not depend on g. In order to observe the graph of $y = h_g(x)$ in more detail, we differentiate $y = h_g(x)$ and $y = h_{\hat{g}}(x)$ with respect to x, and obtain

$$\begin{aligned} \frac{\partial}{\partial x}h_g(x) \quad &\leq \quad f(x+t) = \frac{\partial}{\partial x}h_{\hat{g}}(x), \quad -a-t \leq x \leq -a, \\ \frac{\partial}{\partial x}h_g(x) \quad &\geq \quad -f(x) = \frac{\partial}{\partial x}h_{\hat{g}}(x), \qquad a-t \leq x \leq a. \end{aligned} \tag{8.2}$$

From (8.2) it follows that the set $\{x \mid h_{\hat{g}}(x) \geq h_g(x)\}$ is an interval, which is given by one of the forms $(-\infty, b]$, $[a, \ b]$ and $[a, \ \infty)$, where $-\infty < a < b < \infty$. Figure 1 shows the graphs of $y = h_g(x)$, $y = h_{\hat{g}}(x)$ and $h_f(x)$, where $h_f(x) = F(x+t) - F(x)$ and $F(x) = \int_{-\infty}^{x} f(t)dt$.

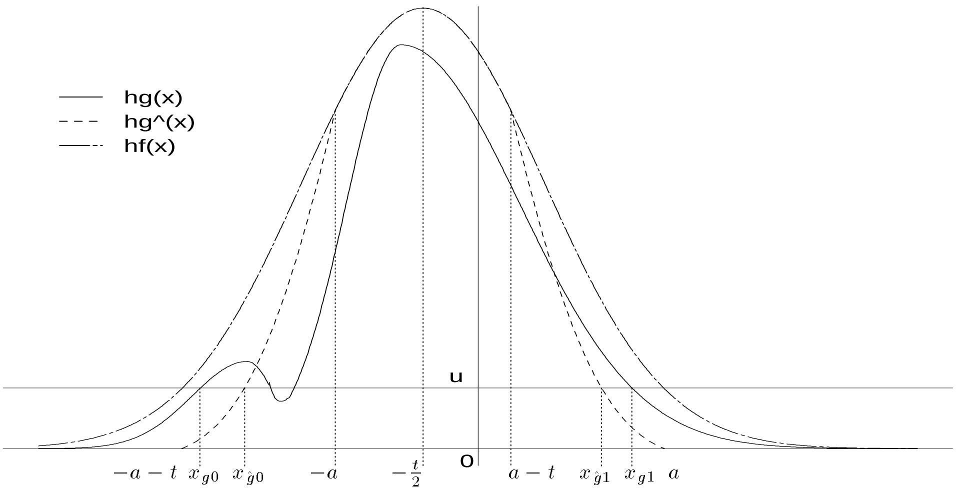

Figure 8.1: The graphs of $y = h_g(x),\ y = h_{\hat{g}}(x),\ y = h_f(x)$

For $0 \le u < \infty$ we define

$$\begin{aligned} x_{g0} &= \inf\{x \mid h_g(x) \ge u\}, \\ x_{g1} &= \sup\{x \mid h_g(x) \ge u\}. \end{aligned}$$

Then it is easy to see

$$G(h_g^{-1}([u,\ \infty))) \le G([x_{g0},\ x_{g1}]) = G(x_{g1}) - G(x_{g0}),$$
$$\hat{G}(h_{\hat{g}}^{-1}([u,\ \infty))) = \hat{G}([x_{\hat{g}0},\ x_{\hat{g}1}]) = \hat{G}(x_{\hat{g}1}) - \hat{G}(x_{\hat{g}0}).$$

We consider the following three cases:

(a) For $u \ge h_{\hat{g}}(-\frac{t}{2})$, we have

$$\hat{G}(h_{\hat{g}}^{-1}([u,\ \infty))) = \hat{G}(\phi) = 0 = G(\phi) = G(h_g^{-1}([u,\ \infty))).$$

(b) For $h_{\hat{g}}(-a) \le u < h_{\hat{g}}(-\frac{t}{2})$, we have

$$[-a,\ a] \supset [x_{\hat{g}0},\ x_{\hat{g}1}] \supset [x_{g0},\ x_{g1}],$$

and hence

$$\begin{aligned} \hat{G}(h_{\hat{g}}^{-1}([u,\ \infty))) &= \hat{G}([x_{\hat{g}0},\ x_{\hat{g}1}]) \ge \hat{G}([x_{g0},\ x_{g1}]) \\ &\ge G([x_{g0},\ x_{g1}]) \ge G(h_g^{-1}([u,\ \infty))). \end{aligned}$$

(c) For $0 \le u < h_{\hat{g}}(-a)$, we have

$$h_{\hat{g}}(x_{\hat{g}1}) = h_g(x_{g1}) = u,$$

and hence

$$\hat{G}(x_{\hat{g}1} + t) - \hat{G}(x_{\hat{g}1}) = G(x_{g1} + t) - G(x_{g1}),$$

Since $x_{\hat{g}1} + t > a$, it follows that

$$\hat{G}(x_{\hat{g}1} + t) = m \ge G(x_{g1} + t).$$

Hence we have

$$G(x_{g1}) \leq \hat{G}(x_{\hat{g}1}).$$

Since $x_{\hat{g}0} \leq -a$, we also have $\hat{G}(x_{\hat{g}0}) = 0$. Therefore

$$\begin{aligned} \hat{G}(h_{\hat{g}}^{-1}([u,\ \infty))) &= \hat{G}(x_{\hat{g}1}) - \hat{G}(x_{\hat{g}0}) = \hat{G}(x_{\hat{g}1}) \\ &\geq G(x_{g1}) \geq G(x_{g1}) - G(x_{g0}) \\ &= G(h_g^{-1}([u,\ \infty))). \end{aligned}$$

The results of (a), (b) and (c) imply that the proposition holds for $0 \leq t < a$. $\square$

Proof of Theorem 4.1 Let $\hat{F}$ be given by (2.7), i.e.,

$$\hat{F} \;=\; c\,(F_0 - \hat{W}) + \gamma\,\Delta_0.$$

Then, from the symmetry and unimodality of f_0 it follows that for $^\forall F \in \mathcal{P}_{F_0}(c,\ \gamma)$ and for $^\forall t \geq 0$

$$\begin{aligned} P_F(|X - \operatorname*{med}_F Y| \leq t) &\leq \sup_y P_F(|X - y| \leq t) \\ &\leq P_{\hat{F}}(|X| \leq t) = P_{\hat{F}}(|X - \operatorname*{med}_{\hat{F}} Y| \leq t). \qquad (\operatorname*{med}_{\hat{F}} Y = 0) \end{aligned}$$

This implies the distribution G_F of $|X - \mathrm{med}_F\, Y|$ under F is stochastically smallest under $F = \hat{F}$. Hence we have

$$B^-_{\mathrm{MAD}}(c,\ \gamma) = \mathrm{MAD}(\hat{F}) = a_1\, G_{\hat{F}}^{-1}\left(\frac{1}{2}\right).$$

Noting

$$G_{\hat{F}}(t) \;=\; \begin{cases} \gamma, & \text{if } t = 0, \\ 2cF_0(t) - c + \gamma, & \text{if } 0 < t \leq F_0^{-1}(\frac{c-\gamma+1}{2c}), \\ 1, & \text{if } t \geq F_0^{-1}(\frac{c-\gamma+1}{2c}), \end{cases}$$

we have

$$G_{\hat{F}}^{-1}\left(\frac{1}{2}\right) \;=\; F_0^{-1}\left(\frac{2c - 2\gamma + 1}{4c}\right).$$

It is obvious that $G_{\hat{F}}^{-1}(\frac{1}{2}) = 0$ holds for $\gamma \geq \frac{1}{2}$. This completes the proof. $\square$

Proof of Theorem 4.2. Suppose that $0 \leq \gamma < \frac{1}{2}$. We first show that $B^-_{\mathrm{S}}(c,\ \gamma) = \mathrm{S}(\hat{F})$, where $\hat{F}$ is given by (2.7). For each x let a_x be the positive number such that

$$F_0(x + a_x) - F_0(x - a_x) \;=\; \frac{1-\gamma}{c},$$

and let

$$F^*_x \;=\; c\,(F_0 - W^*_x) + \gamma\Delta_x, \tag{8.3}$$

where

$$W_x^*(B) \;=\; F_0(B \cap [x - a_x,\ x + a_x]^c), \quad {}^\forall B \in \mathcal{B},$$

and Δ_x is the probability measure with mass 1 at x. Then it is clear that

$$g_{F_x^*}(x) \;=\; \inf\{g_F(x)\ :\ F \in \mathcal{P}_{F_0}(c,\ \gamma)\}, \quad {}^\forall x \in R.$$

Since $0 \le \gamma < \frac{1}{2}$, we can see

$$P_{F_x^*}(|x - Y| \le g_{F_x^*}(x)) \;=\; \frac{1}{2}.$$

Hence, by $g_{F_x^*}(x) < a_x$ we have

$$F_0(x + g_{F_x^*}(x)) - F_0(x - g_{F_x^*}(x)) \;=\; \frac{1-2\gamma}{2c}.$$

From the symmetry and unimodality of f_0, it follows that $g_{F_x^*}(x)$ and $g_{\hat{F}}(x)$ are strictly increasing in $|x|$ and symmetric about the origin. Hence, for any $F \in \mathcal{P}_{F_0}(c,\ \gamma)$

$$\begin{aligned} \mathrm{S}(F) \;&=\; a_2 \operatorname*{med}_F g_F(X) \ge a_2 \operatorname*{med}_F g_{F_X^*}(X) = a_2\, g_{F^*_{\mathrm{med}_F |X|}}(\operatorname*{med}_F |X|) \\ &\ge\; a_2\, g_{F^*_{\mathrm{med}_{\hat{F}} |X|}}(\operatorname*{med}_{\hat{F}} |X|) = a_2\, g_{\hat{F}}(\operatorname*{med}_{\hat{F}} |X|) = \mathrm{S}(\hat{F}). \end{aligned}$$

This implies $B_{\mathrm{S}}^-(c,\ \gamma) = \mathrm{S}(\hat{F})$.

Secondly, we show that $\mathrm{S}(\hat{F}) = a_2\, g^-\left(F_0^{-1}\left(\frac{2c-2\gamma+1}{4c}\right)\right)$. We note that $g_{\hat{F}}(x)$ is the smallest positive solution of

$$\begin{aligned} &c\{(F_0 - \hat{W})(x + g_{\hat{F}}(x)) - (F_0 - \hat{W})(x - g_{\hat{F}}(x))\} \\ &\qquad + \gamma\{\Delta_0(x + g_{\hat{F}}(x)) - \Delta_0(x - g_{\hat{F}}(x))\} + \gamma I(x - g_{\hat{F}}(x) = 0) \ge \frac{1}{2}. \end{aligned}$$

For $0 \le t < x$, we have

$$P_{\hat{F}}(|x - Y| \le t) \;<\; \frac{1}{2},$$

and hence $g_{\hat{F}}(x) \ge x$. We also have

$$x = g_{\hat{F}}(x) \quad \text{iff} \quad x \ge F_0^{-1}\left(\frac{c - 2\gamma + 1}{2c}\right)/2.$$

Next, for $x < g_{\hat{F}}(x)$ we can see

$$F_0(x + g_{\hat{F}}(x)) - F_0(x - g_{\hat{F}}(x)) \;=\; \frac{1-2\gamma}{2c},$$

which implies $g_{\hat{F}}(x) = g^-(x)$. From the symmetry of $\hat{F}$ it follows that

$$\operatorname*{med}_{\hat{F}} |X| = \hat{F}^{-1}\left(\frac{3}{4}\right) = F_0^{-1}\left(\frac{2c - 2\gamma + 1}{4c}\right).$$

Therefore we obtain

$$\mathrm{S}(\hat{F}) = a_2\, g_{\hat{F}}(\operatorname*{med}_{\hat{F}} |X|) = a_2\, g^-\left(F_0^{-1}\left(\frac{2c - 2\gamma + 1}{4c}\right)\right).$$

Because

$$F_0^{-1}\left(\frac{2c - 2\gamma + 1}{4c}\right) \;<\; F_0^{-1}\left(\frac{c - 2\gamma + 1}{2c}\right)/2,$$

(4.4) and (4.5) provide an implicit determination of $B_{\mathrm{S}}^-(c,\ \gamma)$.

When $\gamma \ge \frac{1}{2}$, it immediately follows from $S(\hat{F}) = 0$ that $B_S^-(c,\ \gamma) = 0$. $\square$

Proof of Theorem 4.3. When $0 \le \gamma < \frac{1}{2}$, we can easily see $\mathrm{Q}(\hat{F}) = a_3 K_{\hat{F}}^{-1}(\frac{5}{8}) > 0$, and hence

$$P_{\hat{F}\times\hat{F}}(X - Y \le a_3^{-1}\mathrm{Q}(\hat{F})) \;=\; \frac{5}{8}.$$

This equation is written as

$$\begin{aligned} &c^2((F_0 - \hat{W}) \times (F_0 - \hat{W}))(X - Y \le a_3^{-1}\mathrm{Q}(\hat{F})) \\ &\quad +2c\gamma((F_0 - \hat{W}) \times \Delta_0)(X - Y \le a_3^{-1}\mathrm{Q}(\hat{F})) + \gamma^2(\Delta_0 \times \Delta_0)(X - Y \le a_3^{-1}\mathrm{Q}(\hat{F})) = \frac{5}{8}, \end{aligned}$$

which reduces to (4.7).

On the other hand, when $\gamma \ge \frac{1}{2}$, we see

$$\mathcal{P}_{\hat{F}\times\hat{F}}(X - Y < 0) < \frac{5}{8} \le \mathcal{P}_{\hat{F}\times\hat{F}}(X - Y \le 0).$$

This implies $\mathrm{Q}(\hat{F}) = 0$. □

Proof of Theorem 4.4. When $0 \le \gamma < \dfrac{1}{2}$, we can see for any $F \in \mathcal{P}_{F_0}(c,\ \gamma)$

$$\inf_{t\in[0,\ \frac{1}{2}]} \left| F^{-1}\left(t + \frac{1}{2}\right) - F^{-1}(t) \right| \ge \left| \hat{F}^{-1}\left(\frac{3}{4}\right) - \hat{F}^{-1}\left(\frac{1}{4}\right) \right|,$$

where where $\hat{F}$ is given in (2.7). This implise

$$B^-_{\mathrm{LMS}}(c, \gamma) = \mathrm{LMS}(\hat{F}) = a_4 \left| \hat{F}^{-1}\left(\frac{3}{4}\right) - \hat{F}^{-1}\left(\frac{1}{4}\right) \right|.$$

Since

$$\hat{F}^{-1}\left(\frac{3}{4}\right) = F_0^{-1}\left(\frac{2c - 2\gamma + 1}{4c}\right) \quad \text{and} \quad \hat{F}^{-1}\left(\frac{1}{4}\right) = F_0^{-1}\left(\frac{2c + 2\gamma - 1}{4c}\right),$$

we have

$$B^-_{\mathrm{LMS}}(c, \gamma) \;=\; a_4 \left| \hat{F}^{-1}\left(\frac{3}{4}\right) - \hat{F}^{-1}\left(\frac{1}{4}\right) \right| = 2a_4\, F_0^{-1}\left(\frac{2c - 2\gamma + 1}{4c}\right).$$

When $\gamma \ge \frac{1}{2}$, we can easily see $\mathrm{LMS}(\hat{F}) = 0$ and hence $B^-_{\mathrm{LMS}}(c, \gamma) = 0$. □

Proof of Theorem 4.5. Suppose that $0 \le \gamma < \frac{1}{2}$. Then, for any $x \in R$,

$$g_{F_x^*}(x) \;=\; \inf\{g_F(x)\ :\ F \in \mathcal{P}_{F_0}(c,\ \gamma)\},$$

where F_x^* is given by (8.3). Hence

$$\begin{aligned} B^-_{\mathrm{L}}(c,\ \gamma) &= a_5\ \inf\{\inf_x g_F(x) \mid F \in \mathcal{P}_{F_0}(c,\ \gamma)\} \\ &= a_5\ \inf_x\{\inf g_F(x) \mid F \in \mathcal{P}_{F_0}(c,\ \gamma)\} \\ &= a_5\ \inf_x g_{F_x^*}(x) = a_5\ g_{F_0^*}(0) = a_5\ \operatorname*{med}_{\hat{F}} |Y| \\ &= a_5\ F^{-1}\left(\frac{3}{4}\right) = a_1\ F^{-1}\left(\frac{2c - 2\gamma + 1}{4c}\right). \end{aligned}$$

When $\gamma \ge \frac{1}{2}$, we can easily see that $B^-_{\mathrm{L}}(c,\ \gamma) = 0$. □

Proof of Lemma 6.1 We first note that z_λ is strictly increasing in λ. By (6.6) and (6.8) we have for $0 \le (1-\lambda)\frac{a}{1+a} \le 2F_0(z_\lambda) - (1+\lambda)$

$$\tilde{b}_{v,W_{1\lambda}}\left((1-\lambda)\frac{a}{1+a}; F_0\right) = \frac{2a}{1+a}\left(\frac{1-F_0(z_\lambda)+z_\lambda f_0(z_\lambda)}{f_0(z_\lambda)}\right).$$

Then it follows from the unimodality and symmetry of f_0 that

$$\frac{\partial}{\partial z_\lambda}\tilde{b}_{v,W_{1\lambda}}\left((1-\lambda)\frac{a}{1+a}; F_0\right) = -\frac{2a(1-F_0(z_\lambda))f_0'(z_\lambda)}{(1+a)f_0(z_\lambda)^2} > 0. \tag{8.4}$$

Also, by (6.6) and (6.8) we have for $2F_0(z_\lambda) - (1+\lambda) \le (1-\lambda)\frac{a}{1+a} \le 1-\lambda$,

$$\tilde{b}_{v,W_{1\lambda}}\left((1-\lambda)\frac{a}{1+a}; F_0\right) = F_0^{-1}\left(\frac{1+2a+\lambda}{2(1+a)}\right). \tag{8.5}$$

Therefore the lemma follows from (8.4) and (8.5). □

Proof of Theorem 6.1. Since $\tilde{F}_{0,W_{1\lambda}}$ has an even and unimodal density, we have as in (3.2) of He and Simpson (1993),

$$\tilde{d}_v(\tilde{F}_{\frac{\eta}{2},W_{1\lambda}}, \tilde{F}_{-\frac{\eta}{2},W_{1\lambda}}) = 2\,\tilde{F}_{0,W_{1\lambda}}\left(\frac{|\eta|}{2}\right) - (1-\lambda). \tag{8.6}$$

It is easy to see that $\tilde{b}_{v,W_{1\lambda}}(t,\ F_0)$ is the solution $|\eta|$ of the equation

$$2\,\tilde{F}_{0,W_{1\lambda}}\left(\frac{|\eta|}{2}\right) - (1-\lambda) = t. \tag{8.7}$$

Hence, by solving (8.7) in $|\eta|$ we obtain (6.8).

Let $\tilde{F}_{0,W_\lambda}$ be any element of $\tilde{\mathcal{F}}_0$. Then, by the unimodality and symmetry of $\tilde{F}_{0,W_{1\lambda}}$ and $\tilde{F}_{0,W_\lambda}$ we have

$$\tilde{F}_{0,W_{1\lambda}}(x) \le \tilde{F}_{0,W_\lambda}(x), \quad 0 \le x < \infty. \tag{8.8}$$

Since (8.6) also holds for $\tilde{F}_{0,W_\lambda}$, it follows from (8.7) and (8.8) that

$$\tilde{b}_{v,W_{1\lambda}}(t,\ F_0) \ \ge\ \tilde{b}_{v,W_\lambda}(t,\ F_0).$$

This implies that the theorem holds. □

Acknowledgments. The authors would like to thank Dr.Itsuro Kakiuchi at Kobe University and Dr.Shinichi Matsuda at Nanzan University for their helpful comments. This research was supported in part by the Center for Management Studies at Nanzan University and by Nanzan Pache Research Subsidy I-A,2001. The first step of the research was started while the second author was visiting Probability and Statistics Section, University of Sheffield, U.K..

References

Bednarski,T.(1981). On solutions of minimax test problems for special capacities. *Z. Wakreck. verw. Gebiete.* **10**, 269-278.

Chen,Z. (1998). A note on bias robustness of the median. *Statist. Probab. Lett.* **38**, 363-368.

Donoho,D. and Liu,R. (1988). The "automatic" robustness of minimum distance functionals. *Ann. Statist.* **16**, 552-586.

Grübel,R. (1988). The length of the Shorth. *Ann. Statist.* **16**, 619-628.

Hampel,F.R. Ronchrtti,E.M., Rousseeuw,P.J. and Stahel,W.A.(1986). *Robust Statistics: The approach Based on Influence Functions.* Wiley, New York.

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

Huber,P.J. (1964). Robust estimation of a location parameter. *Ann. Math. Statist.* **35**, 73-101.

Huber,P.J. and Strassen,V.(1973). Minimax tests and the Neyman-Peason lemma for capacities. *Ann. Math. Statist.* **44**, 251-263.

Huber,P.J. (1981). *Robust Statistics.* Wiley, New York.

Martin,R.D. and Zamar,R.H.(1993). Bias robust estimation of scale. *Ann. Statist.* **21**, 991-1017.

Rieder,H. (1977). Least favorable pairs for special capacities. *Ann. Statist..* **6**, 1080-1094.

Rieder,H. (1978). A robust asymptotic testing model. *Ann. Statist.* **6**, 1080-1094.

Rieder,H. (1981a). Robustness of one- and two- sample rank tests against gross errors. *Ann. Statist.* **9**, 245-265.

Rieder,H. (1981b). On local asymptotic minimaxity and admissibility in robust estimation. *Ann. Statist.* **9**, 266-277.

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 Leroy, A.M. (1987). Robust Regression And Outlier Detection. Wiley, New York.

Masakazu Ando
Graduate School of Business Administration, Nanzan University.
18 Yamazato-cho, Showa-ku, Nagoya 466-8673, JAPAN
E-mail: anne@nanzan-u.ac.jp

Miyoshi Kimura
Department of Mathematical Sciences, Nanzan University.
27 Seirei-cho, Seto-shi, Aichi, 489-0863, JAPAN
E-mail: kimura@ms.nanzan-u.ac.jp

Department of Information Systems and Quantitative Sciences,
Nanzan University.
18 Yamazato-cho, Showa-ku, Nagoya 466-8673, JAPAN
E-mail: kimura@nanzan-u.ac.jp