One-Step R-Estimation in Linear Models with Stable Errors

Marc Hallin[a,b], Yvik Swan[c], Thomas Verdebout[d] and David Veredas[a]
[a] *Institut de Recherche en Statistique, ECARES, and ECORE, Université libre de Bruxelles, Belgium.*
[b] *Académie Royale de Belgique, ORFE, Princeton University, and CentER, Tilburg University.*
[c] *Département de Mathématique, Université Libre de Bruxelles, Belgium.*
[d] *EQUIPPE-GREMARS, Université Lille Nord de France.*

Abstract

Classical estimation techniques for linear models either are inconsistent, or perform rather poorly, under α-stable error densities; most of them are not even rate-optimal. In this paper, we propose an original one-step R-estimation method and investigate its asymptotic performances under stable densities. Contrary to traditional least squares, the proposed R-estimators remain root-n consistent (the optimal rate) under the whole family of stable distributions, irrespective of their asymmetry and tail index. While stable-likelihood estimation, due to the absence of a closed form for stable densities, is quite cumbersome, our method allows us to construct estimators reaching the parametric efficiency bounds associated with any prescribed values$(\alpha_0,\ b_0)$ of the tail index α and skewness parameter b, while preserving root-n consistency under any $(\alpha,\ b)$ as well as under light-tailed densities. The method furthermore avoids all forms of multidimensional argmin computation. Simulations confirm its excellent finite-sample performances.

Key words: Stable distributions, local asymptotic normality, LAD estimation, R-estimation, asymptotic relative efficiency.

1. Introduction.

Evidence of heavy-tailed behavior and infinite variances in economics and, even more so, in finance and insurance, is overwhelming. In such context, the Gauss-Markov theorem for linear regression[1] no longer holds true, and the usual OLS estimators of regression coefficients lose their theoretical justifications. Much worse: they also lose their traditional[2] their root-n consistency rates. OLS estimators under stable errors thus are not even rate-optimal: Proposition 3.1 in Hallin, Swan, Verdebout and Veredas (2010) indeed establishes the local asymptotic normality, with root-n consistency rates, of linear models with stable errors, irrespective of their tail index and skewness parameter.

This disturbing fact is by no means a new finding: see Wise (1966) or Blattberg and Sargent (1971) for early discussion. Since then, the asymptotic behavior of estimators in linear models with infinite variance and, more specifically, in models with (non Gaussian) stable errors, has attracted much interest, and several alternatives to OLS estimation have been proposed. Those alternative estimators however either suffer from major consistency problems, or are strictly inefficient and can be improved: see Section 1.1 for a brief review. The objective of this paper is to show how *one-step R-estimation* allows for a tractable and quite substantial rate-optimal improvement.

1.1. Regression parameter estimation under stable errors.

Before turning to R-estimation methods, let us briefly explain why classical estimation methods fail to provide fully satisfactory solutions.

[1] Recall that the Gauss-Markov theorem establishes, for errors with finite variance, that OLS estimators are *best linear unbiased* estimators.

[2] Under the classical condition that the regression constants satisfy Assumption (A1) below—an assumption we tacitly make throughout this section.

(a) *OLS estimators.* As already mentioned, the main trouble with OLS estimators is that their consistency rate depends on the tail index α. This follows from the general results by Samorodnitsky *et al.* (2007) on a class of linear unbiased estimators (see point (c) below). That rate is strictly less than the optimal root-n rate, which is a severe drawback. Moreover, the related asymptotic confidence regions and Wald tests cannot be constructed without estimating α itself.

(b) *Stable MLEs.* OLS estimators are the maximum likelihood estimators (MLEs) associated with Gaussian likelihoods; better performances can be expected from stable likelihoods (involving the four parameters of stable densities along with the regression coefficients of interest). A pioneering result by DuMouchel (1973), indeed, shows that, somewhat surprisingly, stable MLEs (for location, scale, the tail index α, and the skewness parameter b) yield a very standard asymptotically normal behavior, with traditional root-n rates. This result easily extends to the regression case[3]. Practical implementation, of course, runs into the problem that non Gaussian stable densities, hence stable likelihoods, cannot be expressed in closed form. For specified tail index α and skewness parameter b, this is not an obstacle anymore thanks to the computationally efficient integral approximations obtained by Zolotarev (1986, 1995), Nolan (1997, 1999) and several others. But in practice, the tail index and the skewness parameter also have to be estimated; the information matrix, moreover, is not block-diagonal (see DuMouchel (1975)), so that the estimators $\hat{\alpha}$ and $\hat{b}$ of α and b cannot simply be plugged into the information matrix when confidence regions or Wald tests are to be constructed for the regression parameters. Although asymptotically optimal, stable-likelihood-based inference in practice thus seems difficult.

(c) *Linear unbiased estimators.* A broad class of linear unbiased estimators, of which OLS estimators are a particular case, has been considered by Samorodnitsky *et al.* (2007), who also provide a quite complete and systematic picture[4] of their asymptotic behavior. Consistency rates, as a rule, crucially depend on the tail index α of the underlying noise, and are strictly less than the optimal root-n ones; asymptotic covariances depend on α as well. All the drawbacks of OLS estimation thus also are present here. The BLUαN (best linear unbiased estimator, relative to some adequate α-norm—limited to $1 < \alpha < 2$) estimators considered in El Barmi and Nelson (1997) suffers from the same problems.

(d) *LAD estimators.* The bad performances of L_2 estimators (OLS) considerably reinforce the attractiveness of the L_1 approach. The so-called LAD (Least Absolute Deviations) estimators (a particular case of more general quantile regression estimators in the Bassett and Koenker (1978) style) indeed, irrespective of the tail index α, achieve (under Assumption (A1)) root-n consistency. The asymptotic properties of LAD estimators in regression models have been studied intensively: see Bassett and Koenker (1978) for the standard case, Knight (1998) or El Bantli and Hallin (1999) for more general results. Contrary to stable MLEs, BLUEs and OLS estimators, the LAD ones, thus, achieve rate-optimal consistency. Constructing the related confidence regions and Wald tests is possible via classical techniques, without any estimation of α. These advantages of LAD estimation in the stable context have been emphasized as early as 1971 by Fama and Roll (1971). On the other hand, LAD estimators, which are optimal under light-tailed double-exponential noise, cannot be efficient under any heavy-tailed stable density. The objective of this paper is to show how LAD estimators can be improved, often quite substantially, without specifying nor estimating the tail index α.

1.2. R-estimation under stable errors.

Estimation methods based on ranks—in short, R-estimation—go back to Hodges and Lehmann (1963), who provide R-estimators for one-sample and two-sample location models (under symmetric distributions,

[3]The situation is quite different for autoregressive and ARMA models (local experiments are no longer of the LAN type), with $n^{1/\alpha}$ consistency rates under tail index α, and convergence in distribution to the maximizer of a random function; see Andrews *et al.* (2009) for recent results in that context.

[4]Under very general assumptions on the asymptotic behavior of the regression constants (more general than Assumptions (A1) and (A2) below), but assuming *symmetric* heavy-tailed errors—an assumption we do not make here.

for the one-sample case), based on the Wilcoxon and van der Waerden (signed) rank statistic. Since then, the technique has been used in a variety of problems, including K-sample location, regression and analysis of variance, time series analysis and elliptical families—see, e.g., Lehmann (1963), Sen (1966), Jurečková (1971), Koul (1971), Jurečková and Sen (1996), Koul and Saleh (1993), Allal *et al.* (2001), Koul (2002), Hallin *et al.* (2006), and many others.

Ranks naturally appear as maximal invariants in semiparametric models where the density f of some unobservable noise constitutes the infinite-dimensional nuisance. Under classical Argmin form, the Hodges-Lehmann or R-estimator $\underset{\sim}{\boldsymbol{\vartheta}}{}^{(n)}_{\rm HL}$ of a parameter $\boldsymbol{\vartheta}$ is defined as

$$\underset{\sim}{\boldsymbol{\vartheta}}{}^{(n)}_{\rm HL} := \text{argmin}_{\mathbf{t}\in\mathbb{R}^K} \left| \underset{\sim}{Q}{}^{(n)}(\mathbf{R}^{(n)}(\mathbf{t}))\right|, \tag{1.1}$$

where $\underset{\sim}{Q}{}^{(n)}(\mathbf{R}^{(n)}(\boldsymbol{\vartheta}_0))$ is a (signed)-rank test statistic for the null hypothesis $\mathcal{H}_0 : \boldsymbol{\vartheta} = \boldsymbol{\vartheta}_0$ (two-sided test). The main advantage of $\underset{\sim}{\boldsymbol{\vartheta}}{}^{(n)}_{\rm HL}$ over more usual M-estimators follows from the fact that (under parameter value $\boldsymbol{\vartheta}$ and error density f, and standard root-n consistency conditions), $n^{1/2}(\underset{\sim}{\boldsymbol{\vartheta}}{}^{(n)}_{\rm HL} - \boldsymbol{\vartheta})$ is asymptotically equivalent to a function which depends on the unknown actual density f but is measurable with respect to the ranks $\mathbf{R}^{(n)}(\boldsymbol{\vartheta})$ of the unobservable noise. The asymptotic relative efficiencies (AREs) of the R-estimator $\underset{\sim}{\boldsymbol{\vartheta}}{}^{(n)}_{\rm HL}$ defined in (1.1) with respect to other R-estimators, or with respect to its Gaussian competitor (OLS or Gaussian MLE, whenever the latter are root-n consistent) are the same as the AREs of the corresponding rank tests with respect to their Gaussian competitors.[5]

The Argmin form (1.1), however, is computationally inconvenient—particularly so in the case of a relatively high-dimensional parameter $\boldsymbol{\vartheta}$. Inspired by Le Cam's one-step estimation method, Hallin *et al.* (2006), in the context of R-estimation of shape matrices in elliptical families, therefore introduced a one-step form of R-estimation. That method, contrary to (1.1), avoids the computational inconvenience of minimizing, over a possibly high-dimensional parameter space, a piecewise constant function of the form $|\underset{\sim}{Q}{}^{(n)}(\mathbf{R}^{(n)}(\mathbf{t}))|$; moreover it also provides, as a by-product, the asymptotic covariance matrix of the R-estimator. On the other hand, one-step methods require the existence of a preliminary rate-optimal consistent (here, root-n consistent) estimator. This role will be played, in the present context, by the LAD estimator, the only one in the existing literature enjoying the required consistency properties. Our R-estimators thus appear as a one-step improvements over the LAD estimators; they yield the same collection of ARE values as the corresponding rank-based tests, the values of which were obtained in Hallin *et al.* (2010).

In this paper, we explain how that one-step method can be implemented for the estimation of the regression parameter of a general linear model with stable errors, and we study the asymptotic performances of the resulting R-estimators . Those R-estimators rely on a rank-based version of Le Cam's one-step methodology which bypasses the nonparametric estimation of cross-information quantities. They are asymptotically normal under any stable density (with standard root-n rate), and efficient at some prespecified stable density $f_{\boldsymbol{\theta}}$. They exhibit the same asymptotic relative efficiencies as the rank-based tests studied in Hallin *et al.* (2010). For specific scores, they outperform LAD estimators, and hence all valid and tractable estimation methods proposed in the literature. In particular, when based on certain stable scores, such as the score associated with the symmetric stable distribution with tail parameter $\alpha = 1.4$ (see Figure 2), they dominate the LAD under any stable distribution with $\alpha \in (1, 2)$. The computational advantages of one-step R-estimators over the more classical Argmin ones lie in the fact that the K-dimensional minimization (1.1) of a non convex piecewise constant rank-based objective function is replaced by the minimization of a continuous, strictly convex L_1 criterion (yielding the preliminary LAD estimator), followed by a one-dimensional optimization problem; the LAD estimator moreover can be obtained exactly as the solution of a linear programming problem. Table 3 below provides numerical evidence of the quite substantial advantages (in terms of bias and mean squared error) of one-step R-estimation over its classical Argmin counterpart.

[5] Since Gaussian methods are generally invalid under stable error densities, AREs in the sequel are taken with respect to double-exponential likelihood procedures, that is, least absolute deviation (LAD) estimators and the regression version of sign tests (the Laplace rank tests).

2. R-estimation of regression coefficients.

2.1. Asymptotics for linear models with stable errors.

The family of α-stable densities is a four-parameter family

$$\Big\{ f_{\boldsymbol{\theta}} = f_{\alpha,b,\gamma,\delta} |\ \boldsymbol{\theta} := (\alpha, b, \gamma, \delta)' \in \boldsymbol{\Theta} = (0,2] \times [-1,1] \times \mathbb{R}^{+} \times \mathbb{R} \Big\}.$$

Writing $f_{\alpha,b}$ for $f_{\alpha,b,1,0}$, we have

$$f_{\alpha,b,\gamma,\delta}(x) = \frac{1}{\gamma} f_{\alpha,b}\left(\frac{x-\delta}{\gamma}\right), \tag{2.2}$$

which characterizes the roles of δ and γ as location and scale parameters, respectively, and that of $f_{\alpha,b}$ as the standardized version of $f_{\alpha,b,\gamma,\delta}$. The parameters α and b determine the shape of the distribution, with α being the *characteristic exponent* (or *tail index*) and b the *skewness* parameter—an interpretation justified by the fact that, for $b = 0$, $f_{\alpha,b,\gamma,\delta}$ is symmetric with respect to δ and, for $0 < b \le 1$ (resp., $-1 \le b < 0$), skewed to the right (resp., to the left)— see Section 1.2 of Samorodnitsky and Taqqu (1994) for details. The notations $F_{\boldsymbol{\theta}}$ and $F_{\alpha,b,\gamma,\delta}$ will be used for the distribution function associated with $f_{\boldsymbol{\theta}}$.

Some particular choices of $\boldsymbol{\theta}$ yield well-known distributions, namely the Gaussian ($\alpha = 2$, any b), the Cauchy ($\alpha = 1$, $b = 0$) and the Lévy ($\alpha = 1/2$, $b = 1$). However, together with the reflected Lévy density, these are the only instances of stable densities that can be expressed explicitly in terms of elementary functions. For all other choices of the parameters, a closed form for $f_{\boldsymbol{\theta}}$ is not possible, and stable distributions either are defined in terms of characteristic functions and inverse Fourier transforms, or via integral formulas (see e.g. Nolan (1997) or Zolotarev (1986)).

Throughout, we consider a vector $\mathbf{X}^{(n)} := (X_1^{(n)}, \ldots, X_n^{(n)})'$ of observations satisfying

$$X_i^{(n)} = a + \sum_{k=1}^{K} c_{ik}^{(n)} \beta_k + \epsilon_i^{(n)},\ i = 1, \ldots, n, \tag{2.3}$$

for some intercept $a \in \mathbb{R}$ and the regression parameters $\boldsymbol{\beta} := (\beta_1, \ldots, \beta_K)' \in \mathbb{R}^K$; $c_{i1}^{(n)}, \ldots, c_{iK}^{(n)}$ ($i = 1, \ldots, n$) are regression constants, and $\{\epsilon_i^{(n)}, i \in \mathbb{N}\}$ is a sequence of nonobservable i.i.d. errors with stable density $f_{\boldsymbol{\theta}}$, $\boldsymbol{\theta} = (\alpha, b, \gamma, 0) \in \boldsymbol{\Theta}$.

The construction of our R-estimators is based on the uniform local asymptotic normality (ULAN) property, with respect to $\boldsymbol{\beta}$, of the regression model (2.3) under stable error densities. That property is established in Hallin *et al.* (2010) under the following technical assumptions. Without loss of generality, we impose that $\sum_{i=1}^n c_{ik}^{(n)} = 0$ for $k = 1, \ldots, K$; letting $\mathbf{c}_i^{(n)} := (c_{i1}^{(n)}, \ldots, c_{iK}^{(n)})'$, $\mathbb{C}^{(n)} := n^{-1} \sum_{i=1}^n \mathbf{c}_i^{(n)} \mathbf{c}_i^{(n)\prime}$, we make the following assumptions on the asymptotic behavior of the regression constants.

Assumption (A1) For all $n \in \mathbb{N}$, $\mathbb{C}^{(n)}$ is positive definite and converges, as $n \to \infty$, to a positive definite matrix $\mathbb{K}^{-2}$.

Assumption (A2) (Noether conditions) For all $k = 1, \ldots, K$, one has

$$\lim_{n\to\infty} \left\{ \max_{1 \le t \le n} \left(c_{tk}^{(n)}\right)^2 \Big/ \sum_{t=1}^{n} \left(c_{tk}^{(n)}\right)^2 \right\} = 0.$$

Denoting by $\mathrm{P}^{(n)}_{\boldsymbol{\theta},a,\boldsymbol{\beta}}$ the probability distribution of $\mathbf{X}^{(n)}$ under (2.3), let

$$Z_i^{(n)}(\boldsymbol{\beta}) := X_i^{(n)} - a - \sum_{k=1}^{K} c_{ik}^{(n)} \beta_k, \qquad i = 1, \ldots, n$$

stand for the residuals associated with the value $\boldsymbol{\beta}$ of the regression parameter: under $\mathrm{P}^{(n)}_{\boldsymbol{\theta},a,\boldsymbol{\beta}}$, the $Z_i^{(n)}(\boldsymbol{\beta})$'s thus are i.i.d. with density $f_{(\alpha,b,\gamma,0)}$. Here and in the sequel, we write $Z_i^{(n)}(\boldsymbol{\beta})$ instead of $Z_i^{(n)}(a, \boldsymbol{\beta})$ for the sake of simplicity. Although the quantity appearing in Proposition 2.1 depends on a, the rank-based statistics $\underset{\sim}{\boldsymbol{\Delta}}_J^{(n)}$ defined in (2.5) below do not, as the $Z_i^{(n)}(\boldsymbol{\beta})$'s only enter the definition through their ranks, which do not depend on a (fortunately so, as a remains an unspecified nuisance). The following result is proved in Hallin *et al.* (2010).

Proposition 2.1 (ULAN, Hallin, Swan, Verdebout and Veredas 2010). *Suppose that Assumptions (A1) and (A2) hold. Fix $\boldsymbol{\theta} = (\alpha, b, \gamma, 0) \in \boldsymbol{\Theta}$. Then, model (2.3) (the family $\{\mathrm{P}_{\boldsymbol{\theta},a,\boldsymbol{\beta}}^{(n)} | \ \boldsymbol{\beta} \in \mathbb{R}^K\}$), is ULAN with respect to $\boldsymbol{\beta}$, with contiguity rate $n^{1/2}$. More precisely, letting $\boldsymbol{\nu}(n) := n^{-\frac{1}{2}}\mathbb{K}^{(n)}$ with $\mathbb{K}^{(n)} := \left(\mathbb{C}^{(n)}\right)^{-1/2}$, for all $\boldsymbol{\beta} \in \mathbb{R}^K$, all sequences $\boldsymbol{\beta}^{(n)}$ such that $\boldsymbol{\nu}^{-1}(n)(\boldsymbol{\beta}^{(n)} - \boldsymbol{\beta}) = O(1)$ and all bounded sequences $\boldsymbol{\tau}^{(n)} \in \mathbb{R}^K$,*

$$\begin{aligned}
\Lambda_{\boldsymbol{\theta},\boldsymbol{\beta}^{(n)}+\boldsymbol{\nu}(n)\boldsymbol{\tau}^{(n)}}^{(n)} &:= \log\left[\mathrm{dP}_{\boldsymbol{\theta},a,\boldsymbol{\beta}^{(n)}+\boldsymbol{\nu}(n)\boldsymbol{\tau}^{(n)}}^{(n)} / \mathrm{dP}_{\boldsymbol{\theta},a,\boldsymbol{\beta}^{(n)}}^{(n)}\right] \\
&= \log\left[\textstyle\prod_{t=1}^n f_{\boldsymbol{\theta}}(Z_t^{(n)}(\boldsymbol{\beta}+\boldsymbol{\nu}(n)\boldsymbol{\tau}^{(n)}))/\prod_{t=1}^n f_{\boldsymbol{\theta}}(Z_t^{(n)}(\boldsymbol{\beta}))\right] \\
&= \boldsymbol{\tau}^{(n)\prime}\boldsymbol{\Delta}_{\boldsymbol{\theta}}^{(n)}(\boldsymbol{\beta}^{(n)}) - \tfrac{1}{2}\boldsymbol{\tau}^{(n)\prime}\boldsymbol{\tau}^{(n)}\mathcal{I}(\boldsymbol{\theta}) + o_P(1)
\end{aligned}$$

under $\mathcal{H}_{\boldsymbol{\theta}}^{(n)}(\boldsymbol{\beta})$ as $n \to \infty$, where, setting $\varphi_{\boldsymbol{\theta}} := -\dot{f}_{\boldsymbol{\theta}}/f_{\boldsymbol{\theta}}$, with $\dot{f}_{\boldsymbol{\theta}}$ the derivative of $x \mapsto f_{\boldsymbol{\theta}}(x)$ and

$$\mathcal{I}(\boldsymbol{\theta}) := \int_{-\infty}^{\infty} \varphi_{\boldsymbol{\theta}}^2(x) f_{\boldsymbol{\theta}}(x)\mathrm{d}x,$$

$\mathcal{I}(\boldsymbol{\theta})\mathbf{I}_K$ is the information matrix and

$$\boldsymbol{\Delta}_{\boldsymbol{\theta}}^{(n)}(\boldsymbol{\beta}) := n^{-1/2}\mathbb{K}^{(n)\prime}\sum_{i=1}^n \varphi_{\boldsymbol{\theta}}\left(Z_i^{(n)}(\boldsymbol{\beta})\right)\mathbf{c}_i^{(n)} \xrightarrow{\mathcal{L}} \mathcal{N}\big(\mathbf{0},\ \mathcal{I}(\boldsymbol{\theta})\mathbf{I}_K\big) \tag{2.4}$$

the central sequence.

ULAN, here as in Hallin *et al.* (2010), is stated under stable distributions, but of course is well known to hold under any density f such that $f^{1/2}$ is differentiable in quadratic mean; $\mathrm{P}_{\boldsymbol{\theta},a,\boldsymbol{\beta}}^{(n)}$, $\varphi_{\boldsymbol{\theta}}$ and $\mathcal{I}(\boldsymbol{\theta})$ then are to be replaced with $\mathrm{P}_{f,a,\boldsymbol{\beta}}^{(n)}$, $\varphi_f := 2D(f^{1/2})/f^{1/2}$ and $\mathcal{I}_f$, where $D(f^{1/2})$ stands for the quadratic mean derivative of $f^{1/2}$ and $\mathcal{I}_f := \int_{-\infty}^{\infty}\varphi_f^2(x)f(x)\mathrm{d}x$. Denote by $\mathcal{F}$ that class of densities and by $\boldsymbol{\Delta}_f^{(n)}(\boldsymbol{\beta})$ the corresponding central sequences.

2.2. One step R-estimators.

The vector $\mathbf{R}^{(n)} = \mathbf{R}^{(n)}(\boldsymbol{\beta}) := (R_1^{(n)}, \ldots, R_n^{(n)})$, where $R_i^{(n)} = R_i^{(n)}(\boldsymbol{\beta})$ denotes the rank of the residual $Z_i^{(n)} = Z_i^{(n)}(\boldsymbol{\beta})$, $i = 1, \ldots, n$, among $Z_1^{(n)}, \ldots, Z_n^{(n)}$, is distribution-free as f and a range over the class of all nonvanishing densities and $\mathbb{R}$, respectively. Throughout, we consider the class of rank-based statistics

$$\underset{\sim}{\boldsymbol{\Delta}}_J^{(n)}(\boldsymbol{\beta}) := n^{-\frac{1}{2}}\mathbb{K}^{(n)\prime}\sum_{i=1}^n J\left(\frac{R_i^{(n)}}{n+1}\right)\mathbf{c}_i^{(n)}, \tag{2.5}$$

where $J : (0,1) \to \mathbb{R}$ is some score generating function satisfying

Assumption (B) The score function $J : (0,1) \to \mathbb{R}$ is not constant, and the difference $J_1 - J_2$ between two right-continuous and square integrable non-decreasing monotone functions J_1 and $J_2 : (0,1) \to \mathbb{R}$.

Strongly unimodal densities f trivially satisfy that assumption.[6] Except for the Gaussian one, stable densities (2.4) are not strongly unimodal. However, $u \mapsto \varphi_f(F^{-1}(u))$ being bounded (in absolute value) and

[6] A density f is called *strongly unimodal* if $f^{1/2}$ is differentiable in quadratic mean and φ_f is monotone increasing; Gaussian, logistic and double exponential densities are strongly unimodal.

continuously differentiable, with a derivative changing signs exactly twice, it has bounded variation, hence can be expressed as the difference between two monotone increasing functions; φ_f therefore also can.

The following result summarizes the asymptotic properties of the rank-based statistics (2.5); see the Appendix for a proof.

Proposition 2.2 *Let Assumptions (A1), (A2) and (B) hold. Then,*

(i) letting $\mathbf{\Delta}_J^{(n)}(\boldsymbol{\beta}) := n^{-\frac{1}{2}}\mathbb{K}^{(n)\prime}\sum_{i=1}^n J(G(Z_i^{(n)}(\boldsymbol{\beta})))\mathbf{c}_i^{(n)}$, *where G stands for the distribution function associated with a density $g \in \mathcal{F}$, we have, under* $\mathrm{P}_{g,a,\boldsymbol{\beta}}^{(n)}$, *as $n \to \infty$,*

$$\underset{\sim}{\mathbf{\Delta}}_J^{(n)}(\boldsymbol{\beta}) - \mathbf{\Delta}_J^{(n)}(\boldsymbol{\beta}) = o_{\mathrm{P}}(1). \tag{2.6}$$

Hence, for $J(u) = \varphi_f(F^{-1}(u))$ with $f \in \mathcal{F}$, $\underset{\sim}{\mathbf{\Delta}}_J^{(n)}(\boldsymbol{\beta})$ is asymptotically equivalent,[7] *under* $\mathrm{P}_{f,a,\boldsymbol{\beta}}^{(n)}$, *to* $\mathbf{\Delta}_f^{(n)}(\boldsymbol{\beta})$*;*

(ii) under $\mathrm{P}_{g,a,\boldsymbol{\beta}}^{(n)}$ $(g \in \mathcal{F})$, $\underset{\sim}{\mathbf{\Delta}}_J^{(n)}(\boldsymbol{\beta})$ is asymptotically normal with mean zero and covariance matrix $\mathcal{J}(J)\mathbf{I}_K$, where $\mathcal{J}(J) := \int_0^1 J^2(u)\mathrm{d}u$;

(iii) under $\mathrm{P}_{g,a,\boldsymbol{\beta}+\boldsymbol{\nu}^{(n)}\boldsymbol{\tau}}$ $(g \in \mathcal{F})$, $\underset{\sim}{\mathbf{\Delta}}_J^{(n)}(\boldsymbol{\beta})$ is asymptotically normal with mean $\mathcal{J}(J,g)\boldsymbol{\tau}$ and covariance matrix $\mathcal{J}(J)\mathbf{I}_K$, where

$$\mathcal{J}(J,g) := \int_0^1 J(u)\varphi_g(G^{-1}(u))\mathrm{d}u; \tag{2.7}$$

(iv) $\underset{\sim}{\mathbf{\Delta}}_J^{(n)}(\boldsymbol{\beta})$ satisfies the asymptotic linearity property

$$\underset{\sim}{\mathbf{\Delta}}_J^{(n)}(\boldsymbol{\beta} + \boldsymbol{\nu}^{(n)}\boldsymbol{\tau}^{(n)}) - \underset{\sim}{\mathbf{\Delta}}_J^{(n)}(\boldsymbol{\beta}) = -\mathcal{J}(J,g)\boldsymbol{\tau}^{(n)} + o_{\mathrm{P}}(1) \tag{2.8}$$

under $\mathrm{P}_{g,a,\boldsymbol{\beta}}^{(n)}$ with $g \in \mathcal{F}$, as $n \to \infty$.

Under the conditions of Proposition 2.1, the Le Cam one-step methodology requires the existence of a preliminary root-n consistent estimator $\hat{\boldsymbol{\beta}}^{(n)}$ of $\boldsymbol{\beta}$. The LAD estimator $\hat{\boldsymbol{\beta}}_{\mathrm{LAD}}^{(n)}$ of $\boldsymbol{\beta}$, which we are considering in the sequel, is one possibility, but any other estimator enjoying root-n consistency under the whole class of stable densities would be an equally valid candidate.

The LAD estimator $(\hat{a}_{\mathrm{LAD}}^{(n)}, \hat{\boldsymbol{\beta}}_{\mathrm{LAD}}^{(n)\prime})'$ of $(a, \boldsymbol{\beta}')'$ is obtained by minimizing the L_1-objective function

$$(\hat{a}_{\mathrm{LAD}}^{(n)}, \hat{\boldsymbol{\beta}}_{\mathrm{LAD}}^{(n)\prime})' := \mathrm{argmin}_{(a,\boldsymbol{\beta})\in\mathbb{R}^{K+1}} \sum_{i=1}^n |Z_i^{(n)}(\boldsymbol{\beta})|.$$

In this context, however, a needs not be estimated, as ranks are insensitive to location shift; we therefore concentrate on $\hat{\boldsymbol{\beta}}_{\mathrm{LAD}}^{(n)}$. In order to control for the uniformity of local behaviors, a discretized version $\hat{\boldsymbol{\beta}}_{\#}^{(n)}$ of $\hat{\boldsymbol{\beta}}_{\mathrm{LAD}}^{(n)}$ should be considered in theoretical asymptotic statements. The discretization trick, which is due to Le Cam, is quite standard in the context of one-step estimation. While retaining root-n consistency, discretized estimators indeed enjoy the important property of *asymptotic local discreteness*, that is, they only take a finite number of distinct values, as $n \to \infty$, in $\boldsymbol{\beta}$-centered balls with $O(n^{-1/2})$ radius. In fixed-n practice, however, such discretizations are irrelevant (the discretization constant can be chosen arbitrarily

[7] Since central sequences are only defined up to $o_{\mathrm{P}}(1)$ terms, $\underset{\sim}{\mathbf{\Delta}}_J^{(n)}(\boldsymbol{\beta})$ thus is a rank-based version of the central sequence $\mathbf{\Delta}_f^{(n)}(\boldsymbol{\beta})$.

large). For the sake of simplicity, we will henceforth tacitly assume that $\hat{\boldsymbol{\beta}}^{(n)}_{\rm LAD}$, in asymptotic statements, has been adequately discretized.

Were $\mathcal{J}^{-1}(J,g)$ a known quantity, the one-step R-estimator of $\boldsymbol{\beta}$ would take (since the asymptotic variance of $\utilde{\boldsymbol{\Delta}}^{(n)}_J$ is proportional to an identity matrix) the following very simple form:

$$\utilde{\tilde{\boldsymbol{\beta}}}^{(n)}_J := \hat{\boldsymbol{\beta}}^{(n)}_{\rm LAD} + \boldsymbol{\nu}^{(n)} \mathcal{J}^{-1}(J,g)\, \utilde{\boldsymbol{\Delta}}^{(n)}_J(\hat{\boldsymbol{\beta}}^{(n)}_{\rm LAD}). \tag{2.9}$$

It readily follows from (2.8) (as well as from standard results on one-step estimation: see, e.g., Proposition 1 in Chapter 6 of Le Cam and Yang (2000)) that

$$\boldsymbol{\nu}^{-1}(n)(\utilde{\tilde{\boldsymbol{\beta}}}^{(n)}_J - \boldsymbol{\beta}) = \mathcal{J}^{-1}(J,g)\, \utilde{\boldsymbol{\Delta}}^{(n)}_J(\boldsymbol{\beta}) + o_{\rm P}(1),$$

hence, that $\boldsymbol{\nu}^{-1}(n)(\utilde{\tilde{\boldsymbol{\beta}}}^{(n)}_J - \boldsymbol{\beta})$ is asymptotically $\mathcal{N}(\mathbf{0}, (\mathcal{J}(J)/\mathcal{J}^2(J,g))\mathbf{I}_K)$ under ${\rm P}^{(n)}_{g,a,\boldsymbol{\beta}}$ $(g \in \mathcal{F})$. This in turn implies that $\boldsymbol{\nu}^{-1}(n)(\utilde{\tilde{\boldsymbol{\beta}}}^{(n)}_J - \boldsymbol{\beta})$, for $J(u) = \varphi_f(F^{-1}(u))$, is asymptotically $\mathcal{N}(\mathbf{0}, \mathcal{J}^{-1}(J)\mathbf{I}_K)$ under ${\rm P}^{(n)}_{f,a,\boldsymbol{\beta}}$, that is, reaches parametric efficiency at correctly specified density $f = g$.

Unfortunately, the scalar *cross-information quantity* $\mathcal{J}(J,g)$ is not known—a phenomenon that does not appear in the usual one-step method, based on the "parametric central sequence" associated with some correctly identified density $f = g$. Under definition (2.9), $\utilde{\tilde{\boldsymbol{\beta}}}^{(n)}_J$ therefore is not a genuine estimator. That cross-information quantity $\mathcal{J}(J,g)$ thus has to be consistently estimated. To obtain such a consistent estimator, we adopt here the idea first developed in Hallin *et al.* (2006) and generalized in Cassart *et al.* (2010).

For all $v > 0$, define $\tilde{\boldsymbol{\beta}}^{(n)}(v) := \hat{\boldsymbol{\beta}}^{(n)}_{\rm LAD} + \boldsymbol{\nu}^{(n)} v\, \utilde{\boldsymbol{\Delta}}^{(n)}_J(\hat{\boldsymbol{\beta}}^{(n)}_{\rm LAD})$, and consider the scalar product

$$h^{(n)}(v) := \big(\utilde{\boldsymbol{\Delta}}^{(n)}_J(\hat{\boldsymbol{\beta}}^{(n)}_{\rm LAD})\big)' \utilde{\boldsymbol{\Delta}}^{(n)}_J(\tilde{\boldsymbol{\beta}}^{(n)}(v)).$$

Proposition 2.2, the consistency and local asymptotic discreteness of $\hat{\boldsymbol{\beta}}^{(n)}_{\rm LAD}$, and the definition of $\tilde{\boldsymbol{\beta}}^{(n)}(v)$ entail that, under ${\rm P}^{(n)}_{g,a,\boldsymbol{\beta}}$ with $g \in \mathcal{F}$,

$$\begin{aligned} h^{(n)}(v) &= \Big(\utilde{\boldsymbol{\Delta}}^{(n)}_J(\hat{\boldsymbol{\beta}}^{(n)}_{\rm LAD})\Big)' \Big(\utilde{\boldsymbol{\Delta}}^{(n)}_J(\hat{\boldsymbol{\beta}}^{(n)}_{\rm LAD}) - \mathcal{J}(J,g)\, n^{1/2}(\mathbb{K}^{(n)})^{-1}(\tilde{\boldsymbol{\beta}}^{(n)}(v) - \hat{\boldsymbol{\beta}}^{(n)}_{\rm LAD})\Big) + o_{\rm P}(1) \\ &= \Big(\utilde{\boldsymbol{\Delta}}^{(n)}_J(\hat{\boldsymbol{\beta}}^{(n)}_{\rm LAD})\Big)' \Big(\utilde{\boldsymbol{\Delta}}^{(n)}_J(\hat{\boldsymbol{\beta}}^{(n)}_{\rm LAD}) - \mathcal{J}(J,g) v\, \utilde{\boldsymbol{\Delta}}^{(n)}_J(\hat{\boldsymbol{\beta}}^{(n)}_{\rm LAD})\Big) + o_{\rm P}(1) \\ &= (1 - \mathcal{J}(J,g)v)\, h^{(n)}(0) + o_{\rm P}(1) \end{aligned} \tag{2.10}$$

for any $v > 0$; this provides the intuition for taking the solution of $h(v) = 0$ as an estimation of $(\mathcal{J}(J,g))^{-1}$. And, provided that $h^{(n)}(0)$ is not $o_{\rm P}(1)$, a consistent estimator of $(\mathcal{J}(J,g))^{-1}$ indeed would be

$$\hat{v}^{(n)} := \inf\{v > 0 : h^{(n)}(v) < 0\}.$$

More precisely, consider a discretization of the positive half-line, with $v_\ell := \ell/c$, $\ell \in \mathbb{N}$, $c > 0$ a (typically, large) discretizing constant, the value of which, however, plays no role in asymptotic statements. Putting

$$v^{(n)}_- := \min\{\ell \text{ such that } h^{(n)}(v^{(n)}_{\ell+1}) < 0\} \quad \text{and} \quad v^{(n)}_+ := v^{(n)}_- + \frac{1}{c}, \tag{2.11}$$

consider the linear interpolation

$$\hat{v}^{(n)} := v^{(n)}_- \left(1 - \frac{h^{(n)}(v^{(n)}_-)}{h^{(n)}(v^{(n)}_-) - h^{(n)}(v^{(n)}_+)}\right) + v^{(n)}_+ \frac{h^{(n)}(v^{(n)}_-)}{h^{(n)}(v^{(n)}_-) - h^{(n)}(v^{(n)}_+)}. \tag{2.12}$$

It follows from Proposition 2.1 in Cassart *et al.* (2010) that, unless $h^{(n)}(0)$ is $o_{\rm P}(1)$, $\widehat{\mathcal{J}}(J,g) := (\hat{v}^{(n)})^{-1}$ provides a consistent estimator of the cross-information quantity $\mathcal{J}(J,g)$. Our one-step R-estimator then is defined as

$$\underset{\sim}{\boldsymbol{\beta}}{}_J^{(n)} := \tilde{\boldsymbol{\beta}}^{(n)}(\widehat{\mathcal{J}}^{-1}(J,g)) = \hat{\boldsymbol{\beta}}^{(n)}_{\text{LAD}} + \boldsymbol{\nu}^{(n)}\widehat{\mathcal{J}}^{-1}(J,g)\,\underset{\sim}{\boldsymbol{\Delta}}{}_J^{(n)}(\hat{\boldsymbol{\beta}}^{(n)}_{\text{LAD}}).$$

Now, in case J is such that $\underset{\sim}{\boldsymbol{\Delta}}{}_J^{(n)}(\hat{\boldsymbol{\beta}}^{(n)}_{\text{LAD}}) = o_{\text{P}}(1)$, that is, if the Laplace or double-exponential score function $u \mapsto J_{\text{L}}(u) := \sqrt{2}\,\text{sign}(u - 1/2)$, is considered, we have (see Proposition 2.4) $\underset{\sim}{\boldsymbol{\beta}}{}_{J_L}^{(n)} = \hat{\boldsymbol{\beta}}^{(n)}_{\text{LAD}} + o_{\text{P}}(n^{-1/2})$ and $\underset{\sim}{\tilde{\boldsymbol{\beta}}}{}_{J_{\text{L}}}^{(n)} = \hat{\boldsymbol{\beta}}^{(n)}_{\text{LAD}} + o_{\text{P}}(n^{-1/2})$, so that our estimator coincides, asymptotically, with the LAD estimator.

The following result (see the Appendix for a proof) summarizes the asymptotic properties of $\underset{\sim}{\boldsymbol{\beta}}{}_J^{(n)}$.

Proposition 2.3 *Let Assumptions (A1), (A2) and (B) hold. Then,* $n^{1/2}(\underset{\sim}{\boldsymbol{\beta}}{}_J^{(n)} - \boldsymbol{\beta})$ *is asymptotically normal with mean zero and covariance matrix* $\big(\mathcal{J}(J)/\mathcal{J}^2(J,g)\big)\mathbb{K}^2$ *under* $\text{P}^{(n)}_{g,a,\boldsymbol{\beta}}$ *with* $g \in \mathcal{F}$*. Therefore, letting* $J(u) = \varphi_f(F^{-1}(u))$*,* $\underset{\sim}{\boldsymbol{\beta}}{}_J^{(n)}$ *achieves the parametric efficiency bound under* $\text{P}^{(n)}_{f,a,\boldsymbol{\beta}}$*.*

In view of Proposition 2.3, the asymptotic relative efficiences of our R-estimators clearly coincide with those of the corresponding tests developed in Hallin *et al.* (2010). More precisely, we have that

$$\text{ARE}_g(J_1/J_2) = \mathcal{J}^2(J_1,g)\mathcal{J}(J_2)/\mathcal{J}^2(J_2,g)\mathcal{J}(J_1), \tag{2.13}$$

where $\text{ARE}_g(J_1/J_2)$ denotes the asymptotic relative efficiency, under density g, of the R-estimator $\underset{\sim}{\boldsymbol{\beta}}{}_{J_1}^{(n)}$, based on the score-generating function J_1, with respect to the R-estimator $\underset{\sim}{\boldsymbol{\beta}}{}_{J_2}^{(n)}$, based on the score-generating function J_2.

Table 1: AREs of R-estimators with respect to LAD estimators

Estimators	Underlying stable density			
	$\alpha = 2$; $b = 0$	$\alpha = 1.8$; $b = 0$	$\alpha = 1.8$; $b = 0.5$	$\alpha = 0.5$; $b = 0.5$
$\underset{\sim}{\boldsymbol{\beta}}{}_{J_{\text{W}}}^{(n)}/\hat{\boldsymbol{\beta}}^{(n)}_{\text{LAD}}$	1.4999	1.3888	1.3984	1.7776
$\underset{\sim}{\boldsymbol{\beta}}{}_{J_{\text{vdW}}}^{(n)}/\hat{\boldsymbol{\beta}}^{(n)}_{\text{LAD}}$	1.5708	1.3056	1.3285	1.251
$\underset{\sim}{\boldsymbol{\beta}}{}_{J_{\text{C}}}^{(n)}/\hat{\boldsymbol{\beta}}^{(n)}_{\text{LAD}}$	0.6759	0.7880	0.7769	2.007
$\underset{\sim}{\boldsymbol{\beta}}{}_{J_{1.8;0}}^{(n)}/\hat{\boldsymbol{\beta}}^{(n)}_{\text{LAD}}$	1.4459	1.4183	1.4222	1.6453
$\underset{\sim}{\boldsymbol{\beta}}{}_{J_{1.8;.5}}^{(n)}/\hat{\boldsymbol{\beta}}^{(n)}_{\text{LAD}}$	1.4452	1.3969	1.4459	1.4432
$\underset{\sim}{\boldsymbol{\beta}}{}_{J_{.5;.5}}^{(n)}/\hat{\boldsymbol{\beta}}^{(n)}_{\text{LAD}}$	0.0925	0.1099	0.1175	21.2364

AREs for R-estimators based on various scores with respect to the LAD estimator. Columns correspond to the (stable) densities under which AREs are computed, rows to the scores considered: Wilcoxon (J_{W}), van der Waerden (J_{vdW}), Cauchy (J_{C}), and three ($\delta = 0$, $\gamma = 1$) stable scores ($J_{\alpha;b}$); recall that the R-estimator based on Laplace scores asymptotically coincides with the LAD estimator (see Proposition 2.4).

Traditional scores (such as the van der Waerden, Wilcoxon and Laplace ones) are associated with some classical light-tailed densities (such as the normal, logistic and double-exponential), leading to the score-generating functions

$$J_{\text{vdW}}(u) = \Phi^{-1}(u), \quad J_{\text{W}}(u) = \frac{\pi}{\sqrt{3}}(2u-1), \quad \text{and} \quad J_{\text{L}}(u) = \sqrt{2}\,\text{sign}(u - 1/2),$$

respectively, where Φ denotes, as usual, the standard normal distribution function. The resulting R-estimators are reaching parametric efficiency under Gaussian, logistic and double-exponential densities, respectively. *Stable scores*, of the form $J_{\boldsymbol{\theta}}(x) = -\dot{f}_{\boldsymbol{\theta}}(F_{\boldsymbol{\theta}}^{-1}(x))/f_{\boldsymbol{\theta}}(F_{\boldsymbol{\theta}}^{-1}(x))$, where $f_{\boldsymbol{\theta}}$ is some stable density, also

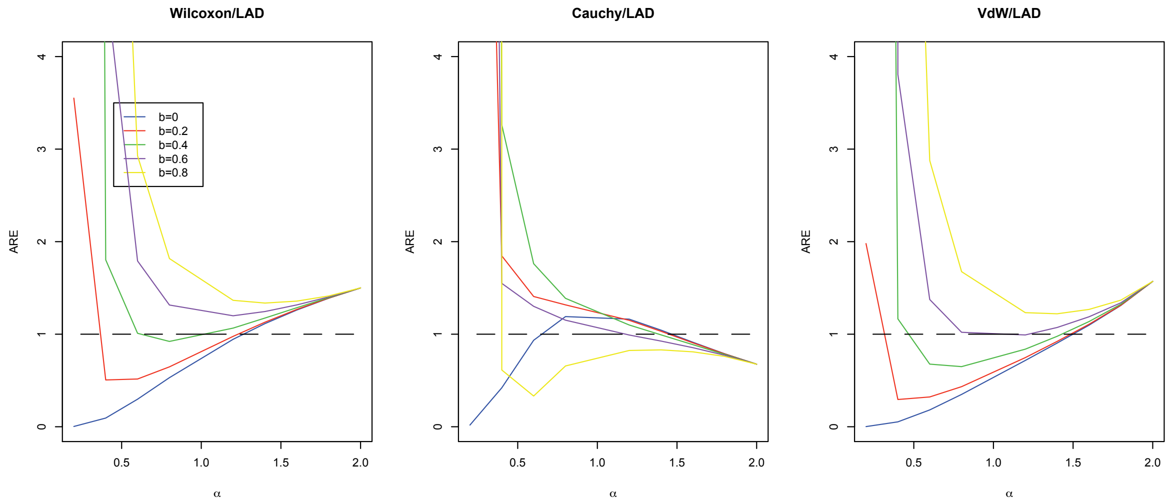

Figure 1: AREs of R-estimators based on Wilcoxon, Cauchy and van der Waerden scores, with respect to the LAD estimator, as a function of α and for various values of b.

can be considered, not under closed form, though; we refer to Appendix B of Hallin *et al.* (2010), where rank tests based on such stable scores are discussed, for details. Table 1 and Figures 1 and 2 provide numerical values of AREs in (2.13) for various estimators and underlying stable densities. Interestingly, the R-estimators based on the stable scores for tail index 1.4 uniformly dominate, irrespective of the asymmetry parameter b, the LAD estimator for all values of $\alpha \in [1, 2]$. Their AREs with respect to LAD estimators moreover culminates in the vicinity of $\alpha = 1.8$, a value which is generally recognized as a reasonable tail index for financial data.[8]

To conclude this section, the following result establishes the asymptotic equivalence between the LAD estimator and the Laplace R-estimator (based on the score function J_{L}); see the Appendix for a proof.

Proposition 2.4 *Let Assumptions (A1) and (A2) hold. Then, the difference* $\underset{\sim}{\boldsymbol{\beta}}{}^{(n)}_{J_{\mathrm{L}}} - \hat{\boldsymbol{\beta}}^{(n)}_{\mathrm{LAD}}$ *is* $o_{\mathrm{P}}(n^{-1/2})$ *as* $n \to \infty$ *under* $\mathrm{P}^{(n)}_{g,a,\boldsymbol{\beta}}$ *for any* $g \in \mathcal{F}$ *such that* g *is strictly positive at the median* $G^{-1}(\frac{1}{2})$.

As a direct consequence, the ARE (under $\mathrm{P}^{(n)}_{g,a,\boldsymbol{\beta}}$ with $g \in \mathcal{F}$) of any estimator $\tilde{\boldsymbol{\beta}}^{(n)}$ with respect to $\hat{\boldsymbol{\beta}}^{(n)}_{\mathrm{LAD}}$ is equal to the ARE of $\tilde{\boldsymbol{\beta}}^{(n)}$ with respect to $\underset{\sim}{\boldsymbol{\beta}}{}^{(n)}_{J_{\mathrm{L}}}$.

3. Finite-sample performance.

This section is devoted to a simulation study of the finite-sample performances of the various R-estimators described in the previous sections and some of their competitors, in order to check whether these performances are are in line with the ARE results of Table 1.

We generated $M = 1000$ samples from two multiple regression models,

$$Y_i^{(1)} = c_{i1} + c_{i2} + \epsilon_i, \quad i = 1, \ldots, n = 100, \tag{3.14}$$

with two regressors, and

$$Y_i^{(2)} = c_{i1} + c_{i2} + c_{i3} + c_{i4} + \epsilon_i, \quad i = 1, \ldots, n = 100, \tag{3.15}$$

[8]Dominicy and Veredas (2010) found that the estimated α for 22 major worldwide market indexes (nine years of daily returns) ranges between 1.55 to 1.90, with an average of 1.75. Similar values have been obtained for other financial assets, e.g. in Mittnik *et al.* (2000) or Deo (2002).

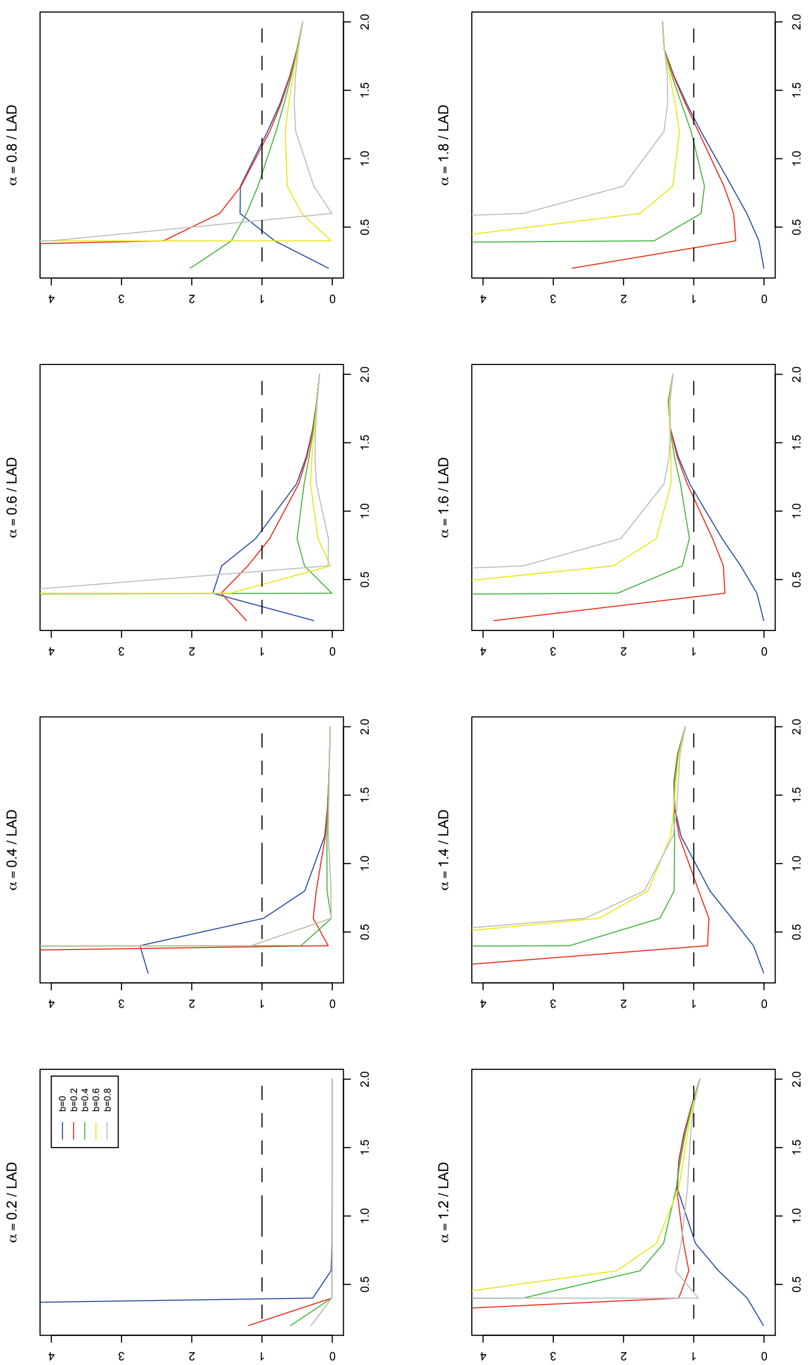

Figure 2: AREs under stable distributions of R-estimators based on various stable scores with respect to the LAD estimator, as a functions of α and b.

with four regressors, both with alpha-stable i.i.d. ϵ_i's. The regression constants c_{ij} (the same ones across the 1000 replications) were drawn (independently) from the uniform distribution on $[-1,1]^2$ and $[-1,1]^4$, respectively. Letting $\mathbf{1}_K := (1,1,\ldots,1) \in \mathbb{R}^K$, the true values of the regression parameters are thus $\boldsymbol{\beta} = \mathbf{1}_2$ in model (3.14) and $\boldsymbol{\beta} = \mathbf{1}_4$ in model (3.15).

Denoting by $\boldsymbol{\beta}^{(n)}(j) = (\beta_1^{(n)}(j), \ldots, \beta_K^{(n)}(j))'$ ($j = 1, \ldots, M$; $K = 2$ or 4 depending on the model) an estimator $\boldsymbol{\beta}^{(n)}$ computed from the jth replication, the empirical bias and empirical mean square error for the first component $\beta_1^{(n)}$ of $\boldsymbol{\beta}^{(n)}$ are

$$\text{BIAS}(\boldsymbol{\beta}^{(n)}) := \frac{1}{M}\sum_{j=1}^{M}(\beta_1^{(n)}(j) - 1), \quad \text{and} \quad \text{MSE}_l(\boldsymbol{\beta}^{(n)}) := \frac{1}{M}\sum_{j=1}^{M}(\beta_1^{(n)}(j) - 1)^2,$$

respectively; models (3.14) and (3.15) being perfectly symmetric, efficiency comparisons can be based on that first component only. These quantities were computed for the least squares $\hat{\boldsymbol{\beta}}^{(n)}_{\text{LS}}$ and the LAD estimators $\hat{\boldsymbol{\beta}}^{(n)}_{\text{LAD}}$, the one-step versions $\underset{\sim}{\boldsymbol{\beta}}^{(n)}_{J_{\text{vdW}}}$, $\underset{\sim}{\boldsymbol{\beta}}^{(n)}_{J_{\text{W}}}$ and $\underset{\sim}{\boldsymbol{\beta}}^{(n)}_{J_{\text{L}}}$ of the van der Waerden, Wilcoxon and Laplace estimators, and the one-step R-estimators $\underset{\sim}{\boldsymbol{\beta}}^{(n)}_{J_{\alpha/\text{b}}}$ associated with the stable scores with tail index α and skewness parameter b ($\alpha = 1.8/b = 0$; $\alpha = 1.8/b = 0.5$; $\alpha = 1.2/b = 0$; $\alpha = 1.2/b = 0.5$; $\alpha = 0.5/b = 0.5$), respectively. For the sake of comparison, we also computed the bias and mean square errors associated with the Argmin (Hodges-Lehmann 1963; Jurečková 1971) versions $\underset{\sim}{\boldsymbol{\beta}}^{(n)}_{\text{HL;W}}$, $\underset{\sim}{\boldsymbol{\beta}}^{(n)}_{\text{HL; vdW}}$ and $\underset{\sim}{\boldsymbol{\beta}}^{(n)}_{\text{HL;1.8/0}}$ of the Wilcoxon, van der Waerden, and stable score ($\alpha = 1.8/b = 0$) R-estimators; the latter were computed via the Nelder-Mead (1965) method.

Results are collected in Table 3 for model (3.14) and Table 3 for model (3.15), and confirm the theoretical findings of the previous sections. Least squares behave quite poorly, and fail miserably as the tail index decreases, while least absolute deviations maintain an overall good performance. The empirical performances of R-estimators are consistent with theoretical ARE rankings. Depending on the scores and the actual underlying tail index and skewness parameter, R-estimators may or may not improve on least absolute deviations. Stable score-based R-estimators, as a rule, outperform least absolute deviations, as expected, under correctly specified values of the tail index.

It is worth noting that one-step R-estimators are doing better than their Hodges-Lehmann counterparts in tmodel (3.15), that is, when the parameter is of dimension four. This is most probably due to computational problems related with the Argmin approach in higher dimension; such problems do not occur in the one-step approach. Further evidence of this phenomenon is provided in Table 3, where we report results for the one-step and Hodges-Lehmann versions of the van der Waerden R-estimator in regression models of the form

$$Y_i^{(2)} = c_{i1} + c_{i2} + \ldots + c_{iK} + \epsilon_i, \quad i = 1, \ldots, n = 100, \tag{3.16}$$

with K regressors, $K = 6,\ 10,\ 15$ (same number of replications; regression constants uniform over $[-1,1]^K$). Irrespective of the underlying stable density, the superiority of the one-step version quite significantly increases with K.

4. Conclusion.

Stable densities constitute a broad and flexible class of probability density functions, allowing for asymmetry and heavy tails. Their theoretical properties make them quite appealing in a variety of applications, including econometric and financial ones. Traditional inference methods, however, in general are not valid in models involving stable error: classical tests no longer satisfy nominal probability level constraints, and estimators, as a rule, are rate-suboptimal. On the other hand, due to the absence of closed-form likelihoods, theoretical optimality results are not easily derived. And, still for the same reason, their practical implementation is all but straightforward.

Table 2: Empirical bias and mean square error for various estimators of $\boldsymbol{\beta}$ in model (3.14)

Estimator		Underlying stable density (α/b)					
		$\alpha=2/b=0$	$\alpha=1.8/b=0$	$\alpha=1.8/b=0.5$	$\alpha=1.2/b=0$	$\alpha=1.2/b=0.5$	$\alpha=0.5/\beta=0.5$
$\hat{\boldsymbol{\beta}}^{(n)}_{\rm LS}$	(Bias)	.00193	-.00134	.01385	.18680	-.19255	740527.6
	(MSE)	.06770	.19459	.27336	124.46	88.070	5.3560e+14
$\hat{\boldsymbol{\beta}}^{(n)}_{\rm LAD}$	(Bias)	.00167	-.00087	.00502	.02995	.00646	-.02438
	(MSE)	.10674	.10411	.11638	.11560	.13396	.23233
$\underset{\sim}{\boldsymbol{\beta}}^{(n)}_{J_{\rm vdW}}$	(Bias)	.00256	-.00136	.00694	.03376	-.00243	.00745
	(MSE)	.06878	.07694	.08545	.15165	.14499	.49418
$\underset{\sim}{\boldsymbol{\beta}}^{(n)}_{J_{\rm W}}$	(Bias)	.00076	.00015	.00920	.02957	-.00147	-.00165
	(MSE)	.07234	.07454	.08366	.12060	.12219	.29830
$\underset{\sim}{\boldsymbol{\beta}}^{(n)}_{J_{\rm L}}$	(Bias)	.00167	-.00087	.00502	.02995	.00646	-.02438
	(MSE)	.10674	.10411	.11638	.11560	.13396	.23232
$\underset{\sim}{\boldsymbol{\beta}}^{(n)}_{J_{1.8/0}}$	(Bias)	.00250	.00063	.00883	.03046	.00068	.00267
	(MSE)	.07088	.07457	.08310	.12976	.12820	.36304
$\underset{\sim}{\boldsymbol{\beta}}^{(n)}_{J_{1.8/.5}}$	(Bias)	.00187	-.00119	.01057	.03284	-.00037	.00284
	(MSE)	.07104	.07683	.08139	.13562	.12398	.34625
$\underset{\sim}{\boldsymbol{\beta}}^{(n)}_{J_{1.2/0}}$	(Bias)	.00424	.00353	.01373	.02155	-.00363	.01652
	(MSE)	.11613	.09812	.11040	.09641	.10971	.17458
$\underset{\sim}{\boldsymbol{\beta}}^{(n)}_{J_{1.2/.5}}$	(Bias)	.00670	-.00418	.01609	.02735	.00310	-.00199
	(MSE)	.11416	.10382	.10822	.11455	.08917	.11282
$\underset{\sim}{\boldsymbol{\beta}}^{(n)}_{J_{.5/.5}}$	(Bias)	.01070	.03350	.00357	.04768	-.01671	.00466
	(MSE)	.22575	.28311	.24386	.35926	.18999	.12103
$\underset{\sim}{\boldsymbol{\beta}}^{(n)}_{\rm HL;\, vdW}$	(Bias)	-.01668	-.01040	-.00253	.04306	-.01664	.11740
	(MSE)	.07936	.08958	.09508	.20227	.20441	1.1934
$\underset{\sim}{\boldsymbol{\beta}}^{(n)}_{\rm HL;\, W}$	(Bias)	-.00672	-.02019	-.01113	-.01052	-.03408	-.24449
	(MSE)	.08225	.09071	.09702	.16290	.14918	.82852
$\underset{\sim}{\boldsymbol{\beta}}^{(n)}_{\rm HL;\, 1.8/0}$	(Bias)	-.02274	-.02834	-.01923	-.01504	-.05129	-.24827
	(MSE)	.09066	.10291	.10488	.18247	.19072	.96871

Empirical bias and MSE of the least square $\hat{\boldsymbol{\beta}}^{(n)}_{\rm LS}$, the LAD $\hat{\boldsymbol{\beta}}^{(n)}_{\rm LAD}$ and various rank-based estimators computed from 1000 replications of model (3.14) with sample size n=100, under various stable error distributions.

Table 3: Empirical bias and mean square error for various estimators of $\boldsymbol{\beta}$ in model (3.15)

Estimator		Underlying stable density (α/b)					
		$\alpha=2/b=0$	$\alpha=1.8/b=0$	$\alpha=1.8/b=0.5$	$\alpha=1.2/b=0$	$\alpha=1.2/b=0.5$	$\alpha=0.5/b=0.5$
$\hat{\boldsymbol{\beta}}^{(n)}_{\mathrm{LS}}$	(Bias)	.00314	.01367	-.01945	-4.09468	-.09272	-47944.35
	(MSE)	.06339	.30161	.12752	15818.91	39.45292	1.23211e+13
$\hat{\boldsymbol{\beta}}^{(n)}_{\mathrm{LAD}}$	(Bias)	.00693	.00880	-.00774	-.00652	.00352	-.00746
	(MSE)	.09995	.09992	.09548	.08495	.09984	.21871
$\underset{\sim}{\boldsymbol{\beta}}^{(n)}_{J_{\mathrm{vdW}}}$	(Bias)	.00378	.00638	-.01177	-.00763	-.01262	-.01902
	(MSE)	.06463	.06964	.07238	.11369	.11015	.35648
$\underset{\sim}{\boldsymbol{\beta}}^{(n)}_{J_{\mathrm{W}}}$	(Bias)	.00542	.00579	-.01236	-.00624	-.00774	-.01330
	(MSE)	.06811	.06847	.06988	.09038	.09127	.22657
$\underset{\sim}{\boldsymbol{\beta}}^{(n)}_{J_{\mathrm{L}}}$	(Bias)	.00693	.00880	-.00774	-.00652	.00352	-.00746
	(MSE)	.09995	.09992	.09548	.08495	.09984	.21871
$\underset{\sim}{\boldsymbol{\beta}}^{(n)}_{J_{1.8/0}}$	(Bias)	.00499	.00531	-.01221	-.00445	-.00980	-.01629
	(MSE)	.06755	.06735	.07021	.09908	.09562	.27044
$\underset{\sim}{\boldsymbol{\beta}}^{(n)}_{J_{1.8/.5}}$	(Bias)	.00339	.00526	-.01109	-.00438	-.01151	-.01722
	(MSE)	.06686	.06914	.06977	.10095	.09397	.25358
$\underset{\sim}{\boldsymbol{\beta}}^{(n)}_{J_{1.2/0}}$	(Bias)	.00802	.00608	-.01297	.00682	.00404	.00226
	(MSE)	.10763	.09229	.08986	.07061	.08406	.13542
$\underset{\sim}{\boldsymbol{\beta}}^{(n)}_{J_{1.2/.5}}$	(Bias)	.00291	.00024	-.01401	.00396	-.00231	-.00573
	(MSE)	.10332	.09233	.08567	.09036	.07037	.07636
$\underset{\sim}{\boldsymbol{\beta}}^{(n)}_{J_{.5/.5}}$	(Bias)	.03400	.03653	-.02823	-.05925	-.00469	-.01970
	(MSE)	.30150	.35030	.28818	.43049	.18807	.19423
$\underset{\sim}{\boldsymbol{\beta}}^{(n)}_{\mathrm{HL;\,vdW}}$	(Bias)	.00401	.00634	-.01208	-.00704	-.01234	-.02138
	(MSE)	.06513	.06968	.07266	.11310	.10956	.38167
$\underset{\sim}{\boldsymbol{\beta}}^{(n)}_{\mathrm{HL;\,W}}$	(Bias)	.00513	.00623	-.01285	-.00547	-.00755	-.01470
	(MSE)	.06854	.06855	.07006	.09010	0.09100	.23734
$\underset{\sim}{\boldsymbol{\beta}}^{(n)}_{\mathrm{HL;\,1.8/0}}$	(Bias)	.00494	.00582	-.01245	-.00396	-.01081	-.01793
	(MSE)	.06783	.06753	.07037	.09854	.09594	.28729

Empirical bias and MSE of the least square $\hat{\boldsymbol{\beta}}^{(n)}_{\mathrm{LS}}$, the LAD $\hat{\boldsymbol{\beta}}^{(n)}_{\mathrm{LAD}}$ and various rank-based estimators computed from 1000 replications of model (3.15) with sample size n=100, under various stable error distributions.

Table 4: One-step R-estimation versus Argmin

Estimator		Underlying stable density (α/b)					
		$\alpha=2/b=0$	$\alpha=1.8/b=0$	$\alpha=1.8/b=0.5$	$\alpha=1.2/b=0$	$\alpha=1.2/b=0.5$	$\alpha=0.5/b=0.5$
$K=6$							
$\underset{\sim}{\beta}^{(n)}_{J_{\rm vdW}}$	(Bias)	-.01991	-.00485	.01084	-.01890	.02246	.00162
	(MSE)	.07707	.08821	.08935	.16485	.15258	.61554
$\underset{\sim}{\beta}^{(n)}_{\rm HL;\, vdW}$	(Bias)	-.19519	-.19834	-.19202	-.36809	-.30435	-.59222
	(MSE)	.24257	.27483	.27461	.58981	.52245	2.51344
$K=10$							
$\underset{\sim}{\beta}^{(n)}_{J_{\rm vdW}}$	(Bias)	-.00877	.00607	.00187	-.00807	-.01376	.06003
	(MSE)	.07834	.09133	.08641	.16835	.15545	1.4346
$\underset{\sim}{\beta}^{(n)}_{\rm HL;\, vdW}$	(Bias)	-.91080	-.89626	-.92196	-1.00979	-.99976	-.97662
	(MSE)	1.04321	1.07289	1.09949	1.50269	1.43327	3.23870
$K=15$							
$\underset{\sim}{\beta}^{(n)}_{J_{\rm vdW}}$	(Bias)	-.00374	-.01421	-.00575	.02479	0.00271	.01123
	(MSE)	.08894	.10969	.10539	.20918	.19621	2.00335
$\underset{\sim}{\beta}^{(n)}_{\rm HL;\, vdW}$	(Bias)	-1.07573	-1.11915	-1.11057	-1.23107	-1.21492	-1.31910
	(MSE)	1.19685	1.33319	1.32890	1.91879	1.88120	4.32374

Empirical bias and MSE of the one-step and Argmin versions $\underset{\sim}{\beta}^{(n)}_{J_{\rm vdW}}$ and $\underset{\sim}{\beta}^{(n)}_{\rm HL;\, vdW}$ of the van der Waerden R-estimator computed from 1000 replications of model (3.16) with $K = 6, 10, 15$, sample size n=100 and various stable error distributions.

In the particular case of linear models with stable errors (with unspecified tail index α and skewness parameter b), Hallin *et al.* (2010) show how rank-based methods provide a powerful and convenient solution to testing problems. In order to do so, they first establish the local asymptotically normal nature (ULAN, with root-n contiguity rates) of linear model experiments with stable errors. In this paper, we extend their approach to estimation problems. More particularly, taking full advantage of the ULAN property, we construct one-step R-estimators for the regression parameter $\boldsymbol{\beta}$. Those estimators are root-n consistent and asymptotically normal, irrespective of the underlying stable density, and their asymptotic covariance matrices are obtained as a by-product of the one-step procedure. Using numerical results derived in Hallin *et al.* (2010), we moreover show how to construct the R-estimators associated with stable scores, achieving parametric optimality at prespecified values of α and b.

A thorough Monte Carlo study confirms the excellent finite-sample performances of our one-step R-estimators, which are shown to outperform not only the traditional OLS and LAD estimator, but also their Argmin or Hodges-Lehmann counterparts.

5. Appendix.

Proof of Proposition 2.2. Point (i) is a direct consequence of the Hájek projection theorem. Points (ii) and (iii) follow from point (i), the central limit theorem and the Le Cam's Third Lemma. As for point (iv), Theorem 3.1 in Jurečková (1969) applies. □

Proof of Proposition 2.3. In view of (2.9), we have that

$$n^{1/2}(\underset{\sim}{\boldsymbol{\beta}}{}_J^{(n)} - \boldsymbol{\beta}) = n^{1/2}(\hat{\boldsymbol{\beta}}^{(n)}_{\rm LAD} - \boldsymbol{\beta}) + \mathbb{K}^{(n)}\widehat{\mathcal{J}}^{-1}(J,g)\underset{\sim}{\boldsymbol{\Delta}}{}_J^{(n)}(\hat{\boldsymbol{\beta}}^{(n)}_{\rm LAD}). \tag{5.17}$$

The consistency of $\widehat{\mathcal{J}}^{-1}(J,g)$ together with point (iv) of Proposition 2.2 entail that, under $\mathrm{P}^{(n)}_{g,a,\boldsymbol{\beta}}$ with $g \in \mathcal{F}$, as $n \to \infty$,

$$\mathbb{K}^{(n)}\widehat{\mathcal{J}}^{-1}(J,g)\underset{\sim}{\boldsymbol{\Delta}}{}_J^{(n)}(\hat{\boldsymbol{\beta}}^{(n)}_{\rm LAD}) \quad = \quad \mathbb{K}^{(n)}\mathcal{J}^{-1}(J,g)\underset{\sim}{\boldsymbol{\Delta}}{}_J^{(n)}(\boldsymbol{\beta}) - n^{1/2}(\hat{\boldsymbol{\beta}}^{(n)}_{\rm LAD} - \boldsymbol{\beta}) + o_{\mathrm{P}}(1). \tag{5.18}$$

Combining (5.17) and (5.18), we readily obtain

$$n^{1/2}(\underset{\sim}{\boldsymbol{\beta}}{}_J^{(n)} - \boldsymbol{\beta}) = \mathbb{K}^{(n)}\mathcal{J}^{-1}(J,g)\underset{\sim}{\boldsymbol{\Delta}}{}_J^{(n)}(\boldsymbol{\beta}) + o_{\mathrm{P}}(1) \tag{5.19}$$

under $\mathrm{P}^{(n)}_{g,a,\boldsymbol{\beta}}$ with $g \in \mathcal{F}$, as $n \to \infty$. The result follows using Proposition 2.2. □

Proof of Proposition 2.4. Without loss of generality, we assume that the ϵ_i's have median zero. In this proof, we show that $n^{1/2}(\hat{\boldsymbol{\beta}}^{(n)}_{\rm LAD} - \boldsymbol{\beta}) = n^{1/2}(\underset{\sim}{\boldsymbol{\beta}}{}^{(n)}_{J_{\rm L}} - \boldsymbol{\beta}) + o_{\mathrm{P}}(1)$. From the proof of Theorem 4.1 in Koenker (2005) (see also Koenker and Basset 1978), we have that (least absolute deviation estimation is equivalent to median regression hence quantile regression with quantile of order $\tau = 1/2$)

$$n^{1/2}(\hat{\boldsymbol{\beta}}^{(n)}_{\rm LAD} - \boldsymbol{\beta}) = \frac{n^{-1/2}}{2g(0)}\mathbb{K}^{(n)}\mathbb{K}^{(n)\prime}\sum_{i=1}^{n}\mathrm{sign}(\epsilon_i)c_i + o_{\mathrm{P}}(1) \tag{5.20}$$

under $\mathrm{P}^{(n)}_{g,a,\boldsymbol{\beta}}$. Now, since $J_{\rm L}(u) = \sqrt{2}\mathrm{sign}(u - 1/2)$, we have that

$$\begin{aligned}
\mathcal{J}(J_{\rm L}, g) \quad &= \quad \sqrt{2}\int_0^1 \mathrm{sign}(u - 1/2)\varphi_g(G^{-1}(u))du \\
&= \quad -\sqrt{2}\int_{-\infty}^{\infty} \mathrm{sign}(G(v) - G(0))g'(v)dv \\
&= \quad \sqrt{2}\int_{-\infty}^{0} g'(v)dv - \sqrt{2}\int_0^{\infty} g'(v)dv = 2\sqrt{2}g(0).
\end{aligned} \tag{5.21}$$

Using (5.21), (5.19) in the proof of Proposition 2.3 and point (i) of Proposition 2.2, we obtain that

$$
\begin{aligned}
n^{1/2}(\underset{\sim}{\boldsymbol{\beta}}^{(n)}_{J_{\rm L}} - \boldsymbol{\beta}) &= \mathbb{K}^{(n)}\mathcal{J}^{-1}(J_{\rm L}, g)\,\underset{\sim}{\boldsymbol{\Delta}}^{(n)}_{J_{\rm L}}(\boldsymbol{\beta}) + o_{\rm P}(1) \\
&= \mathbb{K}^{(n)}\mathcal{J}^{-1}(J_{\rm L}, g)\boldsymbol{\Delta}^{(n)}_{J_{\rm L}}(\boldsymbol{\beta}) + o_{\rm P}(1) \\
&= \frac{n^{-1/2}}{2g(0)}\mathbb{K}^{(n)}\mathbb{K}^{(n)\prime}\sum_{i=1}^{n} \text{sign}\left(G(\epsilon_i) - \frac{1}{2}\right) c_i + o_{\rm P}(1) \\
&= \frac{n^{-1/2}}{2g(0)}\mathbb{K}^{(n)}\mathbb{K}^{(n)\prime}\sum_{i=1}^{n} \text{sign}\,(\epsilon_i)\, c_i + o_{\rm P}(1),
\end{aligned}
$$

which, in view of (5.20), completes the proof. □

Acknowledgements. The authors thank two anonymous referees for their pertinent comments on their original manuscript, which led to a much improved revised version.

Marc Hallin acknowledges the support of the Sonderforschungsbereich “Statistical modelling of nonlinear dynamic processes (SFB 823) of the Deutsche Forschungsgemeinschaft, and a Discovery Grant of the Australian Research Council. The research of Yvik Swan is supported by a Mandat de Chargé de Recherche from the Fonds National de la Recherche Scientifique, Communauté française de Belgique, and that of Thomas Verdebout by a BQR (Bonus Qualité Recherche) of the Université Lille 3. David Veredas acknowledges the financial support of the IAP P6/07 contract, from the IAP program (Belgian Federal Scientific Policy) “Economic policy and finance in the global economy”.

References

[1] Allal, J., A. Kaaouachi and D. Paindaveine, 2001, R-estimation for ARMA models. Journal of Nonparametric Statistics 13, 815-831.

[2] Andrews, B., M. Calder and R. A. Davis, 2009, Maximum likelihood estimation for α-stable autoregressive processes. The Annals of Statistics 37, 1946-1982.

[3] Bassett, G. and R. Koenker, 1978, Asymptotic theory of least absolute error regression. Journal of the American Statistical Association 73, 618-622.

[4] Blattberg, R. and T. Sargent, 1971, Regression with Non-Gaussian Disturbances: Some Sampling Results. Econometrica 39, 501-510.

[5] Cassart, D., M. Hallin and D. Paindaveine, 2010, On the estimation of cross-information quantities in R-estimation, in: J. Antoch, M. Husková and P.K. Sen, (Eds.), Nonparametrics and Robustness in Modern Statistical Inference and Time Series Analysis: A Festschrift in Honor of Professor Jana Jurečková, I.M.S. Collections, Volume 7, 35-45.

[6] Deo, R.S., 2002, On testing the adequacy of stable processes under conditional heteroskedasticity. Journal of Empirical Finance 9, 257270.

[7] Dominicy, Y. and D. Veredas, 2010, The method of simulated quantiles, ECARES WP 2010/09.

[8] DuMouchel, W. H., 1973, On the asymptotic normality of the maximum-likelihood estimate when sampling from a stable distribution. The Annals of Statistics 1, 948-957.

[9] DuMouchel, W. H., 1975, Stable distributions in statistical inference . 2: Information from stably distributed samples. Journal of the American Statistical Association 70, 386-393.

[10] El Bantli, F. and M. Hallin, 1999, L_1-estimation in linear models with heterogeneous white noise. Statistics and Probability Letters 45, 305-315.

[11] El Barmi, H. and P.I. Nelson, 1997, Inference from stable distributions. Proceedings of the Symposium on Estimating Functions. IMS Lecture Notes-Monograph Series, 32, 439-456.

[12] Fama, E. and R. Roll, 1971, Parameter Estimates for Symmetric Stable Distributions. Journal of the American Statistical Association 66, 331-338.

[13] Hallin, M., H. Oja and D. Paindaveine, 2006, Semiparametrically efficient rank-based inference for shape. II. Optimal R-estimation of shape. The Annals of Statistics 34, 2757-2789.

[14] Hallin, M., Y. Swan, T. Verdebout and D. Veredas, 2010, Rank-based testing in linear models with stable errors. Journal of Nonparametric Statistics, to appear.

[15] Hodges, J. L. and E. L. Lehmann, 1963, Estimation of location based on ranks. The Annals of Mathematical Statistics 34, 598-611.

[16] Jurečková, J., 1969, Asymptotic linearity of a rank statistic in regression parameter. The Annals of Mathematical Statistics 40, 1889-1900.

[17] Jurečková, J., 1971, Nonparametric estimate of regression coefficients. The Annals of Mathematical Statistics 42, 1328-1338.

[18] Jurečková, J. and P. K. Sen, 1996, Robust Statistical Procedures: Asymptotics and Interrelations, Wiley, New York.
[19] Knight, K., 1998, Bootstrapping sample quantiles in non-regular cases. Statistics and Probability Letters 37, 259-267.
[20] Koenker, R., 2005, Quantile Regression, Cambridge University Press, Cambridge.
[21] Koul, H. L., 1971, Asymptotic behavior of a class of confidence regions based on ranks in regression. Annals of Mathematical Statistics 42, 466-476.
[22] Koul, H. L., 2002, Weighted empirical processes in dynamic nonlinear models, Lecture Notes in Statistics 166, Springer, New York.
[23] Koul, H. L. and A. Saleh, 1993, R-estimation of the parameters of autoregressive [AR (p)] models. The Annals of Statistics 21, 534-551.
[24] Lehmann, E. L., 1963, Nonparametric confidence intervals for a shift parameter. The Annals of Mathematical Statistics 34, 1507-1512.
[25] Mittnik, S., M.S. Paolella and S.T. Rachev, 2000, Diagnosing and treating the fat tails in financial returns dat. Journal of Empirical Finance 7, 389416.
[26] Nolan, J. P., 1997, Numerical calculation of stable densities and distribution functions. Communications in Statistics and Stochastic Models 13, 759-774.
[27] Nolan, J. P., 1999, An algorithm for evaluating stable densities in Zolotarev's (M) parametrization. Mathematical and Computer Modelling 29, 229-233.
[28] Samorodnitsky, G., S.T. Rachev, J.R. Kurz-Kim, and S.V. Stoyanov, 2007, Asymptotic distributions of unbiased linear estimators in the presence of heavy-tailed stochastic regressors and residuals. Probability and Mathematical Statistics 27, 275-302.
[29] Samorodnitsky, G., and M. S. Taqqu, 1994, Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance, Chapman and Hall, New York.
[30] Sen, P. K., 1966, On a distribution-free method of estimating asymptotic efficiency of a class of nonparametric tests. Annals of Mathematical Statistics 37, 1759-1770.
[31] Wise, J., 1966, Linear estimators for linear regression systems having infinite residual variances. Unpublished manuscript, Department of Economics, University of Hawaii.
[32] Zolotarev, V. M., 1986, One-dimensional stable distributions, Translation of Mathematical Monographs 65, American Mathematical Society, Providence, RI.
[33] Zolotarev, V. M., 1995, On representation of densities of stable laws by special functions. Theory of Probability and its Applications 39, 354-362.