Regression Quantiles and Improved L-Estimation in Linear Models

by

Jana Jureckova, A.K.M. Ehsanes Saleh and Pranab Kumar Sen

Department of Biostatistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 1829

July 1987

REGRESSION QUANTILES AND IMPROVED L-ESTIMATION IN LINEAR MODELS

JANA JUREČKOVÁ, A.K.M.EHSANES SALEH AND PRANAB KUMAR SEN

Charles University, Prague, Carleton University, Ottawa and University of North Carolina, Chapel Hill.

ABSTRACT. For the usual linear model, bearing the plausibility of a redundant subset of parameters, pre-test and Stein-rule estimators based on the trimmed least squares estimation theory are considered. Compared to parallel M-estimators, the proposed L-estimators are computationally simpler and are scale-equivariant too. In the light of asymptotic distributional risks, the relative (risk-)efficiency results for these trimmed L-estimators and their improved versions are studied in detail. Positive-rule L-estimators are also considered in this context.

1. INTRODUCTION. Consider the usual regression model

$$Y_i = \underset{\sim}{\beta}'\underset{\sim}{x}_i + e_i \ ; \ \underset{\sim}{x}_i' = (x_{oi},\ldots,x_{pi}) \ , \ x_{oi} = 1, \text{ for } i=1,\ldots,n, \tag{1.1}$$

where $\underset{\sim}{\beta} = (\beta_o,\ldots,\beta_p)'$ is a vector of unknown parameters, the $\underset{\sim}{x}_i$ are known vectors of regression constants, and the errors e_i are independent and identically distributed random variables (i.i.d.r.v.) with a (unknown) continuous distribution function F, defined on the real line R . We partition $\underset{\sim}{\beta}' = (\underset{\sim}{\beta}_1', \underset{\sim}{\beta}_2')$ where $\underset{\sim}{\beta}_j$ is a p_j-vector, j=1,2; $p_1+p_2 = p+1$. We are primarily interested in (robust) estimation of $\underset{\sim}{\beta}_1$ when $\underset{\sim}{\beta}_2$, though unknown, is suspected to be close to a **pivot** $\underset{\sim}{\beta}_2^o$, which we may take (without any loss of generality) as $\underset{\sim}{0}$. For situations involving such an uncertain prior , we may refer to Saleh and Sen (1987) where improved least squares estimators (LSE) of $\underset{\sim}{\beta}_1$ have been studied. Since the LSE are known to be generally non-robust, other possibilities for such improved estimators (retaining robustness) include the R-estimators and M-estimators which were treated earlier by Saleh and Sen (1986) and Sen and Saleh (1987), among others. However, generally, these R- and M-estimators are computationally quite cumbersome , and , moreover, the

---

AMS 1980 Subject Classification No.: 62C15, 62F10, 62G05, 62H12

Key Words and Phrases: *Asymptotic distributional risk; asymptotic distributional risk-efficiency; linear models; L-estimation; local alternatives; minimaxity; preliminary test estimator; quadratic loss; regression quantiles; robustness; shrinkage estimation; trimmed least squares estimators.*

M-estimators are not usually scale-equivariant, and their studentized versions are generally less attractive. For these reasons, we propose to incorporate the **trimmed least squares estimators (TLSE)** in this improved estimation problem. For a good account of the TLSE, we may refer to Jurečková and Sen (1984) where the earlier works by Koenker and Bassett (1978), Ruppert and Carroll (1980), Jurečková (1983 a,b, 1984) and others are all cited.

Corresponding to pre-specified left and right trimming proportions α_1, $1-\alpha_2$, we may denote a TLSE of β in (1.1) by $L_n(\alpha_1,\alpha_2)$; here $0 < \alpha_1 < \alpha_2 < 1$. This estimator for the full model in (1.1) is termed an **unrestricted TLSE (UTLSE)**. Next, we partition $x_i = (x_{i1}', x_{i2}')'$ and consider the restricted model $Y_i = \beta_1' x_{i1} + e_i$, $i=1, \ldots, n$. For this model, the TLSE of β_1, denoted by $L^*_{n1}(\alpha_1,\alpha_2)$, is termed **a restricted** TLSE (RTLSE). We also let $L_n(\alpha_1,\alpha_2) = (L_{n1}'(\alpha_1,\alpha_2), L_{n2}'(\alpha_1,\alpha_2))'$. If $\beta_2 = 0$, then generally $L^*_{n1}(\alpha_1,\alpha_2)$ has a smaller risk (with quadratic loss) than $L_{n1}(\alpha_1,\alpha_2)$. However, when $\beta_2 \neq 0$, $L^*_{n1}(\alpha_1,\alpha_2)$ may not only become biased, it may as well be inconsistent and inefficient(relative to the TLSE). A **pre-test estimator (PTE)** may usually be used to eliminate this undesirability of the RTLSE : A preliminary test of the null hypothesis H_0: $\beta_2 = 0$ may be incorporated in choosing between the UTLSE and RTLSE ; we term this estimator as a **preliminary test TLSE (PTTLSE)**. In the case of the usual LSE, it has been observed [viz., Saleh and Sen (1984a,b)] that a PTE is a good comprom se : Unlike the RLSE, it does not have unbounded risk (when β_2 moves away from 0), while, like the RLSE, near the pivot (0), the risk of the PTE is smaller than the LSE. However, the PTE may not dominate either the ULSE or RLSE. Some further improvements are possible if p_1 and p_2 are both greater than 2. Shrinkage (or **Stein-rule**) versions of the LSE (SLSE) can be constructed and in some well defined manner, the SLSE may dominate the ULSE (although it may not dominate the PTE); we may refer to Saleh and Sen (1987) where other works have also been discussed in detail. We propose to extend this picture to the general case of TLSE and advocate the use of these improved versions .Our primary emphasis will be on the PTTLSE and the **shrinkage TLSE (STLSE)**.

Along with the preliminary notions, the different versions of the TLSE are introduced in Section 2. Section 3 deals with the notion of asymptotic distributional risk (ADR) where local alternatives (to the pivot) play a prominant role. In this context, we need to study also the asymptotic properties of tests for $H_0: \beta_2 = 0$ based on these TLSE, and these are presented in Section 4. The main results on the ADR along with the relative dominance picture are considered in Section 5. A brief treatment of the positive-rule TLSE (PRTLSE) is also included in this section.

2. THE PROPOSED VERSIONS OF TLSE

Corresponding to the model (1.1), we write $Y_n = (Y_1,\ldots,Y_n)'$ and $X_n' = (x_1, \ldots, x_n)$, so that we may rewrite (1.1) as $Y_n = X_n\beta + e_n$; $e_n = (e_1,\ldots,e_n)'$. Also, we partition X_n as (X_{1n}, X_{2n}) where X_{in} is $n\times p_i$, for i=1,2. Then, the restrained model is $Y_n = X_{n1}\beta_1 + e_n$. We denote by

$$nQ_n = X_n'X_n \quad \text{and partition } Q_n = ((Q_{nij}))_{i,j=1,2}, \tag{2.1}$$

where Q_{nij} is of order $p_i\times p_j$, i,j=1,2. We make the following assumptions.

(A) There exists a positive definite (p.d.) matrix Q $(= ((Q_{ij}))_{i,j=1,2})$, such that

$$\lim_{n\to\infty} Q_n = Q \quad \text{and} \quad \max_{1\le i\le n}\{x_i'Q_n^{-1}x_i\} = O(n^{\frac{1}{2}}), \text{ as } n\to\infty. \tag{2.2}$$

For later use, we denote by

$$Q_n^{-1} = ((Q_n^{ij}))_{i,j=1,2} \quad \text{and } Q^{-1} = ((Q^{ij}))_{i,j=1,2} = ((Q^*_{ij}))_{i,j=1,2} = Q^*. \tag{2.3}$$

(B) Let α_1, α_2 be positive numbers such that $0 < \alpha_1 < \alpha_2 < 1$, and consider a compact interval $J(\alpha_1,\alpha_2,\eta) = [F^{-1}(\alpha_1) - \eta, F^{-1}(\alpha_2)+\eta] = J_1(\alpha_1;\eta) \cup J_0(\alpha_1,\alpha_2,\eta) \cup J_2(\alpha_2,\eta)$, where $J_i(\alpha_i,\eta) = [F^{-1}(\alpha_i)-\eta, F^{-1}(\alpha_i)+\eta]$, i=1,2 and $J_0(\alpha_1,\alpha_2,\eta) = [F^{-1}(\alpha_1)+\eta, F^{-1}(\alpha_2) - \eta]$. We assume that the d.f. F is absolutely continuous and its density f is positive and continuous on $J_i(\alpha_i,\eta)$, for i=1,2 and some $\eta > 0$; also, we assume that f has a bounded first derivative f' in $J_i(\alpha_i,\eta)$, i=1,2.

Following Koenker and Bassett (1978), we define the unrestricted regression quantiles (URQ) $\tilde{\beta}(\alpha_i)$ $(\in R^{p+1})$, i=1,2 (for the full model in (1.1)) as the solutions (t) to the minimization problem

$$\Sigma_{j=1}^{n}\rho_{\alpha_i}(Y_j - x_j't) = \min, \quad \text{where } \rho_s(x) = x[s - I(x<0)],\ x \in R^1. \tag{2.4}$$

Although the solution in (2.4) may not be generally unique, one can always adapt a convention to choose a unique one. Let $\underset{\sim}{A}_n$ be diagonal matrix with the elements

$$a_{ii} = 0, \text{ if } Y_i \le \underset{\sim}{x}_i' \tilde{\underset{\sim}{\beta}}(\alpha_1) \text{ or } Y_i \ge \underset{\sim}{x}_i' \tilde{\underset{\sim}{\beta}}(\alpha_2)$$
$$= 1 \text{ , otherwise , } \qquad \text{for } i=1,\dots,n. \qquad (2.5)$$

The TLSE of $\underset{\sim}{\beta}$ is then defined by the ordinary LSE based on the subset of the Y_i for which $a_{ii} = 1$, $i=1,\dots,n$. We call this estimator as the UTLSE and denote it by

$$\underset{\sim}{L}_n(\alpha_1,\alpha_2) = (\underset{\sim}{X}_n'\underset{\sim}{A}_n\underset{\sim}{X}_n)^-(\underset{\sim}{X}_n'\underset{\sim}{A}_n\underset{\sim}{Y}_n) = (\underset{\sim}{L}_{1n}'(\alpha_1,\alpha_2), \underset{\sim}{L}_{2n}'(\alpha_1,\alpha_2))' , \qquad (2.6)$$

where $\underset{\sim}{L}_{in}(\alpha_1,\alpha_2)$ is a p_i-vector, $i=1,2$, $p_1+p_2 = p+1$. The RTLSE is defined in a similar manner (but working with the restrained model), and is denoted by

$$\underset{\sim}{L}_{1n}^*(\alpha_1,\alpha_2) = (\underset{\sim}{X}_n'\underset{\sim}{A}_n^*\underset{\sim}{X}_n)^-(\underset{\sim}{X}_n'\underset{\sim}{A}_n^*\underset{\sim}{Y}_n) , \qquad (2.7)$$

where $\underset{\sim}{A}_n^* = \text{Diag}(a_{ii}^*, i=1,\dots,n)$ is defined as in (2.5) with the URQ $\tilde{\underset{\sim}{\beta}}(\alpha_j)$ being replaced by the restricted RQ (RRQ), defined as in (2.4) (but with the reduced model, i.e., $\underset{\sim}{x}_j'\underset{\sim}{t}$ being replaced by $\underset{\sim}{x}_{1j}'\underset{\sim}{t}_1$).

Next, we may note that both the PTE and Stein-rule estimator rest on a suitable test statistic for testing the null hypothesis H_0: $\underset{\sim}{\beta}_2 = \underset{\sim}{0}$. For this purpose, we may proceed as in Jurečková (1983), and define

$$\underset{\sim}{Q}_{22.1}^{(n)} = \underset{\sim}{Q}_{n22} - \underset{\sim}{Q}_{n21}\underset{\sim}{Q}_{n11}^-\underset{\sim}{Q}_{n12} \ , \quad \underset{\sim}{Q}_{11.2}^{(n)} = \underset{\sim}{Q}_{n11} - \underset{\sim}{Q}_{n12}\underset{\sim}{Q}_{n22}^-\underset{\sim}{Q}_{n21} \ ; \qquad (2.8)$$

$$s_n^2 = (\alpha_2 - \alpha_1)^{-2}\{ (n-p_2-1)^{-1} S_n^2 + \alpha_1[\tilde{\beta}_o(\alpha_1)-L_{0n}(\alpha_1,\alpha_2)]^2 + (1-\alpha_2)[\tilde{\beta}_o(\alpha_2)-L_{0n}(\alpha_1,\alpha_2)]$$
$$- [\alpha_1\{\beta_o(\alpha_1) - L_{0n}(\alpha_1,\alpha_2)\}+ (1-\alpha_2)\{\beta_o(\alpha_2)- L_{0n}(\alpha_1,\alpha_2)\}]^2 , \qquad (2.9)$$

$$\mathcal{L}_n^2 = (\underset{\sim}{L}_{2n}(\alpha_1,\alpha_2))' (\underset{\sim}{Q}_{22.1}^{(n)})^-(\underset{\sim}{L}_{2n}(\alpha_1,\alpha_2))/s_n^2 , \qquad (2.10)$$

where the URQ are defined as in (2.4), and

$$S_n^2 = \underset{\sim}{Y}_n'\underset{\sim}{A}_n'[\underset{\sim}{I}_n - \underset{\sim}{X}_n(\underset{\sim}{X}_n'\underset{\sim}{A}_n\underset{\sim}{X}_n)^-\underset{\sim}{X}_n]\underset{\sim}{A}_n\underset{\sim}{Y}_n \qquad (2.11)$$

is the usual residual sum of squares computed from the set of observations for which $a_{ii} = 1$, $i=1,\dots,n$. With this adoption of the preliminary test statistic [in (2.10)], we may then formulate the PTTLSE as

$$\underset{\sim}{L}_{1n}^{PT}(\alpha_1,\alpha_2) = \underset{\sim}{L}_{1n}^*(\alpha_1,\alpha_2)I(\mathcal{L}_n^2 \le \ell_{n,\varepsilon}) + \underset{\sim}{L}_{1n}(\alpha_1,\alpha_2)I(\mathcal{L}_n^2 > \ell_{n,\varepsilon}), \qquad (2.12)$$

where $\ell_{n,\varepsilon}$ stands for the critical level of $\mathcal{L}_n^2$ at the significance level ε : $0 < \varepsilon < 1$. Note that for the PTE, we only need that p_1,p_2 are positive integers, not necessarily greater than 2.

A Stein-rule or shrinkage estimator is generally adapted to a given quadratic loss function ; in this specific problem, this calls for a p.d. matrix $\underset{\sim}{W}$ of order $p_1 \times p_1$, and the actual estimator depends on this choice of $\underset{\sim}{W}$. We may define

$$d_n = \text{smallest characteristic root of } (n\underset{\sim}{W}(\underset{\sim}{Q}_{11.2}^{(n)})^{-1}) = ch_{p_1}(n\underset{\sim}{W}(\underset{\sim}{Q}_{11.2}^{(n)})^{-1}), \quad (2.13)$$

and following the lines of Saleh and Sen (1987) [where the case of the classical LSE has been treated in detail], we may consider the following STLSE :

$$\underset{\sim}{L}_{1n}^{S}(\alpha_1,\alpha_2) = \underset{\sim}{L}_{1n}^{*}(\alpha_1,\alpha_2) + [\, \underset{\sim}{I}_{p_1} - c d_n \underset{\sim}{W}^{-1} \underset{\sim}{Q}_{11.2}^{(n)} n^{-1} \mathcal{L}_n^{-2} \,][\, \underset{\sim}{L}_{1n}(\alpha_1,\alpha_2) - \underset{\sim}{L}_{1n}^{*}(\alpha_1,\alpha_2)], \quad (2.14)$$

where

$$0 < c < c(p_1,p_2) \text{ is a positive shrinkage factor.} \quad (2.15)$$

We shall provide bounds for $c(p_1,p_2)$ later on. In passing we may remark that for this linear model problem, the use of the usual Mahalanobis distance for the loss function has been advocated in many cases [viz., Saleh and Sen (1987)], and if the same is done here, we are led to a specific choice of $\underset{\sim}{W} = n^{-1}\underset{\sim}{Q}_{11.2}^{(n)}$, so that d_n in (2.13) then reduces to 1, and (2.14) simplifies to

$$\underset{\sim}{L}_{1n}^{S*}(\alpha_1,\alpha_2) = \underset{\sim}{L}_{1n}^{*}(\alpha_1,\alpha_2) + (1 - c\mathcal{L}_n^{-2})[\underset{\sim}{L}_{1n}(\alpha_1,\alpha_2) - \underset{\sim}{L}_{1n}^{*}(\alpha_1,\alpha_2)] , \quad (2.16)$$

which is another convex combination of the UTLSE and RTLSE. We shall mainly be concerned with the specific STLSE in (2.16) and discuss about the shrinkage factor in (2.15).

3. WHITHER ADR ?

Asymptotic considerations dominate the formulation of the TLSE , and as we shall see in the next section that in this sense, the theory of PTE and STLSE can also be formulated in a very unified manner. On the contrary, the treatment for the small sample case is very much dependent on the underlying F, and we may not be able to draw a picture pertaining to the general pattern. While the justifications for asymptotic considerations in the study of ADR are given in detail in Sen (1984), Sen and Saleh (1985), among other places, we may without much repetitions of their arguments mention the following salient points.

First, we formulate the risk function with a quadratic loss. As has been mentioned in the previous section, we would like to incorporate the Mahalanobis distance in the formulation of the loss function. Specifically, we take the loss in estimating β_1 by an estimator b_n as

$$D(b_n, \beta_1) = [n(b_n - \beta_1)' Q_{11.2}^{(n)}(b_n - \beta_1)], \tag{3.1}$$

so that the risk is given by

$$\rho(b_n, \beta_1) = E\{D(b_n, \beta_1)\} = \mathrm{Tr}(nQ_{11.2}^{(n)} E[(b_n - \beta_1)(b_n - \beta_1)']). \tag{3.2}$$

Even in this simple form, there are some problems connected with the evaluation of this risks for the various versions of the TLSE considered in Section 2. Note that the RQ are non-linear functions of the observations and hence, not only the a_{ii} in (2.5) are stochastic in nature, they are also non-linear functions of the observations. Whereas the asymptotic distribution theory of these RQ and the TLSE, mostly developed in Jurečková (1983, 1984), relate to the usual weak convergence results (or to some weak representations in terms of independent r.v.'s) which may not be strong enough to justify the moment convergence results needed for (3.2). Even if they were justified, there is another basic problem. For the UTLSE, computation of the mean product matrix is simplest, and this matrix is independent of the nuisance parameter β_2 . However, for the RTLSE, this mean product matrix depends on β_2 also (when computed for the full model in (1.1)), and for any (fixed) $\beta_2 \neq 0$, as n increases, the risk in (3.2) becomes indefinitely large (i.e., unbounded); although at $\beta_2 = 0$, this risk is generally smaller than that of the UTLSE. Thus, in an asymptotic setup, for any fixed $\beta_2 \neq 0$, the RTLSE becomes heavily biased and inefficient estimator of β_1 . This suggests that in our setup, (3.2) may depend as well on the unknown β_2 , and also, we need to localize β_2 in the neighbourhood of the pivot so as to obtain a meaningful picture. A similar consideration holds for the PTTLSE in (2.12). For any fixed $\beta_2 \neq 0$, asymptotically, the PTE reduces to the UTLSE if the test based on $\mathcal{L}_n^2$ in (2.10) is consistent (as will be shown in the next section). A similar asymptotic equivalence result holds for the STLSE in (2.16) [or (2.14)]. However, in the case of the STLSE, we have an additional problem. Since $\mathcal{L}_n^{-2}$ appears in the estimator, computation of the second moment of an STLSE may require negative moments of the test statistics. In the normal theory case, by appeal to the usual non-central chi square d.f.'s for the test statistics, these negative moments can be evaluated by making use of the Stein-identities, in the non-normal or non-linear cases, convergence of the negative moments may demand quite restrictive regularity conditions. To eliminate these drawbacks , we prescribe the following :

(i) Confine to a sequence of local alternatives, viz.,

$$H_n : \beta_2 = \beta_{2,n} = n^{-\frac{1}{2}} \lambda_2, \text{ where } \lambda_2 \text{ (fixed)} \in R^{p_2} ; \tag{3.3}$$

note that the null hypothesis case is covered by letting $\lambda_2 = 0$.

(ii) Show that for an estimator b_n of β_1, under consideration, when $\{H_n\}$ holds,

$$\lim_{n\to\infty} P\{ n^{\frac{1}{2}}(b_n - \beta_1) \le x \mid H_n \} = G(x;\lambda_2,Q,\gamma) \text{ exists}, \forall\, x \in R^{p_1} , \tag{3.4}$$

where Q is defined by (2.2) and γ is a suitable scale factor, depending on the d.f. F.

(iii) Instead of the risk in (3.2), use the asymptotic distribution in (3.4) to formulate the asymptotic distributional risk (ADR) as

$$\rho^*(b, \beta_1; \lambda_2 ,Q) = \mathrm{Tr}(Q_{11.2}\{\int\cdots\int xx'dG(x ; \lambda_2,Q,\gamma) \}). \tag{3.5}$$

We shall prescribe the use of this ADR instead of the actual risk computed from (3.2), and we shall see later on that this results in considerable relalaxation of the regularity conditions. On the other hand, we should keep in mind that the comparative results based on (3.5) may not necessarily apply to the actual risk situation when some of these extra regularity conditions may not hold.

4. SOME ASYMPTOTICS ON TLSE AND $\mathcal{L}_n$

As we have discussed in Section 3, for our study of the ADR results, we need to study the consistency of the test based on $\mathcal{L}_n^2$ in (2.10) and also the asymptotic distribution theory of the different versions of the TLSE leading to appropriate forms for (3.4). Towards this, we start with a basic asymptotic result on TLSE discussed in detail in Jurečková (1984). Let

$$\psi_F(E_i) = F^{-1}(\alpha_1) \vee E_i \wedge F^{-1}(\alpha_2) \; ; \; E_i = Y_i - x_i'\beta \; , \; i=1,\ldots,n, \tag{4.1}$$

$$\psi_i = \psi_F(E_i) - \int_{R^1} \psi_F(x)dF(x) \; , \; i=1,\ldots,n, \quad \psi_n =(\psi_1,\ldots, \psi_n)' , \tag{4.2}$$

$$\delta = (\alpha_2 - \alpha_1)^{-1}\int_{\alpha_1}^{\alpha_2} F^{-1}(u)du \; , \text{ and } e_1 = (1,0,\ldots,0)' . \tag{4.3}$$

Then, the following result is due to Jurečková (1984) :

$$n^{\frac{1}{2}}(L_n(\alpha_1,\alpha_2) - \beta - \delta e_1) = n^{-\frac{1}{2}} (\alpha_2 -\alpha_1)^{-1}Q^*X_n'\psi_n + O_p(n^{-\frac{1}{4}}) , \tag{4.4}$$

so that by virtue of (2.2), (2.3) and the central limit theorem , we obtain that

$$n^{\frac{1}{2}}(\underset{\sim}{L}_n(\alpha_1,\alpha_2) - \underset{\sim}{\beta} - \delta \underset{\sim}{e}_1) \sim \mathcal{N}_{p+1}(\underset{\sim}{0} , \sigma^2(F,\alpha_1,\alpha_2)\underset{\sim}{Q}^*) , \qquad (4.5)$$

where

$$\sigma^2(F,\alpha_1,\alpha_2) = (\alpha_2-\alpha_1)^{-2}\{ \int_{\alpha_1}^{\alpha_2}(F^{-1}(u)- \delta)^2 du +\alpha_1(F^{-1}(\alpha_1) -\delta)^2 + (1- \alpha_2)(F^{-1}(\alpha_2)- \delta)^2 - [\alpha_1(F^{-1}(\alpha_1) - \delta)+ (1 -\alpha_2)(F^{-1}(\alpha_2) - \delta)]^2 \}. \qquad (4.6)$$

In this context, we may note that only the estimator of the intercept β_o may be asymptotically biased; the others are all asymptotically unbiased. If, in particular, F is symmetric and we choose $\alpha_2 = 1 - \alpha_1$, then $\delta = 0$, so that asymptotic unbiasedness will be achieved for all the p+1 components of the UTLSE. It also follows from Jurečková (1984) that irrespective of H_0 being true or not,

$$s_n^2 \rightarrow \sigma^2(F,\alpha_1,\alpha_2), \text{ in probability, as } n \rightarrow \infty . \qquad (4.7)$$

Looking back at (2.10), (2.2), (4.7) and the stochastic convergence of $\underset{\sim}{L}_{n2}(\alpha_1,\alpha_2)$ to $\underset{\sim}{\beta}_2$ [ensured by (4.5)] , we immediately conclude that for any fixed $\underset{\sim}{\beta}_2 \in R^{p_2}$,

$$n^{-1}\mathcal{L}_n^2 \xrightarrow{p} (\underset{\sim}{\beta}_2')\underset{\sim}{Q}_{22.1}^{-1}(\underset{\sim}{\beta}_2) \ (> 0, \text{ whenever } \underset{\sim}{\beta}_2 \neq \underset{\sim}{0}), \text{ as } n \rightarrow \infty , \qquad (4.8)$$

while, by (4.5), (2.10), (4.7) and the Cochran theorem on quadratic forms in (asymptotically) normal vectors, we conclude that under H_0: $\underset{\sim}{\beta}_2 = \underset{\sim}{0}$,

$$\mathcal{L}_n^2 \xrightarrow{\mathcal{D}} \chi_{p_2}^2 , \text{ so that } \ell_{n,\varepsilon} \rightarrow \chi_{p_2,\varepsilon}^2 , \text{ as } n \rightarrow \infty , \qquad (4.9)$$

where $\chi_{q,\varepsilon}^2$ is the upper $100\varepsilon\%$ point of the central chi square d.f. with q degrees of freedom (DF). From (4.8) and (4.9), we readily conclude that the test based on $\mathcal{L}_n^2$ in (2.10) is consistent against any $\underset{\sim}{\beta}_2 \neq \underset{\sim}{0}$. We denote by $H_q(x;\Delta)$ the noncentral chi square d.f. with q DF and noncenrtrality parameter Δ , $x \in R^+$. Then, by virtue of (2.10), (4.5), (4.7) and the Cochran theorem, we obtain that

$$\lim_{n\to\infty} P\{ \mathcal{L}_n^2 \leq x \mid H_n \} = H_{p_2}(x ;\Delta) , \ x \in R^+ , \qquad (4.10)$$

where

$$\Delta = (\underset{\sim}{\lambda}_2'\underset{\sim}{Q}_{22.1} \underset{\sim}{\lambda}_2)/\sigma^2(F,\alpha_1,\alpha_2) . \qquad (4.11)$$

Other distributional results will be presented in the next section.

5. ADR RESULTS FOR THE TLSE

In the sequel, we shall take $\alpha_1 = 1-\alpha_2 =\alpha : 0 < \alpha<\frac{1}{2}$, and assume that F is symmetric about 0, so that δ ,defined by (4.3) is equal to 0. Also, to simplify notations, instead of (α_1,α_2), we shall use α , in the statistics as well as in

(4.6) and elsewhere. Then, we have the following

THEOREM 5.1. *Under the assumed regularity conditions, the following holds:*

(i) $\lim_{n\to\infty} P\{n^{\frac{1}{2}}(\underset{\sim}{L}_{1n}(\alpha) - \underset{\sim}{\beta}_1) \le \underset{\sim}{x} \mid H_n\} = G_{p_1}(\underset{\sim}{x};\ \underset{\sim}{0}, \sigma^2(F,\alpha).\underset{\sim}{Q}^*_{11})$, (5.1)

(ii) $\lim_{n\to\infty} P\{n^{\frac{1}{2}}(\underset{\sim}{L}^*_{1n}(\alpha) - \underset{\sim}{\beta}_1) \le \underset{\sim}{x} \mid H_n\} = G_{p_1}(\underset{\sim}{x} + \underset{\sim}{Q}^{-1}_{11}\underset{\sim}{Q}_{12}\underset{\sim}{\lambda}_2; \underset{\sim}{0}, \sigma^2(F,\alpha)\underset{\sim}{Q}^*_{11.2})$, (5.2)

(iii) $\lim_{n\to\infty} P\{n^{\frac{1}{2}}(\underset{\sim}{L}^{PT}_{1n}(\alpha) - \underset{\sim}{\beta}_1) \le \underset{\sim}{x} \mid H_n\} = H_{p_2}(\chi^2_{p_2,\varepsilon};\Delta)\, G_{p_1}(\underset{\sim}{x} + \underset{\sim}{Q}_{12}\underset{\sim}{Q}^{-1}_{22}\underset{\sim}{\lambda}_2; \underset{\sim}{0}, \sigma^2(F,\alpha)\underset{\sim}{Q}^*_{11.2})$

$$+ \int_{E(\lambda_2)} G_{p_1}(\underset{\sim}{x} - \underset{\sim}{Q}^*_{12}\underset{\sim}{Q}^{*-1}_{22}\underset{\sim}{z}; \underset{\sim}{0}, \sigma^2(F,\alpha)\underset{\sim}{Q}^*_{11.2})\, dG_{p_2}(\underset{\sim}{z};\ \underset{\sim}{0}, \sigma^2(F,\alpha)\underset{\sim}{Q}^*_{22}), \quad (5.3)$$

where

$$E(\underset{\sim}{\lambda}_2) = \{\ \underset{\sim}{z} : (\underset{\sim}{z} + \underset{\sim}{\lambda}_2)'\underset{\sim}{Q}_{22.1}(\ \underset{\sim}{z} + \underset{\sim}{\lambda}_2) \ge \chi^2_{p_2,\varepsilon}.\sigma^2(F,\alpha)\ \}, \quad (5.4)$$

and $G_q(x;\ \underset{\sim}{\mu}, \underset{\sim}{\Sigma})$ *stands for the q-variate normal d.f. with mean vector* $\underset{\sim}{\mu}$ *and dispersion matrix* $\underset{\sim}{\Sigma}$.

Note that (5.1) follows directly from (4.5), while (5.2) follows from the asymptotic representation theorem for TLSE [viz., Jurečková (1984)]. The rest of the proof of this theorem is quite similar to that of Saleh and Sen (1987) [dealing with the ordinary LSE], and hence, the details are omitted. Moreover, using (4.5) and virtually repeating the proof of Theorem 3.2 of Saleh and Sen (1987), we arrive at the following.

THEOREM 5.2. *Under the same regularity conditions as in Theorem 5.1,*

$$n^{\frac{1}{2}}(\ \underset{\sim}{L}^{S*}_{1n}(\alpha) - \underset{\sim}{\beta}_1\ ) \xrightarrow{\mathfrak{D}} \underset{\sim}{Q}^*_1\underset{\sim}{U} + c\left\{\frac{\underset{\sim}{Q}^{-1}_{11}\underset{\sim}{Q}_{12}(\ \underset{\sim}{Q}^*_1\underset{\sim}{U} + \underset{\sim}{\lambda}_2)}{[(\underset{\sim}{Q}^*_2\underset{\sim}{U} + \underset{\sim}{\lambda}_2)'\underset{\sim}{Q}_{22}(\ \underset{\sim}{Q}^*_2\underset{\sim}{U} + \underset{\sim}{\lambda}_2)]/\sigma^2(F,\alpha)}\right\}, \quad (5.5)$$

where

$$\underset{\sim}{U} \sim \mathcal{N}_{p+1}(\underset{\sim}{0},\ \sigma^2(F,\alpha)\underset{\sim}{Q})\ \textit{and}\ \underset{\sim}{Q}^* = (\ \underset{\sim}{Q}^{*\prime}_1,\ \underset{\sim}{Q}^{*\prime}_2\ )' = \underset{\sim}{Q}^{-1}. \quad (5.6)$$

Now, Theorems 5.1 and 5.2 provide the details for the first two steps in (3.3) and (3.4), so that we may then use (3.5) to compute the desired ADR results. Note that by definition in (3.5) and Theorem 5.1,

$$\rho^*(\underset{\sim}{L}_1(\alpha), \underset{\sim}{\beta}_1,\ \underset{\sim}{\lambda}_2,\ \underset{\sim}{Q}) = p_1\sigma^2(F,\alpha) \quad [\text{ as } \underset{\sim}{Q}^*_{11} = \underset{\sim}{Q}^{-1}_{11.2}\ ],\quad \underset{\sim}{\lambda}_2 \in R^{p_2}. \quad (5.7)$$

Thus, the UTLSE has a constant ADR over $\underset{\sim}{\lambda}_2$ (in any compact subspace in R^{p_2}). In a similar manner, we obtain that

$$\rho^*(\underset{\sim}{L}^*_1(\alpha), \underset{\sim}{\beta}_1, \underset{\sim}{\lambda}_2, \underset{\sim}{Q}) = \sigma^2(F,\alpha)\mathrm{Tr}(\underset{\sim}{Q}_{11.2}\underset{\sim}{Q}^{-1}_{11}) + (\underset{\sim}{\lambda}'_2\underset{\sim}{M}\underset{\sim}{\lambda}_2), \quad (5.8)$$

where

$$\underset{\sim}{M} = \underset{\sim}{Q}_{21}\underset{\sim}{Q}_{11}^{-1}\underset{\sim}{Q}_{11.2}\underset{\sim}{Q}_{11}^{-1}\underset{\sim}{Q}_{12} \ . \tag{5.9}$$

Also,

$$\rho^{*}(\underset{\sim}{L}_1^{PT}(\alpha),\underset{\sim}{\beta}_1,\underset{\sim}{\lambda}_2,\underset{\sim}{Q}) = p_1\sigma^2(F,\alpha) + (\underset{\sim}{\lambda}_2'\underset{\sim}{M}\underset{\sim}{\lambda}_2)[2H_{p_2+2}(\chi^2_{p_2,\varepsilon};\Delta) - H_{p_2+4}(\chi^2_{p_2,\varepsilon};\Delta)] \\ -\sigma^2(F,\alpha)\, H_{p_2+2}(\chi^2_{p_2,\varepsilon};\Delta)\{p_1 - \mathrm{Tr}(\underset{\sim}{Q}_{11.2}\underset{\sim}{Q}_{11}^{-1})\} , \tag{5.10}$$

where Δ is defined by (4.11). Finally, from Theorem 5.2 and (3.5), we obtain that

$$\rho^{*}(\underset{\sim}{L}_1^{S*}(\alpha),\underset{\sim}{\beta}_1,\underset{\sim}{\lambda}_2,\underset{\sim}{Q}) = p_1\sigma^2(F,\alpha) - c\sigma^2(F,\alpha)\{2\mathrm{Tr}(\underset{\sim}{M}^*)E(\chi^{-2}_{p_2+2}(\Delta)) - c\mathrm{Tr}(\underset{\sim}{M}^*)E(\chi^{-4}_{p_2+2}(\Delta)) \\ - (c+4)(\underset{\sim}{\lambda}_2'\underset{\sim}{M}\underset{\sim}{\lambda}_2)E(\chi^{-4}_{p_2+4}(\Delta))/\sigma^2(F,\alpha)\ \}, \tag{5.11}$$

where

$$\underset{\sim}{M}^* = \underset{\sim}{Q}_{22.1}^{-1/2}\underset{\sim}{M}\underset{\sim}{Q}_{22.1}^{-1/2} \quad , \quad \chi_q^{-2r}(\Delta) = \{\chi_q^2(\Delta)\}^{-r}, \tag{5.12}$$

and $\chi_q^2(\Delta)$ has the d.f. $H_q(x;\Delta)$. The derivations of these results [given the two theorems on thier asymptotic distributions] are very similar to the case of the usual LSE, treated in detail in Saleh and Sen (1987), and hence, we omit these details here.

Having obtained these ADR results, our next task is to compare the estimators UTLSE, RTLSE, PTTLSE and STLSE in the light of their ADR. This picture is also very much similar to the case of the classical LSE ; the only difference is that the usual variance $\sigma^2(F)$ of the d.f. F is replaced here by $\sigma^2(F,\alpha)$. Thus, the relative picture of (5.7) and (5.8) remains the same as in the case of the LSE, excepting that there the noncentrality parameter had a divisor $\sigma^2(F)$, where as here we have $\sigma^2(F,\alpha)$. It follows from Bickel and Lehmann (1975) that if $\mathcal{F}_1$ stands for the class of symmetric (absolutely continuous)distributions on R^1, then for $\alpha \in [0.05,0.10], F \in \mathcal{F}_1$, $\sigma^2(F)/\sigma^2(F,\alpha)$ is bounded from below by $(1-2\alpha)^2$, whereas the upper bound can be indefinitely large. Since we have for our TLSE,

$$(5.8)/(5.7) = p_1^{-1}\mathrm{Tr}(\underset{\sim}{Q}_{11.2}\underset{\sim}{Q}_{11}^{-1}) + p_1^{-1}(\underset{\sim}{\lambda}_2'\underset{\sim}{M}\underset{\sim}{\lambda}_2)/\sigma^2(F,\alpha) , \tag{5.13}$$

we conclude that the smaller is the value of $\sigma(F,\alpha)$ compared to $\sigma(F)$, the larger is the second term on the right hand side of (5.13) [compared to the case of the LSE], so that the faster will be the increase of the relative risk of the RTLSE (with respect to the UTLSE). This explains the lack of robustness (in terms of the ADR) of the RTLSE with small departures from the pivot; the picture may even be worse when the UTLSE is more efficient than the ordinary LSE. At the pivot, of course, the RTLSE is generally better than the UTLSE (unless $\underset{\sim}{Q}_{11.2} = \underset{\sim}{Q}_{11}$, i.e., $\underset{\sim}{Q}_{12}$ is a null matrix). Thus, though the TLSE may be preferred on the ground of efficiency and robustness against heavy tailed distributions, in this greater domain, the RTLSE may not turn out to be a good competitor of the UTLSE , although none dominates the other in the light of their ADR. If we compare the PTTLSE with the UTLSE, we get a picture similar to the case of the usual LSE [viz., Saleh and Sen (1987)] with the only change that in the expression for the ADR (and Δ), we have $\sigma^2(F,\alpha)$ instead of $\sigma^2(F)$. Thus, here also, at the pivot, the PTTLSE has usually a smaller ADR than the UTLSE (larger ADR than the RTLSE), the relative ADR first increases as $\underset{\sim}{\lambda}_2$ moves away from $\underset{\sim}{0}$, then after attaining a maximum (greater than 1) , it continues to stay above the line 1 and approaches the upper asymptote 1 as $||\underset{\sim}{\lambda}_2|| \to \infty$. Thus, the PTTLSE does not have an unbounded ADR (like the RTLSE) although it does not dominate the RTLSE or the UTLSE. The picture of the ADR of the PTTLSE reveals a lot of robustness aspect, and it only requires that $\underset{\sim}{Q}_{12}$ is non-null.

For the ADR of the STLSE, we have so far taken the shrinkage factor c to be arbitrary. Let us comment on this choice , so that the STLSE has some good ADR properties. If we compute (5.11) at the pivot (i.e., $\underset{\sim}{\lambda}_2 = \underset{\sim}{0}$), we obtain that it would be smaller than (5.7) if $2E(\chi^{-2}_{p_2+2}(0)) > cE(\chi^{-4}_{p_2+2}(0))$. Now , we know that $E(\chi_q^{-2}(0)) = (q-2)^{-1}$ and $E(\chi_q^{-4}(0)) = \{(q-2)(q-4)\}^{-1}$. Thus, in order that at $\underset{\sim}{\lambda}_2 = \underset{\sim}{0}$, (5.11) is smaller than (5.7), we need that $0 < c < 2(p_2-2)$ (which in turn demands that $p_2 > 2$). However, this simple condition on c may not be enough to ensure the asymptotic dominance of the STLSE over UTLSE. Towards this, we define

$$\underset{\sim}{M}^o = \underset{\sim}{Q}_{12}\underset{\sim}{Q}_{22}^{-1}\underset{\sim}{Q}_{21}\underset{\sim}{Q}_{11}^{-1} \quad \text{and } h = ch_1(\underset{\sim}{M}^o)/Tr(\underset{\sim}{M}^o) , \tag{5.14}$$

so that $0 < h \leq 1$. Note that if $\underset{\sim}{Q}_{12}$ is of rank 1, then $ch_1(\underset{\sim}{M}^o) > 0$, but the other characteristic roots are all equal to 0, so that $h = 1$. Then proceeding as in the case of the usual LSE [viz., Saleh and Sen (1987)], we conclude that in the light of the ADR, the STLSE dominates the UTLSE if the following Condition holds :

$$0 < c < 2(p_2 - 2) \quad \text{and} \quad h(c+4) \leq 2 \ . \tag{5.15}$$

In order that c satisfies both the inequalities (and is positive), we need that $h < 1/2$, which in turn , requires that $\underset{\sim}{M}^o$ has rank at least equal to 3 (i.e., both p_1 and p_2 are greater than 2). Actually, we may note that

$$\begin{aligned} Tr(\underset{\sim}{M}^o) &= Tr(\underset{\sim}{Q}_{22}^{-1}\underset{\sim}{Q}_{21}\underset{\sim}{Q}_{11}^{-1}\underset{\sim}{Q}_{12}) = Tr(\underset{\sim}{Q}_{22}^{-1}[\ \underset{\sim}{Q}_{22} - \underset{\sim}{Q}_{22.1}]) = Tr(\underset{\sim}{I}_{p_2} - \underset{\sim}{Q}_{22}^{-1}\underset{\sim}{Q}_{22.1}) \\ &= Tr(\underset{\sim}{I}_{p_1} - \underset{\sim}{Q}_{11}^{-1}\underset{\sim}{Q}_{11.2}) \leq p^* = \min(p_1, p_2) \ ; \end{aligned} \tag{5.16}$$

$$ch_1(\underset{\sim}{M}^o) = 1 - ch_{p_1}(\underset{\sim}{Q}_{11.2}\underset{\sim}{Q}_{11}^{-1}) \ [\text{ or } 1 - ch_{p_2}(\underset{\sim}{Q}_{22.1}\underset{\sim}{Q}_{22}^{-1})\ ] \ , \tag{5.17}$$

so that the condition that $h < 1/2$ may directly be verified by computing the characteristic roots of $\underset{\sim}{Q}_{11}^{-1}\underset{\sim}{Q}_{11.2}$ (or of $\underset{\sim}{Q}_{22}^{-1}\underset{\sim}{Q}_{22.1}$). In any case, for this STLSE to dominate the UTLSE (in the light of their ADR), we need that $p^* \geq 3$, and moreover, h to be less than 1/2. Compared to the case of the PTTLSE, here the conditions on p_1, p_2 and h are more restrictive, and there may be some cases where the PTTLSE works out very conveniently, while the STLSE may fail to yield any dominance result. Actually, comparing (5.10) and (5.11), we may remark that even if (5.15) holds, at the pivot $\underset{\sim}{\lambda}_2 = \underset{\sim}{0}$, (5.11) is generally larger than (5.10) [unless, ε, the significance level of the preliminary test on the pivot, is large], so that the STLSE may not generally dominate the PTTLSE. However, the STLSE , under (5.15), is asymptotically minimax (in the light of the ADR), while the PTTLSE is not so. Thus, in making a choice between the PTTLSE and STLSE, we need to take into account the design matrix $\underset{\sim}{Q}$, the specific values of p_1 and p_2 , and the asymptotic minimax considerations. Categorically, we may not be able to advocate the use of either one of them in all cases, although both of them fare well comared to the UTLSE and RTLSE. In this context, we may remark that the smaller is the ratio $\sigma^2(F,\alpha)/\sigma^2(F)$, the lesser will be the relative gain in the improvent due to PTE or shrinkage in the case of the TLSE over the case of the ordinary LSE. Thus, in one hand, in order to achieve more robustness and efficiency, one may use the TLSE instead of the LSE (in linear models) , but in that case, the relative improvement in (5.10) or (5.11) over (5.7) may be smaller compared to the case of the LSE.

For improved estimation of the mean vector of a multivariate normal distribution with known or unknown dispersion matrix, a modification of the Stein-rule estimator (called the **positive-rule estimator** (PRE)) has been considered, and this PRE has usually a smaller risk than the Stein-rule estimator. Tempted by this, we are naturally led to consider a PRE version of (2.16). We define a **positive rule** TLSE (PRTLSE) [under a Mahalanobis distance type loss function] as

$$\underset{\sim}{L}_{1n}^{S+}(\alpha_1,\alpha_2) = \underset{\sim}{L}_{1n}^{*}(\alpha_1,\alpha_2) + (1 - c\,\underset{\sim}{\mathcal{L}}_n^{-2})^{+}[\,\underset{\sim}{L}_{1n}(\alpha_1,\alpha_2) - \underset{\sim}{L}_{1n}^{*}(\alpha_1,\alpha_2)\,], \qquad (5.18)$$

where

$$(1 - c\,\underset{\sim}{\mathcal{L}}_n^{-2})^{+} = \max\{\,0,\ 1 - c\,\underset{\sim}{\mathcal{L}}_n^{-2}\,\}, \qquad (5.19)$$

and the other notations are all adapted from Section 2. This PRTLSE differs from the STLSE only on the set where $\underset{\sim}{\mathcal{L}}_n^{2} < c$, and in this sense, the PRTLSE overshrinks the estimator towards the RTLSE. By virtue of the asymptotic representation in (4.4), we are able to reduce this problem in an asymptotic setup to the standard multinormal case where the existing results [viz., Corollary 5.3.1 of Anderson (1984)] can be readily used to show that under $\{H_n\}$ in (3.3) and the assumed regularity conditions, the ADR of the PRTLSE in (5.18) is smaller than that of the STLSE when (5.15) holds. Thus, the PRTLSE dominates the STLSE in the light of their ADR. However, in this context, it may be remarked that (5.18) may not agree with (2.12) even if we let $c = \ell_{n,\varepsilon}$, and generally, for small values of ε , the PRTLSE may not dominate the PTTLSE. Moreover, if instead of the simplified form in (2.16), we take an arbitrary quadratic loss and define the Stein-rule estimator as in (2.14), then the corresponding positive-rule version may not generally enjoy this small ADR property (over (2.14)).However, given that in linear models, this Mahalanobis type loss function is quite appropriate, we are naturally to advocate the use of the PRTLSE instead of the STLSE.

ACKNOWLEDGEMENTS

This research was partially supported by the NSERC Grant No. A3088 (Canada) and National Heart, Lung and Blood Institute (U.S.), Contract No. NIH-NHLBI-71-2243-L from the National Institutes of Health.

REFERENCES

ANDERSON, T.W. (1984). *An Introduction to Multivariate Statistical Analysis*, 2nd Ed. John Wiley, New York.

BICKEL, P.J. AND LEHMANN, E.L. (1975). Descriptive statistics for nonparametric model, II. *Ann. Statist. 3*, 1045-1069.

JUREČKOVÁ, J. (1983a). Winsorized least squares estimator and its M-estimator counterpart. In *Contributions to Statistics: Essays in Honour of Norman L. Johnson* (ed: P.K.Sen), North Holland, Amsterdam, pp. 237-245.

JUREČKOVÁ, J. (1983b). Robust estimators of location and regression parameters and their second order asymptotic relations. *Proc.9th Prague Confer. Infor.Theor. Ran. Proc. Statist. Dec. Func.* (ed: J.A. Visek), Reidel, Dordrecht, 19-32.

JUREČKOVÁ, J. (1984). M-, L- and R-estimators. In *Handbook of Statistics, Volume 4: Nonparametric Methods* (eds: P.R.Krishnaiah and P.K.Sen), North Holland, Amsterdam pp. 463-485.

JUREČKOVÁ, J. (1985). Robust estimators of location and their second order asymptotic relations. *A Celebration of Statistics: The ISI Centenary Volume* (Eds. A.C. Atkinson and S.E. Fienberg), Springer-Verlag, New York, pp. 377-392.

JUREČKOVÁ, J. AND SEN, P.K. (1984). On adaptive scale-equivariant M-estimators in linear models. *Statist. Dec. 2*, *Suppl.*1, pp.31-46

KOENKER, R.W. AND BASSETT, G.W. (1978). Regression quantiles. *Econometrica 46*, 33-50.

RUPPERT, D. AND CARROLL, R.J. (1980). Trimmed least squares estimation in the linear model. *Jour. Amer. Statist. Assoc. 75*, 828-838.

SALEH, A.K.M.E. AND SEN, P.K. (1984a). Least squares and rank order preliminary test estimation in general linear models. In *Statistics: Applications and New Directions* (Eds. J.K.Ghosh and J.Roy), Statist.Publ. Soc., Calcutta, pp. 237-253.

SALEH, A.K.M.E. AND SEN, P.K. (1984b). Nonparametric Preliminary Test Inference. In *Handbook of Statistics, Volume 4: Nonparametirc Methods* (Eds. P.R. Krishnaiah and P.K.Sen), North Holland, Amsterdam, pp. 275-297.

SALEH, A.K.M.E. AND SEN, P.K. (1986). On shrinkage R-estimation in a multiple regression model. *Commun. Statist., Theor. Meth. 15*, 2229-2244.

SALEH, A.K.M.E. AND SEN, P.K. (1987). Relative performance of Stein-rule and preliminary test testimators in linear models: Least squares theory. *Commun. Statist. Theor. Meth. 16*, 461-476.

SEN, P.K. (1984). A James-Stein detour of U-statistics. *Commun. Statist. Theor. Meth. A 13*, 2725-2747.

SEN, P.K. AND SALEH, A.K.M.E. (1985). On some shrinkage estimators of multivariate locations. *Ann. Statist. 13*, 272-281.

SEN, P.K. AND SALEH, A.K.M.E. (1987). On preliminary test and shrinkage M-estimation in linear models. *Ann. Statist.* 15, in press.