Optimum robust testing in linear models

By Christine Müller

Free University of Berlin
October 1995

Abbreviated title: Optimum robust testing in linear models

Summary

Robust tests for linear models are defined by Wald-type tests based on asymptotically linear estimators. The asymptotic bias of the level in shrinking contamination neighbourhoods is used as robustness criterion which coincides with the robustness criterion derived via influence functions in Heritier and Ronchetti (1994). But in opposite to Heritier and Ronchetti the assumption of Fréchet differentiability of the test statistics is not used. Therefore most robust tests can be derived. By regarding additionally the asymptotic power of the tests also admissible robust tests and most efficient robust tests are derived. Thereby, for most efficiency the determinant of the covariance matrix of the underlying asymptotically linear estimator is minimized. It is shown that at the classical D-optimal designs the most robust tests and the most efficient robust tests have a very simple form. Moreover, the D-optimal designs provide the highest robustness and the highest efficiency under robustness constraints within all designs so that the D-optimal designs are also the optimal designs for robust testing. For two examples the most robust tests and the most efficient robust tests are given explicitly.

AMS 1991 subject classifications. Primary 62F35, 62K05; secondary 62J05, 62J10.
Key words and phrases. Linear model, shrinking contamination, robust tests, asymptotically linear estimators, bias of the level, most robustness, efficiency, D-optimality, optimal design.

1 Introduction

A general linear model

(1.1) $Y_N = A(d_N)\beta + Z_N$

is considered, where $Y_N = (Y_{1N}, ..., Y_{NN})'$ is the vector of real valued observations, $Z_N = (Z_{1N}, ..., Z_{NN})'$ the vector of error variables, $\beta \in I\!R^r$ an unknown parameter vector, $d_N = (t_{1N}, ..., t_{NN})' \in \mathcal{T}^N$ the vector of known experimental conditions, i.e. the design, $A(d_N) = (a(t_{1N}), ..., a(t_{NN}))' \in I\!R^{N\times r}$ the design matrix, and $a : \mathcal{T} \to I\!R^r$ the vector of known "regression" functions. Realizations of the N-dimensional random vectors Y_N and Z_N are denoted by $y_N = (y_{1N}, ..., y_{NN})'$ and $z_N = (z_{1N}, ..., z_{NN})'$, respectively. In this linear model a hypothesis of the general form

$$H_0 : L\beta = l$$

should be tested against the alternative

$$H_1 : L\beta \neq l,$$

where $l \in I\!R^s$ and $\varphi(\beta) = L\beta$ is a linear aspect of the unknown parameter vector β with given matrix $L \in I\!R^{s\times r}$ of rank s.

For outlier robust estimation of β Hampel (1878) and Krasker (1980) proposed to use the gross-error-sensitivity based on the influence function as robustness measure and derived sufficient conditions for most efficient robust M-estimators by using the trace of the asymptotic covariance matrix of the estimators as efficiency measure. Thereby the influence function especially exists if the M-functional is Fréchet differentiable. But Fréchet differentiability implies asymptotic linearity of the estimator which is a common property of estimators. Therefore Bickel (1981, 1984) and Rieder (1985, 1987, 1994) investigated the asymptotic behaviour of asymptotically linear estimators in shrinking contamination neighbourhoods and showed that for appropriate contamination neighbourhoods this leads to the same robustness measure as the approach based on influence functions. They also showed that the sufficient conditions for most efficient robust estimators are also necessary. Kurotschka and Müller (1992) extended these results for estimation of general linear aspects $\varphi(\beta) = L\beta$ of β with $L \in I\!R^{s\times r}$. Müller (1994a) investigated also the influence of the design on the robust estimators. She showed that at the classical A-optimal designs, which minimize the trace of the covariance matrix of the non-robust Gauss-Markov estimators, the most robust estimators and the most efficient robust estimators have the simplest form. She also showed that the A-optimal designs provide most robustness and most efficiency under robustness constraints within all designs. This result bases mainly on complete characterizations of the most robust and the most efficient estimators for all designs. These characterizations are possible because the robustness measure as well as the efficiency measure are convex functions of the score/influence functions of the estimators so that the results of the convex analysis could be applied. But the solutions with respect to this robustness measure and this efficiency measure are not invariant with respect to nonsingular transformations of the linear aspect $\varphi(\beta) = L\beta$. For estimation this is no grave disadvantage. But for tests this invariance property is very important because the optimal tests and designs should not depend on different equivalent formulations of the hypothesis.

For testing a hypothesis of the form $H_0 : \beta = 0$ or $H_0 : \beta_1 = 0$, where β_1 is a subvector of β, there exists several attemps to derive criteria for outlier robustness and robust tests. Many of these robustness criteria were proposed by transferring robustness criteria for estimators to the test statistics, and as for estimation usually the asymptotic behaviour of the test statistics were regarded. See for example Ronchetti (1982), Hampel et al. (1986), Staudte and Sheather (1990), Markatou and Hettmansperger (1990), Akritas (1991), Markatou et al. (1991), Silvapulle (1992a,b), Coakley and Hettmansperger (1992, 1994), Markatou and He (1994), Heritier and Ronchetti (1994). In particular Heritier and Ronchetti (1994) derived a robustness concept based on the influence function of the test statistic by assuming Fréchet differentiability of the test statistic. To avoid Fréchet differentiability and to take a planned experiment into account we here use a robustness concept which Müller (1992a,c) proposed. In this robustness concept the test statistics are Wald-type test statistics based on asymptotically linear estimators and their behaviour in shrinking contamination neighbourhoods is investigated to derive the robustness property of these tests. As for estimation it turns out that this approach leads to the same robustness measure as the approach based on influence functions of Heritier and Ronchetti (1994). But in opposite to the approach of Heritier and Ronchetti we show that our approach also includes the most robust test statistics which are not Fréchet differentiable. This is important because in this paper we will derive most robust tests and most efficient robust tests as it was possible for estimation.

Up to now only Ronchetti (1982) derived optimal robust tests within the class of τ-tests. These τ-tests have a complicated asymptotic distribution so that also the optimal robust tests are of complicated form. See also Hampel et al. (1986), Chapter 7. In our approach the asymptotic distribution of the Wald-type tests is the simple chi-squared distribution. But in opposite to estimation we have for the Wald-type tests the problem that the robustness measure is not anymore a convex function of the score/influence functions. It is invariant with respect to nonsingular transformation of the linear aspect as it should be for testing. As efficiency measure we should use also a measure which provides invariant solutions with respect to the nonsingular transformations. An appropiate measure satisfying this property and describing the asymptotic power of the test is the determinant of the asymptotic covariance matrix of the underlying estimator of the Wald-type test. But this efficiency measure is also not a convex function of the score/influence functions. Hence, the results of the convex analysis can not be applied as for estimation. Therefore most robust tests and most efficient robust tests can be characterized only for some special designs. But nevertheless it is possible to show that at the classical D-optimal designs most robustness and most efficiency under robustness constraints is achieved within all designs. Moreover we will see that at the D-optimal designs the most robust tests and the most efficient robust tests have a very simple form. Hence, the D-optimal designs play for robust testing the same role as the A-optimal designs for robust estimation.

Recall that the classical D-optimal design is that design at which the determinant of the covariance matrix of the nonrobust Gauss-Markov estimator and the volume of the ellipsoid given by a fixed power value of the nonrobust F-test are minimized. If we generalize the designs d_N to design measures δ (briefly also called designs) then this means that a design δ_D is D-optimal for $L\beta$ in a set Δ of designs if and only if

$$\delta_D \in \arg\min\{\det(L\, I(\delta)^- L');\ \delta \in \Delta\},$$

where $\det(M)$ is the determinant of a matrix M, M^- is the g-inverse of a matrix M, i.e. $M\,M^-\,M = M$, and $I(\delta) := \int a(t)\,a(t)'\,\delta(dt)$ is the information matrix of the design δ. (see for example Pázman (1986), p. 100, or Pukelsheim (1993), p. 135).

The paper is organized as follows: In Section 2 the ALE-tests, the Wald-type tests based on asymptotically linear estimators, are defined and their asymptotic behaviour in linear models with shrinking contamination is derived. The asymptotic behaviour provides in particular the asymptotic bias of the level and the asymptotic power of the Wald-type tests which are used in Section 3 for deriving the robustness measure and the efficiency measure, respectively. In Section 4 the most robust tests and designs, i.e. the tests and the designs with minimum asymptotic bias of the level, are derived. In Section 5 the most efficient robust tests and designs, i.e. the tests and designs which maximize the asymptotic power under a bias bound b, are derived. Most of the results of this section reduce the power (the efficiency) of the test to the determinant of the asymptotic covariance matrix of the underlying estimator. But also admissible tests are characterized by generalizing a result of Krasker and Welsch (1982). In Section 6 two examples are given, which show the applicability of the results for models with qualitative factors as well as for models with quantitative factors. In the first example the most robust test and the most efficient robust test are given for the problem of testing the equality of the effects in a one-way lay-out model. In the second example the optimal robust tests are presented for a test problem in a quadratic regression model. In Section 7 all proofs are given.

2 ALE-tests and their asymptotic behaviour in contaminated linear models

To study the asymptotic behaviour of tests in the linear model (1.1) we assume that the sequence of designs $(d_N)_{N\in\mathbb{N}}$ is converging to an asymptotic design measure δ with finite support in the following sense:

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} e_{t_{nN}}(\{t\}) = \delta(\{t\})$$

for all $t \in \text{supp}(\delta)$, where $\text{supp}(\delta)$ is the support of δ and e_t is the Dirac measure on $t \in \mathcal{T}$. Briefly δ is also called a design. The results of this paper will also hold if the experimental conditions $t_{1N}, ..., t_{NN}$ are realizations of independent and identically distributed random variables $T_{1N}, ..., T_{NN}$, each with distribution δ, so that d_N is the realization of a random design (see Kurotschka and Müller (1992) and Müller (1994b)). But for brevity we here regard only deterministic design sequences $(d_N)_{N\in\mathbb{N}}$.

In the ideal (central) linear model (1.1) it is assumed that the error variables $Z_{1N}, ...,$ Z_{NN} have unknown variance σ^2 so that $Z_{1N}/\sigma, ..., Z_{NN}/\sigma$ are independent and identically distributed according to a symmetric distribution $P = P^{Z_{nN}/\sigma}$ with mean 0, variance 1 and bounded Lebesgue density $f(z)$. For the contaminated linear model (1.1), which takes outliers into account, we assume as in Müller (1994b) that $Z_{1N}/\sigma, ..., Z_{NN}/\sigma$ are independent but distributed according to a contaminated distribution of the form

$$\begin{aligned} Q_{nN}(dz) &:= P^{Z_{nN}/\sigma}(dz) \\ &= (1 - N^{-1/2}\,\epsilon\, c(t_{nN}))P(dz) + N^{-1/2}\,\epsilon\, c(t_{nN})\, g(z, t_{nN})\, P(dz), \end{aligned}$$

where $\epsilon > 0$, $c(t) \geq 0$, $g(z,t) \geq 0$, $\int c(t)\,\delta(dt) \leq 1$ and $\int g(z,t)\,P(dz) = 1$ for all $t \in \mathcal{T}$ and $z \in \mathbb{R}$. All sequences $(Q^N_{\epsilon,c,g})_{N\in\mathbb{N}} = (Q^N)_{N\in\mathbb{N}} := (\bigotimes_{n=1}^N Q_{nN})_{N\in\mathbb{N}}$ define the shrinking contamination neighbourhood $\mathcal{U}_\epsilon$ around the ideal (central) model $(P^N)_{N\in\mathbb{N}} := (\bigotimes_{n=1}^N P)_{N\in\mathbb{N}}$. If P^N is the distribution of Z_N then P^N_β denotes the distribution of Y_N, and if Q^N is the distribution of Z_N then Q^N_β denotes the distribution of Y_N.

For testing $H_0 : L\beta = l$ in such contaminated linear model appropriate tests can be based on estimators for $\varphi(\beta) = L\beta$, for example on asymptotically linear estimators $\widehat{\varphi}_N : \mathbb{R}^N \times \mathcal{T}^N \to \mathbb{R}^s$ for $\varphi(\beta)$. The class of asymptotically linear estimators has the advantage that it includes many well-known estimators as R-estimators, M-estimators and one-step M-estimators, and the asymptotically linear estimators are convenient for studying the asymptotic behaviour of the estimators in the shrinking contamination neighbourhood as Bickel (1984), Rieder (1985, 1987, 1994) and Kurotschka and Müller (1992) have shown. While for estimation the asymptotic linearity of the estimator should hold under all $(P^N_\beta)_{N\in\mathbb{N}}$ with $\beta \in \mathbb{R}^r$, for deriving robust tests the asymptotic linearity should be specified in the sense that the estimators should be asymptotically linear under all $(P^N_{\beta_N})_{N\in\mathbb{N}}$ with $\beta_N = \beta + N^{-1/2}\overline{\beta}$ and $L\beta = l$. Usually $(P^N_{\beta_N})_{N\in\mathbb{N}}$ provides contiguous alternatives to the distributions $(P^N_\beta)_{N\in\mathbb{N}}$ of the hypothesis $H_0 : L\beta = l$. Hence, we define the asymptotic linearity here as follows.

Definition 2.1 *An estimator $\widehat{\varphi}_N$ for $\varphi(\beta) = L\beta$ is asymptotically linear for contiguous alternatives of $H_0 : L\beta = l$ with influence function ψ if $\psi : I\!R \times \mathcal{T} \to I\!R^s$ satisfies the conditions*

$$(2.1) \qquad \int \psi(z,t)\, P(dz) = 0 \quad \textit{for all } t \in \mathcal{T},$$

$$(2.2) \qquad \int \psi(z,t)\, a(t)'z\, P(dz)\, \delta(dt) = L,$$

$$(2.3) \qquad \int \psi(z,t)^2\, P(dz)\, \delta(dt) < \infty,$$

and

$$(2.4) \qquad \lim_{N\to\infty} P^N_{\beta_N} \Big(\Big\{ y_N \in I\!R^N;\ \sqrt{N} | \widehat{\varphi}(y_N, d_N) - \varphi(\beta_N) - \frac{\sigma}{N} \sum\nolimits_{n=1}^N \psi\left(\frac{y_{nN} - a(t_{nN})'\beta_N}{\sigma}, t_{nN} \right) \Big| > \epsilon \Big\} \Big) = 0$$

for all $\epsilon > 0$, $\sigma \in I\!R^+$ and $\beta_N = \beta + N^{-1/2}\overline{\beta}$ with $\beta, \overline{\beta} \in I\!R^r$ and $L\beta = l$.

In Müller (1994b) conditions were given so that in a designed experiment a one-step M-estimator is asymptotically linear for all $\beta \in I\!R^r$. But under these conditions the asymptotic linearity for all $\beta \in I\!R^r$ implies property (2.4) if $P^N_{\beta_N}$ is contiguous to P^N_β, which is the case for example for the normal distribution $P = n_{(0,1)}$. Hence, for the rest of the paper we will assume that $P^N_{\beta_N}$ is contiguous to P^N_β. In particular (2.4) is satisfied for score functions ψ satisfying condition (2.2), (2.3) and

$$(2.5) \qquad \begin{array}{l} \psi(z,t) = \psi_0(z,t) + h(t)\mathrm{sgn}(z), \text{ where for all } t \in T_0 \\ \psi_0(\cdot,t) \text{ is antisymmetric, continuous and there exists } \lambda_1(t), ..., \lambda_\Lambda(t) \text{ so that ,} \\ \psi_0(\cdot,t) \text{ has bounded and continuous derivatives on } I\!R \setminus \{\lambda_1(t), ..., \lambda_\Lambda(t)\}. \end{array}$$

Note that score functions ψ satisfying (2.2), (2.3) and (2.5) can be discontinuous so that the corresponding M-functionals are not Fréchet differentiable.

Replacing the Gauss-Markov estimator $\widehat{\varphi}^{LS}_N$ in the test statistic of the classical F-test for $H_0 : L\,\beta = l$ by an asymptotically linear estimator for contiguous alternatives we can generalize the classical F-test. Recall that the test statistic of the F-test is

$$\tau^{LS}_N(y_N, d_N) = N(\widehat{\varphi}^{LS}_N(y_N, d_N) - l)'\, C^{LS}_N(y_N, d_N)^{-1} (\widehat{\varphi}^{LS}_N(y_N, d_N) - l),$$

where $C^{LS}_N(y_N, d_N)$ is converging in probability to the asymptotic covariance matrix of the Gauss-Markov estimator (see for example Christensen (1987), p. 40 f.). Hence, in the generalization of the F-test statistic we also substitute C^{LS}_N by any function $C_N : I\!R^N \times \mathcal{T}^N \to I\!R^{s\times s}$ so that $C_N(y_N, d_N)$ is converging in probability to the asymptotic covariance matrix $\sigma^2\, C(\psi, \delta)$ of the asymptotically linear estimator, where

$$C(\psi, \delta) := \int \psi(z,t)\, \psi(z,t)'\, P(dz)\, \delta(dt).$$

Then, the generalized F-test statistic for testing $H_0 : L\beta = l$ is of Wald-type (compare with Wald (1943), Markatou et al. (1991), Silvapulle (1992a), Heritier and Ronchetti (1994)). Here such tests will be called *ALE-tests* and their corresponding test statistics are called *ALE-test statistics* because they are based on asymptotically (A) linear (L) estimators (E). Those tests are also closely connected with the score-type tests (see Markatou et al. (1991), Markatou and He (1994), Heritier and Ronchetti (1994)).

Definition 2.2 *A test statistic $\tau_N : I\!R^N \times \mathcal{T}^N \to I\!R$ is an ALE-test statistic with influence function ψ for testing $H_0 : L\beta = l$ (briefly for testing $L\beta$) if*

$$\tau_N(y_N, d_N) = N(\widehat{\varphi}_N(y_N, d_N) - l)' \, C_N(y_N, d_N)^{-1} \, (\widehat{\varphi}_N(y_N, d_N) - l),$$

where $\widehat{\varphi}_N$ is asymptotically linear for contiguous alternatives of $H_0 : L\beta = l$ with influence function ψ and $\lim_{N\to\infty} C_N = \sigma^2 \, C(\psi, \delta)$ in probability for $(P_\beta^N)_{N\in I\!N}$ with $L\beta = l$. A test based on an ALE-test statistic is called an ALE-test.

An ALE-test statistic is asymptotically chi-squared distributed in a shrinking contamination neighbourhood. Thereby we can regard not only shrinking neighbourhoods around $(P_\beta^N)_{N\in I\!N}$ with $L\beta = l$ but also around $(P_{\beta_N}^N)_{N\in I\!N}$ with $L\beta_N = l + N^{-1/2}\gamma$.

Theorem 2.1 *If τ_N is an ALE-test statistic for testing $H_0 : L\beta = l$ with influence function ψ then τ_N has an asymptotic chi-squared distribution, i.e.*

$$\mathcal{L}(\tau_N | Q_{\beta_N}^N) \stackrel{N\to\infty}{\longrightarrow} \chi^2(s, [\gamma + \sigma\, b(\psi, (Q^N)_{N\in I\!N})]' \, [\sigma^2 \, C(\psi, \delta)]^{-1} [\gamma + \sigma\, b(\psi, (Q^N)_{N\in I\!N})])$$

for all $(Q^N)_{N\in I\!N} = (Q_{\epsilon,c,g}^N)_{N\in I\!N} \in \mathcal{U}_\epsilon$ and all $\beta_N = \beta + N^{-1/2}\overline{\beta} \in I\!R^r$ with $L\beta_N = l + N^{-1/2}\gamma$, where

$$b(\psi, (Q_{\epsilon,c,g}^N)_{N\in I\!N}) := \epsilon \int \psi(z,t)\, c(t)\, g(z,t)\, P(dz)\, \delta(dt)$$

and $\chi^2(s, \lambda)$ is the chi-squared distribution with s degrees of freedom and noncentrality parameter λ.

3 Efficiency and robustness of ALE-tests

Because $b(\psi, (Q^N)_{N\in I\!N}) = 0$ if $(Q^N)_{N\in I\!N} = (P^N)_{N\in I\!N}$ Theorem 2.1 provides the asymptotic level and the asymptotic power for ideal distributions P^N. In particular the rejection set should be chosen as $(\chi^2_{1-\alpha,s,0}, \infty)$ to get an asymptotic α level test for the ideal distributions. Thereby $\chi^2_{1-\alpha,s,0}$ is the $(1-\alpha)$-quantile of the chi-squared distribution $\chi^2(s,0)$. Then the asymptotic power for contiguous ideal alternatives $(P^N_{\beta_N})_{N\in I\!N}$ with $L\beta_N = l + N^{-1/2}\gamma \neq l$, is

$$\text{(3.1)} \qquad \lim_{N\to\infty} P^N_{\beta_N}(\tau_N > \chi^2_{1-\alpha,s,0}) = 1 - \mathcal{X}^2_{s,\gamma'[\sigma^2\,C(\psi,\delta)]^{-1}\gamma}(\chi^2_{1-\alpha,s,0}),$$

where $\mathcal{X}_{s,\lambda}$ is the distribution function of the chi-squared distribution $\chi^2(s,\lambda)$. For the efficiency of an ALE-test we will regard this asymptotic power (3.1). This power is maximized if $\gamma'[\sigma^2\,C(\psi,\delta)]^{-1}\gamma$ is maximized for all $\gamma \in I\!R^s$.

For the robustness of an ALE-test we regard its maximum bias of the asymptotic level for contaminated distributions Q^N_β with $L\,\beta = l$, i.e.

$$\begin{aligned} b_\epsilon(\psi) \ := \ \max\{ & \lim_{N\to\infty}(Q^N_\beta(\tau_N > \chi^2_{1-\alpha,s,0}) - \alpha);\ (Q^N)_{N\in I\!N} \in \mathcal{U}^0_\epsilon \\ & \text{and } \beta \in I\!R^r \text{ with } L\beta = l\}. \end{aligned}$$

The bias $b_\epsilon(\psi)$ is bounded if and only if $\psi'\,C(\psi,\delta)^{-1}\psi$ is bounded. This shows the following theorem.

Theorem 3.1 *For an ALE-test for testing $H_0 : L\beta = l$ with influence function ψ the following inequalities are equivalent:*

$$\begin{aligned} &a)\quad b_\epsilon(\psi) \leq 1 - \mathcal{X}^2_{s,b}(\chi^2_{1-\alpha,s,0}) - \alpha, \\ &b)\quad \epsilon^2\,\|\psi'\,C(\psi,\delta)^{-1}\psi\|_\delta \leq b, \end{aligned}$$

where $\|\psi'\,C(\psi,\delta)^{-1}\psi\|_\delta := \max_{(z,t)\in I\!R\times \text{supp}(\delta)} \psi(z,t)'\,C(\psi,\delta)^{-1}\,\psi(z,t)$.

Theorem 3.1 shows that the asymptotic bias of a test based on a M-estimator is bounded if the self-standardized gross-error-sensitivity of the M-estimator is bounded (for the definition of the self-standardized gross-error-sensitivity see Krasker and Welsch (1982), Ronchetti and Rousseeuw (1985), Hampel et al. (1986)). This result corresponds with a result of Heritier and Ronchetti (1994) who derived the robustness of tests via influence functions by assuming Fréchet differentiability of the test statistic.

Because the maximum asymptotic bias of the level and the asymptotic power depend on the ALE-test only via its influence function we will have in Section 3 and 4 optimization problems within the set of influence functions. Because usually the influence functions are of the form $\psi(z,t) = M\,a(t)\,\rho(z,t)$ (see for example Hampel et al. (1986)) we will regard optimization problems within

$$\begin{aligned} \Psi^*(\delta, L) \ := \ \{ & \psi : I\!R\times\mathcal{T} \to I\!R^s;\ \psi \text{ satisfies conditions (2.1) - (2.3) and} \\ & \psi(z,t) = M\,a(t)\,\rho(z,t) \text{ for all } (z,t) \in I\!R\times\mathcal{T} \text{ for some} \\ & M \in I\!R^{s\times r} \text{ and } \rho : I\!R\times\mathcal{T} \to I\!R\}. \end{aligned}$$

Thereby, it is not clear if every ψ lying in $\Psi^*(\delta, L)$ is the influence function of an ALE-test. But if we find an optimal solution ψ_* within $\Psi^*(\delta, L)$ which satisfies the conditions (2.2),(2.3) and (2.5) then we know from the remarks above that a one-step M-estimator with this score function ψ_* is an AL-estimator with influence function ψ_* so that ψ_* is also an influence function of an ALE-test. We will see that this is always the case.

For finding the optimal design we will regard the set of all designs, so that the linear aspect $\varphi(\beta) = L\,\beta$ is identifiable, i.e.

$$\Delta(L) := \{\delta;\ \text{there exists } K \in I\!R^{s\times r} \text{ with } L = K\, I(\delta)\}.$$

Sometimes we will regard also the set of designs with fixed support $\mathcal{D} \subset \mathcal{T}$, i.e.

$$\Delta_{\mathcal{D}} := \{\delta;\ \text{supp}(\delta) = \mathcal{D}\}.$$

4 Most robust tests and designs

In Theorem 3.1 it was shown that the asymptotic bias of the level of an α level ALE-test for testing $H_0 : L\beta = l$ is a strictly increasing function of $\|\psi'\, C(\psi, \delta)^{-1}\, \psi\|_\delta$, so that most robust tests and designs should minimize $\|\psi'\, C(\psi, \delta)^{-1}\, \psi\|_\delta$. Thereby let be

$$b_0^T(\delta, L) := \min\{\|\psi'\, C(\psi, \delta)^{-1}\, \psi\|_\delta;\ \psi \in \Psi^*(\delta, L)\}$$

the minimum asymptotic bias for testing $L\beta$ at δ.

Definition 4.1

a) An ALE-test with influence function ψ_0 is most robust for testing $L\beta$ at δ if

$$\psi_0 \in \arg\min\{\|\psi'\, C(\psi, \delta)^{-1}\, \psi\|_\delta;\ \psi \in \Psi^*(\delta, L)\}.$$

b) A design δ_0 is most robust for testing $L\beta$ in Δ if

$$\delta_0 \in \arg\min\{b_0^T(\delta, L);\ \delta \in \Delta\}.$$

Note that the optimization criterion $\|\psi'\, C(\psi, \delta)^{-1}\, \psi\|_\delta$ is not convex in ψ, but it is invariant with respect to nonsingular transformations of $L\,\beta$.

Theorem 4.1
a) For all $\delta \in \Delta(L)$ *we have* $b_0^T(\delta, L) \geq s$.
b) If δ_D *is D-optimal for* $L\beta$ *in* $\Delta_{\mathcal{D}} \subset \Delta(L)$ *then* $b_0^T(\delta_D, L) = s$, *i.e.* δ_D *is most robust for testing* $L\beta$ *in* $\Delta(L)$, *and* ψ_{01} *given by* $\psi_{01}(z,t) := L\, I(\delta_D)^- \, a(t)\, sgn(z) \sqrt{\frac{\pi}{2}}$ *is the influence function of a most robust ALE-test for testing* $L\beta$ *at* δ_D.

Note that ψ_{01} is the influence function of the L_1 estimator and that the M-functional of this estimator is not Fréchet differentiable. Hence ψ_{01} can not be regarded within the approach of Heritier and Ronchetti (1994).

Theorem 4.1 holds also for designs which are D-optimal in a modified model. For that define for a function $h : \mathcal{T} \to I\!R^+ \setminus \{0\}$

$$a_h(t) := h(t)^{-1} a(t), \; L_h := L\, J_h^- \, I_h(\delta),$$

where

$$J_h := \int a_h(t)\, a(t)'\, \delta(dt), \; I_h(\delta) := \int a_h(t) a_h(t)' \delta(dt).$$

Corollary 4.2 *If there exists some function* $h : \mathcal{T} \to I\!R^+ \setminus \{0\}$ *so that* δ *is D-optimal in* $\Delta_{\mathrm{supp}(\delta)} \subset \Delta(L)$ *for* $\varphi_h(\beta) = L_h\beta$ *in the modified model* $Y_h(t) = a_h(t)'\beta + Z$, *then* $b_0^T(\delta, L) = s$, *i.e.* δ *is most robust for testing* $L\beta$ *in* $\Delta(L)$, *and* ψ_0 *given by* $\psi_0(z,t) = L_h I_h(\delta)^- a_h(t) sgn(z) \sqrt{\frac{\pi}{2}}$ *is the influence function of a most robust ALE-test for testing* $L\beta$ *at* δ.

Note that Theorem 4.1 and Corollary 4.2 are analogous to Lemma 1(a) in Müller (1994a) and Theorem 2 in Müller (1992b), respectively, which show that for estimation at A-optimal designs the minimum bias is attained. Using $h(t) = |a(t)|$ Theorem 4.1 a) and Corollary 4.2 provide in particular Theorem 2' of Ronchetti and Rousseeuw (1985) (or Proposition 1(ii) in Hampel et al. (1986), p. 318) which concerns estimators with minimum self-standardized gross-error-sensitivity.

5 Most efficient robust tests and designs

A good test should not only have a small bias of the level but also a high efficiency, i.e. a high power, and the power should be high at least at the ideal model. Because small bias and high power are contrary properties we use the usual approach of maximizing the power at the ideal model under the side condition that the bias is bounded by some bias bound. Because the asymptotic power at the ideal model and the maximum asymptotic bias of an ALE-test are given by $\gamma' C(\psi,\delta)^{-1}\gamma$ with $\gamma \in I\!R^s$ and $\|\psi' C(\psi,\delta)^{-1}\psi\|_\delta$, respectively, we will maximize $\gamma' C(\psi,\delta)^{-1}\gamma$ for all $\gamma \in I\!R^s$ under the side condition that $\|\psi' C(\psi,\delta)^{-1}\psi\|_\delta \le b$. If we have no bias bound, i.e. $b = \infty$ then the influence function of the Gauss-Markov estimator, i.e. ψ_{GM} given by $\psi_{GM}(z,t) := L\,I(\delta)^- a(t)\,z$, maximizes $\gamma' C(\psi,\delta)^{-1}\gamma$ for all $\gamma \in I\!R^s$ within $\Psi^*(\delta,L)$. But for $b < \infty$ in general it is not possible to find a influence function which maximizes $\gamma' C(\psi,\delta)^{-1}\gamma$ uniformly for all $\gamma \in I\!R^s$, i.e. which minimizes $C(\psi,\delta)$ in the positive-semidefinite sense. We can characterize admissible solutions, but to find optimal solutions we have to regard a real-valued function of $C(\psi,\delta)$. For this real-valued function we will use the determinant because the minimization of $\det(C(\psi,\delta))$ is invariant with respect to nonsingular transformations of the linear aspect $\varphi(\beta) = L\beta$. Then $\det(C(\psi,\delta))$ is not convex in ψ. But also the side condition is not convex in ψ so that in no case the optimizations results of the convex analysis are applicable. As in classical design theory the solutions which minimize $\det(C(\psi,\delta))$ are called "D-optimal".

Definition 5.1

a) An ALE-test with influence function $\psi_{b,\delta}$ is admissible for testing $L\beta$ at δ with bias bound b if $\|\psi_{b,\delta}' C(\psi_{b,\delta},\delta)^{-1}\psi_{b,\delta}\|_\delta \le b$ and $C(\psi,\delta) = C(\psi_{b,\delta},\delta)$ for all $\psi \in \Psi^*(\delta,L)$ with $\|\psi'\,C(\psi,\delta)^{-1}\,\psi\|_\delta \le b$ and $C(\psi,\delta) \le C(\psi_{b,\delta},\delta)$.

b) An ALE-test with influence function $\psi_{b,\delta}$ is D-optimal for testing $L\beta$ at δ with bias bound b if

$$\psi_{b,\delta} \in \arg\min\Big\{\det(C(\psi,\delta));\ \psi \in \Psi^*(\delta,L)\ \textit{with}\ \|\psi'\,C(\psi,\delta)^{-1}\,\psi\|_\delta \le b\Big\}.$$

c) A design δ_b is D-optimal in Δ for testing $L\beta$ with bias bound b if

$$\delta_b \in \arg\min\ \Big\{\min\Big\{\det(C(\psi,\delta));\ \psi \in \Psi^*(\delta,L)\ \textit{with}\ \|\psi'\,C(\psi,\delta)^{-1}\,\psi\|_\delta \le b\Big\};\delta \in \Delta\Big\}.$$

For deriving admissible and D-optimal tests we will assume that the ideal distribution P is the standard normal distribution, i.e. $P = n_{(0,1)}$. The distribution function of $n_{(0,1)}$ will be denoted by Φ.

At first we characterize the influence function of an admissible ALE-test by generalizing a result of Krasker and Welsch (1982) concerning admissible estimators for estimating β with a self-standardized gross-error-sensitivity bounded by b. For that we use the following matrices:

$$M_b(B) := \int a(t)a(t)'\left[2\Phi\left(\frac{\sqrt{b}}{|Ba(t)|}\right) - 1\right]\ \delta(dt),$$

$$Q_b(B) := \int a(t)a(t)' g\left(\frac{\sqrt{b}}{|Ba(t)|}\right) \delta(dt),$$

$$M_{b,0}(B) := \int a(t)a(t)' \frac{\sqrt{b}}{|Ba(t)|}\sqrt{\frac{2}{\pi}}\, \delta(dt),$$

$$Q_{b,0}(B) := \int a(t)a(t)' \left(\frac{\sqrt{b}}{|Ba(t)|}\right)^2 \delta(dt).$$

Theorem 5.1 *Let $\delta \in \Delta(L)$ and*

$$\psi_{b,\delta}(z,t) \;:=\; L\, M_b(B_b)^- \,a(t)\, \min\left\{|z|, \frac{\sqrt{b}}{|B_b a(t)|}\right\} \operatorname{sgn}(z), \tag{5.1}$$

where

$$a(t)'B_b' B_b\, a(t) = a(t)' M_b(B_b)^-\, L'\, [L\, M_b(B_b)^- Q_b(B_b) M_b(B_b)^- L']^{-1}\, L\, M_b(B_b)^-\, a(t) \quad \text{for all } t \in supp(\delta), \tag{5.2}$$

or

$$\psi_{b,\delta}(z,t) \;:=\; L\, M_{b,0}(B_b)^-\, a(t)\, \frac{\sqrt{b}}{|B_b a(t)|} \operatorname{sgn}(z), \tag{5.3}$$

where

$$\begin{aligned} &a(t)'B_b' B_b\, a(t) \\ &\quad = \; a(t)' M_{b,0}(B_b)^-\, L'\, [L\, M_{b,0}(B_b)^- Q_{b,0}(B_b) M_{b,0}(B_b)^- L']^{-1}\, L\, M_{b,0}(B_b)^-\, a(t) \\ &\quad\quad \text{for all } t \in supp(\delta). \end{aligned} \tag{5.4}$$

Then the ALE-test with influence function $\psi_{b,\delta}$ is admissible for testing $L\beta$ at δ with bias bound b.

Theorem 5.1 can be used to show that at a D-optimal design δ_D the influence functions of admissible ALE-tests for testing $L\beta$ with bias bound b have a very simple form. These influence functions are given by

$$\psi_{b,\delta_D}(z,t) = \begin{cases} L\, I(\delta_D)^-\, a(t) \operatorname{sgn}(z) \sqrt{\frac{\pi}{2}}, & \text{for } b = s, \\ L\, I(\delta_D)^-\, a(t) \operatorname{sgn}(z) \frac{\min\{|z|, \sqrt{b}\, y_b\}}{2\Phi(\sqrt{b}\, y_b) - 1}, & \text{for } b > s, \end{cases} \tag{5.5}$$

where

$$y_b^2 = \frac{1}{s}\, g\left(\sqrt{b}\, y_b\right) > 0 \tag{5.6}$$

with

$$g(y) := \int \min\{|z|, y\}^2\, P(dz).$$

This shows the following lemma. This lemma is analogous to Lemma 1(b) in Müller (1994a), which shows that for estimation with bias bound at A-optimal designs the influence functions of A-optimal AL-estimators have a simple form. Thereby instead of the trace of the asymptotic covariance matrix we have a simple formula for the asymptotic covariance matrix of the underlying AL-estimator. For that define $\overline{W} : (0,\infty)\times[0,\infty) \to I\!R$ by $\overline{W}(c,y) = g(y)/c - y^2$ and $\overline{w} : (0,1) \to [0,\infty)$ implicitely by $\overline{W}(c,\overline{w}(c)) = 0$. Note that y_b of (5.6) satisfies $y_b = (1/\sqrt{b})\,\overline{w}\,(s/b)$. Set also

$$\overline{v}(c) \;=\; \begin{cases} \frac{(2\Phi(\overline{w}(c))-1)^2}{g(\overline{w}(c))} & \text{for } c < 1, \\ \frac{2}{\pi} & \text{for } c = 1. \end{cases}$$

Note also that a solution of (5.6) always exists and can be easily calculated. For, the function $\overline{\omega} : [0,\infty) \to [0,\infty)$ given by $\overline{\omega}(y) = g(y)/c - y^2$ for $c > 0$ has at most one positive root and this root exists if and only if $c < 1$. Thereby the root lies in $S := \{y \in [0,\infty);\; 2\Phi(-y) \le c\}$, where $\overline{\omega}$ is concave on S. Hence, the root can be calculated by Newton's method by starting with an interior point of S.

Lemma 5.2 *If δ_D is D-optimal for $L\beta$ in $\Delta_{\mathcal{D}} \subset \Delta(L)$ and $b \ge s$ then the ALE-test with influence function ψ_{b,δ_D} given by (5.5) is admissible for testing $L\beta$ with bias bound b at δ_D and its asymptotic covariance matrix satisfies*

$$C(\psi_{b,\delta_D},\delta_D) = \frac{1}{\overline{v}\left(\frac{s}{b}\right)}\, L\, I(\delta_D)^-\, L'.$$

If $a(\tau_1),...,a(\tau_I)$ are linearly independent and δ is not D-optimal in $\Delta_{\mathrm{supp}(\delta)}$, then we have a modification of Lemma 5.2 which gives at least a lower bound for the asymptotic covariance matrix. For ψ with $\psi(z,t) = M\,a(t)\,\rho(z,t)$ set

$$v_\rho(t) \;:=\; \frac{(\int \rho(z,t)\,z\,P(dz))^2}{\int \rho(z,t)^2 P(dz)}$$

and

$$\delta_\rho(\{t\}) \;:=\; \frac{1}{\sum_{\tau\in\mathcal{D}} v_\rho(\tau)\,\delta(\{\tau\})}\, v_\rho(t)\,\delta(\{t\}).$$

Lemma 5.3 *If $\delta \in \Delta_{\mathcal{D}} \subset \Delta(L)$, $\mathcal{D} = \{\tau_1,...,\tau_I\}$, $a(\tau_1),...,a(\tau_I)$ are linearly independent and $\psi \in \Psi^*(\delta,L)$ with $\|\psi' C(\psi,\delta)^{-1}\psi\|_\delta \le b$ and $\psi(z,t) = M\,a(t)\,\rho(z,t)$, then*

$$C(\psi,\delta) \ge \frac{1}{\overline{v}\left(\frac{s}{b}\right)}\, L\, I(\delta_\rho)^-\, L'.$$

With Lemma 5.2 and Lemma 5.3 we can show that at a D-optimal design δ_D we have D-optimality of an ALE-test with influence function ψ_{b,δ_D} given by (5.5). Moreover, the D-optimal design is also D-optimal for testing with bias bound b in $\Delta_{\mathcal{D}}$ and this holds not only within $\Delta_{\mathcal{D}}$ with linearly independent regressors but also for any D-optimal design within designs of $\Delta(L)$ with finite support. Therefore we have a theorem which is analogous to the theorem in Müller (1994a), which shows that the A-optimal designs are also optimal for robust estimation.

Theorem 5.4 *If $\Delta_0 = \{\delta \in \Delta(L);\ supp(\delta)$ is finite $\}$, $b \geq s$ and δ_D is D-optimal for $L\beta$ in Δ_0, then δ_D is D-optimal in Δ_0 for testing $L\beta$ with bias bound b and an ALE-test with influence function ψ_{b,δ_D} given by (5.5) is asymptotically D-optimal for testing $L\beta$ at δ_D with bias bound b.*

Theorem 5.4 in particular shows that the D-optimal design for testing with a bias bound b does not depend on the bias bound b.

6 Examples

Example 6.1 (One-way lay-out)
In a one-way lay-out model with four levels the observation at $t_{nN} = i$ is given by

$$Y_{nN} = \beta_i + Z_{nN} = a(t_{nN})'\beta + Z_{nN}$$

for $n = 1, ..., N$, where $\beta = (\beta_1, \beta_2, \beta_3, \beta_4)' \in I\!R^4$, $a(t) = (1_1(t), 1_2(t), 1_3(t), 1_4(t))' \in I\!R^4$ and $\mathcal{T} = \{1, 2, 3, 4\}$. Usually in this model the hypothesis $H_0 : L\beta := (\beta_2 - \beta_1, \beta_3 - \beta_1, \beta_4 - \beta_1)' = 0$, or equivalently $H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4$, against $H_1 : L\beta = (\beta_2 - \beta_1, \beta_3 - \beta_1, \beta_4 - \beta_1)' \neq 0$ should be tested. The D-optimal design for $L\beta$ in $\Delta(L)$ is

$$\delta_D = \frac{1}{4}(e_1 + e_2 + e_3 + e_4).$$

According to Theorem 4.1 this design δ_D is most robust for testing $H_0 : L\beta = 0$, i.e.

$$b_0^T(\delta_D, L) = 3 = \min\{b_0^T(\delta, L);\ \delta \in \Delta(L)\},$$

and the influence function of a most robust ALE-test at δ_D has the form

$$(6.1) \quad \psi_{01}(z, t) = \begin{cases} (-1, -1, -1)' \operatorname{sgn}(z)\, 4\sqrt{\frac{\pi}{2}} & \text{for } t = 1, \\ (1_2(t), 1_3(t), 1_4(t))' \operatorname{sgn}(z)\, 4\sqrt{\frac{\pi}{2}} & \text{for } t \neq 1. \end{cases}$$

According to Theorem 5.4 for every $b \geq b_0^T(\delta_D, L) = 3$ the D-optimal design δ_D is also D-optimal for testing with bias bound b and the influence function of a D-optimal ALE-test for testing at δ_D with bias bound b is given by (5.5). For $b = b_0^T(\delta_D, L) = 3$ it coincides with ψ_{01} given by (6.1) and for $b > b_0^T(\delta_D, L) = 3$ it has the form

$$\psi_{b,\delta_D}(z,t) = \begin{cases} (-1,-1,-1)' \operatorname{sgn}(z)\, 4\, \frac{\min\{|z|, \sqrt{b}\, v_b\}}{2\Phi(\sqrt{b}\, v_b)-1}, & \text{for } t = 1, \\ (1_{\{2\}}(t), 1_{\{3\}}(t), 1_{\{4\}}(t))' \operatorname{sgn}(z)\, 4\, \frac{\min\{|z|, \sqrt{b}\, v_b\}}{2\Phi(\sqrt{b}\, v_b)-1}, & \text{for } t \neq 1, \end{cases}$$

with $v_b^2 = (1/3)\, g(\sqrt{b}\, v_b) > 0$, i.e. $v_b = (1/\sqrt{b})\, \overline{w}\,(3/b)$. For example, for $b = 4$ we have $v_b = 0.240$.

Example 6.2 (Quadratic regression)
In a quadratic regression model the observation at t_{nN} is given by

$$Y_{nN} = \beta_0 + \beta_1 t_{nN} + \beta_2 t_{nN}^2 + Z_{nN} = a(t_{nN})'\beta + Z_{nN}$$

for $n = 1, ..., N$, where $\beta = (\beta_0, \beta_1, \beta_2)' \in I\!R^3$ and $a(t) = (1, t, t^2)' \in I\!R^3$. Assume that $\mathcal{T} = [0,1]$ and that the interesting aspect is $L\beta := (\beta_0, \beta_1 + \beta_2)' \in I\!R^2$. In particular, testing the hypothesis $H_0 : L\beta = (l_1, 0)$ means that the hypothesis is tested that the quadratic function is symmetric on $[0,1]$ and is equal to l_1 at the end points $t = 0$ and $t = 1$.

The D-optimal design for $L\beta$ in $\Delta_{\mathcal{D}}$ with $\mathcal{D} = \{0,1\}$ is

$$\delta_D = \frac{1}{2}(e_0 + e_1).$$

Hence, according to Theorem 4.1 δ_D is most robust for testing $H_0 : L\beta = 0$ in $\Delta(L)$ with

$$b_0^T(\delta_D, L) = 2 = \min\{b_0^T(\delta, L);\ \delta \in \Delta(L)\},$$

and the influence function of a most robust ALE-test at δ_D has the form

$$(6.2) \qquad \psi_{01}(z,t) = \begin{cases} (1,-1)' \operatorname{sgn}(z)\, 2\, \sqrt{\frac{\pi}{2}} & \text{for } t = 0, \\ (0,1)' \operatorname{sgn}(z)\, 2\, \sqrt{\frac{\pi}{2}} & \text{for } t = 1. \end{cases}$$

According to Theorem 5.4 for every $b \geq b_0^T(\delta_D, L) = 2$ the D-optimal design δ_D is also D-optimal for testing with bias bound b and the influence function of a D-optimal ALE-test statistic for testing at δ_D with bias bound b is given by (5.5). For $b = b_0^T(\delta_D, L) = 2$ this influence function coincides with ψ_{01} given by (6.2) and for $b > b_0^T(\delta_D, L) = 2$ it has the form

$$\psi_{b,\delta_D}(z,t) = \begin{cases} (1,-1)' \operatorname{sgn}(z)\, 2\, \frac{\min\{|z|, \sqrt{b}\, v_b\}}{2\Phi(\sqrt{b}\, v_b)-1}, & \text{for } t = 0, \\ (0,1)' \operatorname{sgn}(z)\, 2\, \frac{\min\{|z|, \sqrt{b}\, v_b\}}{2\Phi(\sqrt{b}\, v_b)-1}, & \text{for } t = 1, \end{cases}$$

with $v_b^2 = (1/2)\, g(\sqrt{b}\, v_b) > 0$, i.e. $v_b = (1/\sqrt{b})\, \overline{w}\,(2/b)$. For example, for $b = 3$ we have $v_b = 0.377$.

7 Proofs

Proof of Theorem 2.1. Consider the following part of τ_N:

$$\sqrt{N} C_N^{-1/2} (\widehat{\varphi}_N - l) = \sqrt{N} C_N^{-1/2} (\widehat{\varphi}_N - \varphi(\beta_N)) + C_N^{-1/2} \gamma.$$

Because $(P^N_{\beta_N})_{N \in I\!N}$ is contiguous to $(P^N_\beta)_{N \in I\!N}$ the covariance estimator C_N also converges to $\sigma^2 C(\psi, \delta)$ in probability $(P^N_{\beta_N})_{N \in I\!N}$. Thus because $\widehat{\varphi}_N$ is weakly asymptotically linear for contiguous alternatives the distribution $\mathcal{L}(\sqrt{N} C_N^{-1/2} (\widehat{\varphi}_N - \varphi(\beta_N)) | P^N_{\beta_N})$ behaves asymptotically like

$$\begin{aligned} &\mathcal{L}\left(N^{-1/2} [\sigma^2 C(\psi, \delta)]^{-1/2} \sum_{n=1}^N \sigma \, \psi \left(\frac{y_{nN} - a(t_{nN})' \beta_N}{\sigma}, t_{nN} \right) \middle| P^N_{\beta_N} \right) \\ = \; &\mathcal{L}(N^{-1/2} \sum_{n=1}^N C(\psi, \delta)^{-1/2} \psi(z_{nN}, t_{nN}) | P^N). \end{aligned}$$

The third lemma of LeCam (see Hájek and Šidák (1967), p. 208) provides that $\mathcal{L}(N^{-1/2} \sum_{n=1}^N C(\psi, \delta)^{-1/2} \psi(z_{nN}, t_{nN}) | Q^N)$ is asymptotically normally distributed with mean $C(\psi, \delta)^{-1/2} b(\psi, (Q^N)_{N \in I\!N})$ and covariance matrix $E_{s \times s}$, and in particular that $(Q^N)_{N \in I\!N}$ is contiguous to $(P^N)_{N \in I\!N}$ for all $(Q^N)_{N \in I\!N} \in \mathcal{U}_\epsilon$. Then also $(Q^N_{\beta_N})_{N \in I\!N}$ is contiguous to $(P^N_{\beta_N})_{N \in I\!N}$ such that $C_N^{-1/2} \gamma$ converges to $[\sigma^2 C(\psi, \delta)]^{-1/2} \gamma$ in probability for $(Q^N_{\beta_N})_{N \in I\!N}$ and $\mathcal{L}(\sqrt{N} [\sigma^2 C(\psi, \delta)]^{-1/2} (\widehat{\varphi}_N - l)) | Q^N_{\beta_N})$ is asymptotically normally distributed with mean $[\sigma^2 C(\psi, \delta)]^{-1/2} (\gamma + \sigma \, b(\psi, (Q^N)_{N \in I\!N})$ and covariance matrix $E_{s \times s}$. □

Proof of Theorem 3.1. ¿From Theorem 2.1 and the fact that $\mathcal{X}_{s,b}(k)$ is decreasing in b for all $k > 0$ we get at once that a) is equivalent with

$$\max\{b(\psi, (Q^N)_{N \in I\!N})' C(\psi, \delta)^{-1} b(\psi, (Q^N)_{N \in I\!N}); \; (Q^N)_{N \in I\!N} \in \mathcal{U}_\epsilon\} \le b. \tag{7.1}$$

The equivalence of (7.1) and b) follows from the results in Rieder (1985, 1994) and Kurotschka and Müller (1992) about the maximum asymptotic bias of asymptotically linear estimators in shrinking contamination neighbourhoods by interpreting the function $C(\psi, \delta)^{-1/2} \psi$ as influence function. □

To derive the assertions for D-optimal designs we need a modification of the equivalence theorem for D-optimality of Kiefer and Wolfowitz (1960).

Lemma 7.1 *Let $\Delta_\mathcal{D} \subset \Delta(L)$. Then δ_D is D-optimal for $L\beta$ in $\Delta_\mathcal{D}$ if and only if*

$$a(t)' I(\delta_D)^- L' [L \, I(\delta_D)^- L']^{-1} L \, I(\delta)^- a(t) = s \text{ for all } t \in \mathcal{D}.$$

Proof. If besides $L\beta$ also β is identifiable at δ_D then δ_D is D-optimal for $L\beta$ in $\Delta_{\mathcal{D}}$ if and only if

$$(7.2) \quad a(t)'\, I(\delta_D)^{-1}\, L'\, [L\, I(\delta_D)^{-1}\, L']^{-1}\, L\, I(\delta)^{-1}\, a(t) = s \text{ for all } t \in \mathcal{D}$$

(see Kiefer (1974), Pázman (1986), p.117). If β is not identifiable at δ_D then the linear hull of $\{a(t);\ t \in \mathcal{D}\}$ is of dimension $q < r$. Hence there exists a matrix $A \in I\!R^{r\times q}$ of rank q and a g-inverse A^- of A such that $a(t) = A\, A^- a(t)$ for all $t \in \mathcal{D}$ and $L = L\,(A^-)'\, A'$. Setting $\tilde\beta := A'\beta$, $\tilde a := A^- a$ and $\tilde L := L(A^-)'$ we have a reparametrized model with $a(t)'\beta = \tilde a(t)'\tilde\beta$ for all $t \in \mathcal{D}$, $\tilde L\tilde\beta = L\beta$ and $(\int \tilde a \tilde a'\, d\delta_D)^{-1} = A' I(\delta_D)^- A$. In particular $\tilde\beta$ is identifiable at δ_D. Then the assertion follows by applying (7.2) on the reparametrized model. □

Proof of Theorem 4.1. Assertion a) follows by a straightforward generalization of the proof of Theorem 2' in Ronchetti and Rousseeuw (1985) (or Proposition 1 (ii) in Hampel et al. (1986), p. 318).
b) Because of $\int \psi_{01}\, \psi_{01}'\, d(P \otimes \delta_D) = L\, I(\delta_D)^-\, L'\, \frac{\pi}{2}$ Lemma 7.1 provides

$$\psi_{01}(z,t)'\, C(\psi_{01}, \delta_D)^{-1} \psi_{01}(z,t) = a(t)'\, I(\delta_D)^-\, L'\, (L\, I(\delta_D)^-\, L')^{-1}\, L\, I(\delta_D)^-\, a(t) = s$$

for all $t \in \mathcal{D}$. Hence we have $b_0^T(\delta_D, L) \le s$. Because a) yields $b_0^T(\delta_D, L) \ge s$ the assertion follows. □

Proof of Corollary 4.2. Because $h(t) > 0$ for all $t \in \mathcal{T}$ there exists $K_1 \in I\!R^{s\times r}$ and $K_2 \in I\!R^{r\times r}$ with $L = K_1 J_h$ and $J_h = K_2 I_h(\delta)$. Then we have

$$\begin{aligned} &\int \psi_0(z,t)\, a(t)'\, z\, P \otimes \delta(dz, dt) = L_h\, I_h(\delta)^-\, J_h \\ &= \quad K_1\, J_h\, J_h^-\, I_h(\delta)\, I_h(\delta)^-\, I_h(\delta)\, K_2' = K_1\, J_h = L. \end{aligned}$$

Hence $\psi_0 \in \Psi^*(\delta, L)$ so that the assertion follows by applying Theorem 4.1 on the modified model. □

Proof of Theorem 5.1. Because $L = K\, M_b(B_b)$ and $L = K_0\, M_{b,0}(B_b)$ for some $K,\ K_0 \in I\!R^{s\times r}$, it is easy to see that $\psi_{b,\delta}$ is an element of $\Psi^*(\delta, L)$ with $\|\psi_{b,\delta}' C(\psi_{b,\delta}, \delta)^{-1} \psi_{b,\delta}\|_\delta \le b$. To show that $\psi_{b,\delta}$ is admissible the proof of Krasker and Welsch (1982) is generalized:

At first assume

$$\psi_{b,\delta}(z,t) \;:=\; L\, M_b(B_b)^-\, a(t)\, \min\left\{|z|, \frac{\sqrt{b}}{|B_b a(t)|}\right\} \operatorname{sgn}(z).$$

Setting $V := L\, M_b(B_b)^-\, Q_b(B_b)\, M_b(B_b)^-\, L'$, $Q := V^{-1/2} L\, M_b(B_b)^-$, $D := V^{-1/2} L$ and $\psi_B(z,t) := V^{-1/2}\psi_{b,\delta}$ one gets because of (5.2) $|B_b\, a(t)| = |Q\, a(t)|$ for all $t \in \operatorname{supp}(\delta)$ and

$$\psi_B(z,t) = Q\, a(t)\, \min\left\{|z|, \frac{\sqrt{b}}{|Q\, a(t)|}\right\} \operatorname{sgn}(z).$$

Because of $\int \psi_B(z,t)a(t)'z\,P(dz)\delta(dt) = D$, Theorem 1 in Kurotschka and Müller (1992) provides

(7.3) $\quad \psi_B \in \arg\min\{\mathrm{tr}(C(\psi,\delta)); \psi \in \Psi^*(\delta, D) \text{ with } \|\psi'\psi\|_\delta \le b\}.$

Assume that there exists $\psi_0 \in \Psi^*(\delta, L)$ with

(7.4) $\quad C(\psi_0,\delta) < C(\psi_{b,\delta},\delta)$

and

(7.5) $\quad \|\psi_0' C(\psi_0,\delta)^{-1}\psi_0\|_\delta \le b.$

Then $\int V^{-1/2}\psi_0(z,t)a(t)'z\,P(dz)\delta(dt) = V^{-1/2}L = D$ and

$$\begin{aligned} & V^{-1/2}\,C(\psi_0,\delta)\,V^{-1/2} < V^{-1/2}\,C(\psi_{b,\delta},\delta)\,V^{-1/2} = \int \psi_B\psi_B'\,d(P\otimes\delta) \\ = \;& V^{-1/2}L\,M_b(B_b)^- Q_b(B_b) M_b(B_b)^- L'V^{-1/2} = E_{r\times r}, \end{aligned}$$

which implies with (7.5)

$$\|\psi_0' V^{-1/2} V^{-1/2}\psi_0\|_\delta \le \left\| \psi_0' V^{-1/2}\left(V^{-1/2}\,C(\psi_0,\delta)\,V^{-1/2}\right)^{-1} V^{-1/2}\psi_0 \right\|_\delta \le b.$$

Hence, property (7.3) provides for $V^{-1/2}\psi_0$

$$\begin{aligned} & \mathrm{tr}(V^{-1/2}\,C(\psi_0,\delta)\,V^{-1/2}) = \mathrm{tr}(C(V^{-1/2}\,\psi_0,\delta)) \ge \mathrm{tr}(C(\psi_B,\delta)) \\ = \;& \mathrm{tr}(V^{-1/2}\,C(\psi_{b,\delta},\delta)\;\;V^{-1/2}) \end{aligned}$$

which is a contradiction to (7.4).

For

$$\psi_{b,\delta}(z,t) \;:=\; L\,M_{b,0}(B_b)^-\,a(t)\,\frac{\sqrt{b}}{|B_b a(t)|}\,\mathrm{sgn}(z)$$

the assertion follows as above. We have only to replace $M_b(B_b)$ and $Q_b(B_b)$ by $M_{b,0}(B_b)$ and $Q_{b,0}(B_b)$. $\Box$

Proof of Lemma 5.2. Set

$$B_b = \begin{cases} (L\,I(\delta_D)^-\,L')^{-1/2}\,L\,I(\delta_D)^-\,\frac{1}{\sqrt{s}\,y_b} & \text{for } b > s, \\ (L\,I(\delta_D)^-\,L')^{-1/2}\,L\,I(\delta_D)^- & \text{for } b = s. \end{cases}$$

Then the equivalence theorem for D-optimality (Lemma 7.1) provides $|B_b\,a(t)| = 1/y_b$ for $b > s$ and $|B_b\,a(t)| = \sqrt{s}$ for $b = s$ for all $t \in \mathcal{D}$ so that the admissibility of ψ_{b,δ_D} follows from Theorem 5.1. Moreover, we obtain for $b > s$ because of $\sqrt{b}\,y_b = \overline{w}\,(s/b)$

$$C(\psi_{b,\delta_D},\delta_D) = \int \psi_{b,\delta_D}\,\psi_{b,\delta_D}\,d(P\otimes\delta_D) = \frac{1}{\overline{v}\left(\frac{s}{b}\right)} L\,I(\delta_D)^-\,L'$$

and for $b = s$

$$C(\psi_{b,\delta_D}, \delta_D) = \int \psi_{b,\delta_D}\, \psi_{b,\delta_D}\, d(P \otimes \delta_D) = \frac{1}{\overline{v}(1)} L\, I(\delta_D)^{-}\, L'. \square$$

For $s = 1$ the admissible solution for a bias bound b is also universally optimal for the bias bound b. In particular, we get the following corollary of Theorem 5.1, which is useful to prove Lemma 5.3.

Corollary 7.2 *If*

$$\rho_b(z) = \begin{cases} sgn(z) & \textit{for } b = 1, \\ sgn(z)\ \min\{|z|, \sqrt{b}\, y\} & \textit{for } b > 1, \end{cases}$$

where $y^2 = g(\sqrt{b}\, y) > 0$*, then we have*

$$\frac{\int \rho_b(z)^2\, P(dz)}{(\int \rho_b(z)\, z\, P(dz))^2} \leq \frac{\int \rho(z)^2\, P(dz)}{(\int \rho(z)\, z\, P(dz))^2}$$

for all $\rho : I\!R \to I\!R$ *with* $\max_{z \in I\!R} \frac{\rho(z)^2}{\int \rho(y)^2\, P(dy)} \leq b$*,* $\int \rho(z)\, P(dz) = 0$ *and* $\int \rho(z)^2\, P(dz) < \infty$*.*

Proof. The assertion follows from Theorem 5.1 by setting $a(t) = 1$ for all $t \in \mathcal{T}$ and $L = 1$ and by using $B_b = 1/y$ for $b > 1$ and $B_b \neq 0$ arbitrarily for $b = 1$ (see also Corollary 9.8 in Müller (1995)). $\square$

Lemma 7.3 $\overline{v} : (0, 1] \to I\!R$ *is decreasing and concave.*

Proof. Define

$$\begin{aligned} \tilde{w}(b) &:= \overline{w}\left(\frac{1}{b}\right) \text{ and } u(b) := \overline{v}\left(\frac{1}{b}\right) \text{ for } b \geq 1, \\ l(y) &:= 2\Phi(y) - 1 - 2y\Phi'(y) \text{ and } h(y) := y\Phi(-y) - \Phi'(y) \text{ for } y \in I\!R, \end{aligned}$$

and $w := \tilde{w}(b)$ for a fixed $b > 1$. Then $g(y) = l(y) + 2y^2\Phi(-y)$, $l(y) > 0$ and $h(y) < 0$ for $y > 0$ and with the implicit function theorem we get for fixed $b > 1$

$$\tilde{w}'(b) = \frac{g(w)\, w}{2\, b\, l(w)} > 0.$$

This implies

$$u'(b) = \frac{-(2\Phi(w) - 1)\, 2\, h(w)}{w} > 0 \tag{7.6}$$

and

$$u''(b) = \frac{\tilde{w}'(b)\,2}{w^2}\,[h(w)\,l(w) - (2\Phi(w)-1)\,w\,\Phi(-w)] < 0.$$

In particular we have

$$(7.7)\quad b\,u''(b) + 2\,u'(b) = \frac{1}{w\,l(w)}\,[h(w)\,l(w)\,g(w) - (2\Phi(w)-1)\,f(w)] < 0$$

because $f(y) := y\Phi(-y)\,g(y) + 4h(y)l(y) > 0$ for $y > 0$. See also Lemma 12.7 in Müller (1995). Because of (7.6) and (7.7) we get

$$\overline{v}'(c) = u'\left(\frac{1}{c}\right)\,\frac{-1}{c^2} < 0$$

and

$$\overline{v}''(c) = \frac{1}{c^4}\,u''\left(\frac{1}{c}\right) + \frac{2}{c^3}\,u'\left(\frac{1}{c}\right) < 0.\square$$

Proof of Lemma 5.3. Set

$$\begin{aligned}
A_{\mathcal{D}} &:= (a(\tau_1),...,a(\tau_I))',\\
M(\rho) &:= \int a(t)\,a(t)'\,\rho(z,t)\,z\,P(dz)\,\delta(dt),\\
Q(\rho) &:= \int a(t)\,a(t)'\,\rho(z,t)^2\,P(dz)\,\delta(dt),\\
D &:= \mathrm{diag}(\delta(\{\tau_1\}),...,\delta(\{\tau_I\})),\\
D_1(\rho) &:= \mathrm{diag}(\textstyle\int \rho(z,\tau_1)\,z\,P(dz),...,\int \rho(z,\tau_I)\,z\,P(dz)),\\
D_2(\rho) &:= \mathrm{diag}(\textstyle\int \rho(z,\tau_1)^2\,P(dz),...,\int \rho(z,\tau_I)^2\,P(dz)),
\end{aligned}$$

and

$$D_\rho \;:=\; \mathrm{diag}(\delta_\rho(\{\tau_1\}),...,\delta_\rho(\{\tau_I\})),$$

where $\mathrm{diag}(\lambda_1,...,\lambda_I)$ is the diagonal matrix with diagonal elements $\lambda_1,...,\lambda_I$. Because $a(\tau_1),...,a(\tau_I)$ are linearly independent we have $A_{\mathcal{D}}(A'_{\mathcal{D}}BA_{\mathcal{D}})^- A'_{\mathcal{D}} = B^{-1}$ for all regular $B \in I\!R^{I\times I}$. (See also Lemma 8.10 in Müller (1995)). Note, if $\psi \in \Psi^*(\delta, L)$ exists then $L\beta$ is identifiable at δ so that $L = K\,A_{\mathcal{D}}$ for some $K \in I\!R^{s\times I}$ (see Lemma 1 in Kurotschka and Müller (1992)). Then condition (2.2) in the definition of $\psi \in \Psi^*(\delta, L)$ provides $\psi(z,t) = L\,M(\rho)^-\,a(t)\,\rho(z,t)$ for all $(z,t) \in I\!R \times \mathcal{D}$. This implies

$$\begin{aligned}
(7.8)\quad C(\psi,\delta) &= K\,D_1(\rho)^{-1}\,D_2(\rho)\,D_1(\rho)^{-1}\,D^{-1}\,K'\\
&= \frac{1}{\sum_{t\in\mathcal{D}} v_\rho(t)\,\delta(\{t\})}\,K\,D_\rho^{-1}\,K' = \frac{1}{\sum_{t\in\mathcal{D}} v_\rho(t)\,\delta(\{t\})}\,L\,I(\delta_\rho)^-\,L'.
\end{aligned}$$

Moreover, for all $z \in I\!\!R$ and all $i = 1, ..., I$ we have

$$\begin{aligned} b &\geq \psi(z,\tau_i)' C(\psi,\delta)^{-1} \psi(z,\tau_i) \\ &= \rho(z,\tau_i)^2\, u_i'\, D_1(\rho)^{-1}\, D^{-1}\, K' \left(\textstyle\sum_{t\in\mathcal{D}} v_\rho(t)\, \delta(\{t\})\right) [L\, I(\delta_\rho)^-\, L']^{-1} \\ &\quad K\, D_1(\rho)^{-1}\, D^{-1}\, u_i \\ &= \frac{\rho(z,\tau_i)^2}{\int \rho(y,\tau_i)^2 P(dy)}\, \frac{1}{\delta_\rho(\{\tau\})\, \delta(\{\tau_i\})}\, u_i'\, K'\, [L\, I(\delta_\rho)^-\, L']^{-1}\, K\, u_i \\ &= \frac{\rho(z,\tau_i)^2}{\int \rho(y,\tau_i)^2 P(dy)}\, \frac{\delta_\rho(\{\tau\})}{\delta(\{\tau_i\})}\, a(\tau_i)'\, I(\delta_\rho)^-\, L'[L\, I(\delta_\rho)^-\, L']^{-1}\, L\, I(\delta_\rho)^-\, a(\tau_i). \end{aligned}$$

Setting

$$b_i := \frac{b\, \delta(\{\tau_i\})}{\delta_\rho(\{\tau_i\})\; a(\tau_i)'\, I(\delta_\rho)^-\, L'\; [L\, I(\delta_\rho)^-\, L']^{-1}\, L\, I(\delta_\rho)^-\, a(\tau_i)}$$

we have for $i = 1, ..., I$

$$\max_{z\in I\!\!R} \frac{\rho(z,\tau_i)^2}{\int \rho(y,\tau_i)^2 P(dy)} \leq b_i$$

so that Theorem 4.1 a) applied to the one dimensional case provides $b_i \geq 1$. Defining

$$\rho_0(z,\tau_i) := \begin{cases} \operatorname{sgn}(z) & \text{for } b_i = 1, \\ \operatorname{sgn}(z)\, \min\{|z|, \sqrt{b_i}\, y_i\} & \text{for } b_i > 1, \end{cases}$$

where $y_i^2 = g(\sqrt{b_i}\, y_i) > 0$, we obtain according to Corollary 7.2

$$v_\rho(\tau_i) \leq v_{\rho_0}(\tau_i) = \overline{v}\left(\frac{1}{b_i}\right)$$

for $i = 1, ..., I$. Note that $\sqrt{b_i}\, y_i = \overline{w}\,(1/b_i)$. Because according to Lemma 7.3 $\overline{v}$ is concave we therefore have

$$\begin{aligned} \textstyle\sum_{i=1}^I v_\rho(\tau_i)\, \delta(\{\tau_i\}) &\leq \textstyle\sum_{i=1}^I \overline{v}\left(\frac{1}{b_i}\right)\, \delta(\{\tau_i\}) \leq \overline{v}\left(\textstyle\sum_{i=1}^I \frac{1}{b_i}\, \delta(\{\tau_i\})\right) \\ &= \overline{v}\left(\frac{1}{b}\operatorname{tr}\left(\textstyle\sum_{i=1}^I [L\, I(\delta_\rho)^-\, L']^{-1}\, L\, I(\delta_\rho)^-\, a(\tau_i)\, a(\tau_i)'\, I(\delta_\rho)^-\, L'\, \delta_\rho(\{\tau_i\})\right)\right) \\ &= \overline{v}\left(\frac{1}{b}\operatorname{tr}(E_{s\times s})\right) = \overline{v}\left(\frac{s}{b}\right). \end{aligned}$$

Then (7.8) implies the assertion. □

Proof of Theorem 5.4. Regard any $\delta \in \Delta$ with $\operatorname{supp}(\delta) = \{\tau_1, ..., \tau_I\}$ and any $\psi \in \Psi^*(\delta, L)$ with $\|\psi' C(\psi,\delta)^{-1}\psi\|_\delta \leq b$. If $a(\tau_1), ..., a(\tau_I)$ are not linearly independent, then we can extend the regressors by $\tilde{a}(t)$ so that $\overline{a}(\tau_1), ..., \overline{a}(\tau_I)$ are linearly independent, where $\overline{a}(t) = (a(t)', \tilde{a}(t)')'$. Then we have $\psi \in \overline{\Psi}^*(\delta, \overline{L})$ for some $\tilde{L}$, where $\overline{L} = (L, \tilde{L})$

and $\overline{\Psi}^*(\delta, \overline{L})$ is defined for the extended model given by $\overline{Y} = \overline{a}(t)'\overline{\beta} + Z$. Denoting $\overline{I}(\delta) = \int \overline{a}(t)\, \overline{a}(t)'\, \delta(dt)$ Lemma 5.3 provides

$$\text{(7.9)} \quad \det(C(\psi, \delta)) \geq \left(\frac{1}{\overline{v}\left(\frac{s}{b}\right)} \right)^s \det(\overline{L}\, \overline{I}(\delta_\rho)^- \,\overline{L}')$$

$$\geq \left(\frac{1}{\overline{v}\left(\frac{s}{b}\right)} \right)^s \det(L\, I(\delta_\rho)^- \, L') \geq \left(\frac{1}{\overline{v}\left(\frac{s}{b}\right)} \right)^s \det(L\, I(\delta_D)^- \, L')$$

(see also Lemma A.15 in Müller (1987) for the property $\overline{L}\, \overline{I}(\delta_\rho)^- \,\overline{L}' \geq L\, I(\delta_\rho)^- \, L'$). According to Lemma 5.2 the lower bound in (7.9) is attained by an ALE-test with influence function ψ_{b,δ_D} at δ_D. □

References

[1] AKRITAS, M.G. (1991). Robust M estimation in the two-sample problem. *J. Amer. Statist. Assoc.* **86**, 201-204.

[2] BICKEL, P.J. (1981). Quelque aspects de la statistique robuste. In École d'Été de Probabilités de St. Flour. *Springer Lecture Notes in Math.* **876**, 1-72.

[3] BICKEL, P.J. (1984). Robust regression based on infinitesimal neighbourhoods. *Ann. Statist.* **12**, 1349-1368.

[4] CHRISTENSEN, R. (1987). *Plane Answers to Complex Questions - The Theory of Linear Models.* Springer, New York.

[5] COAKLEY, C.W. and HETTMANSPERGER, T.P. (1992). Breakdown bounds and expected resistance. *Journal of Nonparametric Statistics* **1**, 267-276.

[6] COAKLEY, C.W. and HETTMANSPERGER, T.P. (1994). The maximum resistance of tests. *Austral. J. Statist.* **36**, 225-233.

[7] HÁJEK, J. and ŠIDÁK, Z. (1967). *Theory of Rank Tests.* Academic Press, New York.

[8] HAMPEL, F.R. (1978). Optimally bounding the gross-error-sensitivity and the influence of position in factor space. *Proceedings of the ASA Statistical Computing Section*, ASA, Washington, D.C., 59-64.

[9] HAMPEL, F.R., RONCHETTI, E.M., ROUSSEEUW, P.J. and STAHEL, W.A. (1986). *Robust Statistics - The Approach Based on Influence Functions.* John Wiley, New York.

[10] HERITIER, S. and RONCHETTI, E. (1994). Robust bounded-influence tests in general parametric models. *J. Amer. Statist. Assoc.* **89**, 897-904.

[11] KIEFER, J. (1974). General equivalence theory for optimum designs (approximate theory). *Ann. Statist.* **2**, 849-879.

[12] KIEFER, J. and WOLFOWITZ, J. (1960). The equivalence of two extremum problems. *Can. J. Math.* **12**, 363-366.

[13] KRASKER, W.S. (1980). Estimation in linear regression models with disparate data points. *Econometrica* **48**, 1333-1346.

[14] KRASKER, W.S. and WELSCH, R.E. (1982). Efficient bounded-influence regression estimation. *J. Amer. Statist. Assoc.* **77**, 595-604.

[15] KUROTSCHKA, V. and MÜLLER, Ch.H. (1992). Optimum robust estimation of linear aspects in conditionally contaminated linear models. *Ann. Statist.* **20**, 331-350.

[16] MARKATOU, M. and HE, X. (1994). Bounded influence and high breakdown point testing procedures in linear models. *J. Amer. Statist. Assoc.* **89**, 543-549.

[17] MARKATOU, M. and HETTMANSPERGER, T.P. (1990). Robust bounded-influence tests in linear models. *J. Amer. Statist. Assoc.* **85**, 187-190.

[18] MARKATOU, M., STAHEL, W.A. and RONCHETTI, E. (1991). Robust M-type testing procedures for linear models. In: *Directions in Robust Statistics and Diagnostics (Part I)*, eds. W. Stahel and S. Weisberg, Springer, New York, 201-220.

[19] MÜLLER, Ch.H. (1987). Optimale Versuchspläne für robuste Schätzfunktionen in linearen Modellen. *Ph. D. thesis.* Freie Universität Berlin.

[20] MÜLLER, Ch.H. (1992a). L_1-estimation and testing in conditionally contaminated linear models. In *L_1-Statistical Analysis and Related Methods*, ed. Y. Dodge, North-Holland, Amsterdam, 69-76.

[21] MÜLLER, Ch.H. (1992b). Robust estimation with minimum bias and A-optimal designs. In *PROBASTAT'91, Proceedings of the International Conference on Probability and Mathematical Statistics, Bratislava 1991*, Pázman, A. and J. Volaufová (eds.), 109-115.

[22] MÜLLER, Ch.H. (1992c). Robust testing in conditionally contaminated linear models. *Preprint No. A-92-12*, Freie Universität Berlin, Fachbereich Mathematik und Informatik.

[23] MÜLLER, Ch.H. (1994a). Optimal designs for robust estimation in conditionally contaminated linear models. *J. Statist. Plann. Inference.* **38**, 125-140.

[24] MÜLLER, Ch.H. (1994b). One-step-M-estimators in conditionally contaminated linear models. *Stat. Decis.* **12**, 331-342.

[25] MÜLLER, Ch.H. (1995). Outlier Robust Inference for Planned Experiments. *Habilitation thesis.* Freie Universität Berlin.

[26] PÁZMAN, A. (1986). *Foundations of Optimum Experimental Design.* Reidel, Dordrecht.

[27] PUKELSHEIM, F. (1993). *Optimal Design of Experiments.* John Wiley, New York.

[28] RIEDER, H. (1985). Robust estimation of functionals. *Technical Report.* Universität Bayreuth.

[29] RIEDER, H. (1987). Robust regression estimators and their least favorable contamination curves. *Stat. Decis.* **5**, 307-336.

[30] RIEDER, H. (1994). *Robust Asymptotic Statistics*, Springer, New York.

[31] RONCHETTI, E. (1982). Robust testing in linear models: The infinitesimal approach. *Ph. D. thesis.* Swiss Federal Institute of Technology, Zürich.

[32] RONCHETTI, E. and ROUSSEEUW, P.J. (1985). Change-of-variance sensitivities in regression analysis. *Z. Wahrsch. verw. Gebiete* **68**, 503-519.

[33] SILVAPULLE, M.J. (1992a). Robust Wald-type tests of one-sided hypothesis in the linear model. *J. Amer. Statist. Assoc.* **87**, 156-161.

[34] SILVAPULLE, M.J. (1992b). Robust tests of inequality constraints and one-sided hypotheses in the linear model. *Biometrika* **79**, 621-630.

[35] STAUDTE, R.G. and SHEATHER, S.J. (1990). *Robust Estimation and Testing.* Wiley, New York.

[36] WALD, A. (1943). Tests of statistical hypothesis concerning several parameters when the number of observations is large. *Transactions Amer. Math. Soc.* **3**, 426-482.

Freie Universität Berlin
Fachbereich Mathematik und Informatik
1. Mathematisches Institut
Arnimallee 2-6
D-14 195 Berlin
Germany