Rank-based Inference 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, E.C.A.R.E.S., Université Libre de Bruxelles, Belgium.*
[b]*Académie Royale de Belgique and Center, Tilburg University.* [c]*Département de Mathématique,*
Université Libre de Bruxelles, Belgium. [d]*EQUIPPE-GREMARS, Université Lille III, France.*

Abstract

Linear models with stable error densities are considered. The local asymptotic normality of the resulting model is established. We use this result, combined with Le Cam's third lemma, to obtain local powers of various classical rank tests (Wilcoxon's and van der Waerden's test, the median test, and their counterparts for regression and analysis of variance) under α-stable laws. The same results are used to construct new rank tests achieving parametric optimality at specified stable densities. A Monte-Carlo study is conducted to compare their relative performances.

Keywords : Stable distributions, local asymptotic normality, rank tests, asymptotic relative efficiencies.

1 Introduction.

Due to their attractive stochastic properties, *stable distributions* (also called α-*stable*) are extremely appealing from the point of view of stochastic modelling. They are, indeed, the only nondegenerate distributions with a domain of attraction: nontrivial limits of normalized sums of independent identically distributed terms are necessarily stable. They moreover constitute a quite flexible four-parameter family, beyond location and scale, they also account for tail behavior and asymmetries.

Empirical evidence of non-Gaussian stable behavior is present in a variety of fields, among which economics, insurance, finance, signal processing, teletraffic engineering, ... In all those fields, neglecting heavy tail and asymmetry features leads to underestimating risks, sometimes quite significantly, hence taking adventurous decisions that possibly induce severe losses. Student families (generally with three degrees of freedom or more, that is, with finite moments of order two) therefore are quite popular in such areas; Student densities, however, do not enjoy the stability property, and their tails often are still too light, failing to provide a reasonable account of many stylized facts. Stable distributions, often referred to as *stable Paretian* distributions in the financial literature, are among the most frequently proposed candidates for describing the distribution of stock returns: see Fama (1965a and 1965b), Mandelbrot (1963), Mittnik *et al.* (1998, 2000), etc.

Despite their attractiveness, however, stable distributions, in all those areas, still remain largely absent from daily practice. While acknowledging the fact that observed data do exhibit heavy-tailed and skewed behavior, most practitioners stick to classical statistical and econometric methods. The reason for that somewhat schizophrenic attitude probably originates in a lack of well-documented statistical and econometric procedures for models involving stable densities, and the widespread opinion that they are hardly tractable.

That opinion is, to a large extent, unfounded. Contrary to a very common belief, stable families are extremely well-behaved. DuMouchel (1973) long time ago showed that maximum likelihood estimators of the parameter of stable distributions are asymptotically normal, at standard root-n rate. Pushing his results a bit farther, it can be seen that stable families actually are *locally asymptotically normal* (LAN) in the Le Cam sense, so that their asymptotic structure actually resorts to the familiar Gaussian shift type—the most regular and classical statistical experiment one can think of.

Now, putting local asymptotics and likelihood methods into practice is not so straightforward. The main drawback of stable densities, from an applied statistics perspective, indeed, is the lack of closed-form expressions for likelihoods, hence for their log-derivatives, the corresponding maximum likelihood estimators, likelihood ratio statistics, etc. That lack of closed-form expressions is probably the main reason why practitioners remain reluctant about stable error distributions. The advent of powerful and user-friendly symbolic calculators such as Mathematica, however, has transformed the lack of closed-formulas from a major hindrance into a mere technical nuisance (see Borak *et al.* (2005) for an overview). Under the impulsion of Zolotarev, Nolan, and a few others, a number of computationally efficient integral formulas for stable distributions have been derived over the years (Zolotarev (1986 and 1995), Nolan (1997 and 1999)). These formulas are at the heart of the numerical results which are in the present paper.

The absence of moments, on the other hand, makes asymptotic properties hard to derive, and a number of important problems still remain open. Among them is a full description of the asymptotic behavior of Student's t statistic for location (see for instance the remarkable papers by Logan *et al.* (1973) and Giné *et al.* (1997)). Somewhat surprisingly, little is kown about the power (under stable densities) of this statistic, a problem which, to the best of our knowledge, only has been considered by Benktus *et al.* (2007). This lack of results on performances or asymptotic performances, as a rule, makes it difficult to convince practitioners to adopt any particular method rather than any other one.

Alternatives to likelihood-based methods do exist, though, which all too often are ignored. Fama and Roll (1968) for symmetric stable distributions with tail indices between 1 and 2, and Press (1972) for general α values, provide consistent estimators. A quantile-based approach proposed by Fama and Roll (1971) and further enhanced by McCulloch (1986), and the method of indirect estimation described in Garcia *et al.* (2010) are the most commonly used. Estimation in regression models with symmetric stable errors has also been studied, either empirically (Blattberg and Sargent (1971)) or theoretically (El Barmi and Nelson (1997)). More recently, Andrews *et al.* (2009) study the case of autoregressive time series (both causal and noncausal) with non-Gaussian stable noise and show that, in that model, while the estimators of the parameters of the stable noise are again root-n asymptotically normal, those for the autoregressive parameters are $n^{1/\alpha}$-consistent and converge in distribution to the maximizer of a random function.

Another alternative in the context are ranks. Rank-based methods, which are distribution-free, and therefore valid under arbitrary densities, naturally come into the picture in the context. Yet, nobody, it seems, ever considered them. An explanation might be, again, the absence of power or asymptotic efficiency results.

The objective of this paper is twofold:

(i) establishing the local asymptotic normality (LAN) of linear models with α-stable error distributions (that model includes two-sample location, one-way analysis of variance, and regression models with unspecified intercept); that LAN result settles the parametric efficiency bounds for the problem and, combined with Le Cam's third lemma, allows us to obtain local powers and asymptotic relative efficiencies for a variety of inference procedures, including traditional rank tests (Wilcoxon, van der Waerden or sign tests);

(ii) analyzing the asymptotic performances of classical rank tests and deriving new ones (Cauchy or α-stable score tests) adapted to stable error distributions.

Section 2 presents some properties of stable distributions, and derives a result on the asymptotic behavior of stable score functions, to be used in the sequel. Section 3.1 establishes the local asymptotic normality property we need to study local asymptotic powers. Section 3.2 studies the performances of classical rank tests in the stable context. Asymptotic relative efficiencies are considered in Section 3.3, not only for the classical rank tests, but also for some new rank-based tests achieving the parametric efficiency bounds at some specified stable densities. Numerical values of those AREs are provided. A simulation study is conducted in Section 4, while concluding remarks are given in Section 5. An appendix collects the proofs.

2   Stable distributions: some properties

Denote by $f_{\boldsymbol{\theta}}$ or $f_{\alpha,\beta,\gamma,\delta}$ the α-stable density function with parameter $\boldsymbol{\theta} := (\alpha, \beta, \gamma, \delta)'$. Writing $f_{\alpha,\beta}$ for the standardized density $f_{\alpha,\beta,1,0}$, we have

$$f_{\alpha,\beta,\gamma,\delta}(x) = \frac{1}{\gamma} f_{\alpha,\beta}\left(\frac{x-\delta}{\gamma}\right),\tag{2.1}$$

which characterizes the roles of δ and γ as location and scale parameters, respectively. The interpretation of β as a *skewness parameter* is justified by the fact that, for $\beta = 0$, the distribution is symmetric about δ and, for $\beta = 1$ (resp., $\beta = -1$), the distribution is totally skewed to the right (resp. to the left). The notation $F_{\boldsymbol{\theta}}$ and $F_{\alpha,\beta,\gamma,\delta}$ will be used for the corresponding distribution function.

Certain choices of $\boldsymbol{\theta}$ yield well-known distributions, namely the Gaussian ($\alpha = 2$, any β), the Cauchy ($\alpha = 1$ and $\beta = 0$) and the Lévy ($\alpha = 1/2$ and $\beta = 1$). However, together with the reflection of the Lévy distribution, these are the *only* instances of stable distributions whose density functions can be written explicitly in terms of elementary functions. For all other choices of the parameters, there exist no closed form for $f_{\boldsymbol{\theta}}$, and stable distributions are either defined in terms of the inverse Fourier transform of their characteristic functions, or in the form of integral formulas (see e.g. Nolan (1997) or Zolotarev (1986)).

A number of alternative parameterizations have been proposed in the literature, each having its own advantages and some drawbacks (see Nolan (2010) for an overview). In the sequel, we are working with Zolotarev's so-called M-parametrization, under which the logarithm of the characteristic function $\phi(t; \alpha, \beta, \gamma, \delta)$ of $f_{\alpha,\beta,\gamma,\delta}$ is defined as

$$\log \phi(t; \alpha, \beta, \gamma, \delta) := \begin{cases} it\delta - |\gamma t|^\alpha \left(1 - i\zeta \operatorname{sign}(\mathrm{t})(|\gamma t|^{1-\alpha} - 1)\right) & \text{if } \alpha \neq 1, \\ it\delta - |\gamma t|^\alpha \left(1 + i\beta \operatorname{sign}(\mathrm{t})(2/\pi) \log |\gamma t|\right) & \text{if } \alpha = 1, \end{cases}\tag{2.2}$$

with $\zeta := \zeta(\alpha, \beta) := -\beta\tan(\pi\alpha/2)$. Note that, in this parameterization, $f_{\alpha,\beta}(x) = f_{\alpha,-\beta}(-x)$ for $x > \zeta$ and all admissible (α, β).

It is well-known that *"all (non-degenerate) stable distributions are continuous distributions with an infinitely differentiable density"* (Theorem 1.9 in Nolan (2010)). Hence, from (2.1), these distributions also are continuously differentiable in the location parameter δ. The tail behavior of these distributions is given, for $|x| \to \infty$, by

$$1 - F_{\boldsymbol{\theta}}(|x|) \sim C_{\boldsymbol{\theta}}|x|^{-\alpha}, \tag{2.3}$$

where $C_{\boldsymbol{\theta}}$ is a constant depending on $\boldsymbol{\theta}$ and the sign of x. Relation (2.3), by the way, justifies the terminology *tail index* for the parameter α. Using (2.3), we can further prove the following (see the Appendix for a proof).

Lemma 2.1 *Define, for all admissible $\boldsymbol{\theta}$, the score function $\varphi_{\boldsymbol{\theta}}(x) := -\dot{f}_{\boldsymbol{\theta}}(x)/f_{\boldsymbol{\theta}}(x)$. Then, for all $\alpha < 2$, $\varphi_{\boldsymbol{\theta}}(x) = O(1/x)$ as $x \to \infty$.*

For an overview of other properties of these distributions, see for instance Feller (1971), Samorodnitsky and Taqqu (1994) or Nolan (2010).

3   Rank-based tests for linear models under stable noise.

Denote by $\mathcal{H}_{\boldsymbol{\theta}}^{(n)}(\boldsymbol{\beta})$ the hypothesis under which a vector $\mathbf{X}^{(n)} := (X_1^{(n)}, \ldots, X_n^{(n)})'$ of observations satisfies the equation

$$X_i^{(n)} = a + \sum_{k=1}^{K} c_{ik}^{(n)}\beta_k + \epsilon_i^{(n)}, \; i = 1, \ldots, n, \tag{3.1}$$

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

This model (which encompasses the two-sample and one-factor analysis of variance models) is of obvious theoretical and practical importance, and its analysis, under second-order moment assumptions, is covered in all standard statistical textbooks. Yet, the stable case, with infinite variances, remains much less explored. In particular, relatively little is known about the efficiency, under stable densities, of testing procedures for the null hypothesis $\mathcal{H}_{\boldsymbol{\theta}}^{(n)}(\boldsymbol{\beta})$. In this paper, we are providing such efficiency results for the class of rank-based tests. Such tests indeed naturally come into the picture, as their validity extends to completely unspecified error densities, among which the stable ones, irrespective of the actual values of α, β and γ.

3.1   Local Asymptotic Normality

The main theoretical tool throughout will be the local asymptotic normality (LAN) property of the regression model (3.1) under stable error densities. Such a property, which is of independent interest, has never been considered, to the best of our knowledge, in the context of stable error distributions.

Letting $\bar{c}_k^{(n)} := n^{-1}\sum_{i=1}^{n} c_{ik}^{(n)}$, $\mathbf{c}_i^{(n)} := (c_{i1}^{(n)}, \ldots, c_{iK}^{(n)})'$, $\mathbb{C}^{(n)} := n^{-1}\sum_{i=1}^{n} \mathbf{c}_i^{(n)}\mathbf{c}_i^{(n)'}$, and $\mathbb{K}^{(n)} := \left(\mathbb{C}^{(n)}\right)^{-1/2}$, suppose that the following technical assumptions are fulfilled.

4

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

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

Denote by $\mathrm{P}_{\boldsymbol{\theta}, a, \boldsymbol{\beta}}^{(n)}$ the probability distribution of $\mathbf{X}^{(n)}$ under (3.1), and let

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

be the residual for the value $\boldsymbol{\beta}$ of the parameter: under $\mathrm{P}_{\boldsymbol{\theta}, a, \boldsymbol{\beta}}^{(n)}$, the $Z_i^{(n)}(\boldsymbol{\beta})$'s are i.i.d. with density $f_{(\alpha, \beta, \gamma, 0)}$. The following then holds (see the Appendix for a proof).

Proposition 3.1 *(LAN) Suppose that Assumptions (A1) and (A2) hold. Let $\boldsymbol{\nu}(n) := n^{-\frac{1}{2}} \mathbb{K}^{(n)}$ and fix $\boldsymbol{\theta} = (\alpha, \beta, \gamma, 0) \in \boldsymbol{\Theta}$, $a \in \mathbb{R}$. Then, model (3.1), with parameter $\boldsymbol{\beta} \in \mathbb{R}^K$, is LAN. More precisely, for all $\boldsymbol{\beta} \in \mathbb{R}^K$, and all bounded sequence $\boldsymbol{\tau}^{(n)} \in \mathbb{R}^K$,*

$$
\Lambda_{\boldsymbol{\theta}, \boldsymbol{\beta} + \boldsymbol{\nu}(n)\boldsymbol{\tau}^{(n)}}^{(n)} \ := \ \log \frac{d\mathrm{P}_{\boldsymbol{\theta}, a, \boldsymbol{\beta} + \boldsymbol{\nu}(n)\boldsymbol{\tau}^{(n)}}^{(n)}}{d\mathrm{P}_{\boldsymbol{\theta}, a, \boldsymbol{\beta}}^{(n)}} = \log \frac{\prod_{t=1}^n f_{\boldsymbol{\theta}} \left(Z_t^{(n)} \left(\boldsymbol{\beta} + \boldsymbol{\nu}(n)\boldsymbol{\tau}^{(n)} \right) \right)}{\prod_{t=1}^n f_{\boldsymbol{\theta}} \left(Z_t^{(n)}(\boldsymbol{\beta}) \right)}
$$

$$
= \ \boldsymbol{\tau}^{(n)'} \boldsymbol{\Delta}_{\boldsymbol{\theta}}^{(n)}(\boldsymbol{\beta}) - \tfrac{1}{2} \boldsymbol{\tau}^{(n)'} \boldsymbol{\tau}^{(n)} \mathcal{I}(\boldsymbol{\theta}) + o_P(1)
$$

under $\mathcal{H}_{\boldsymbol{\theta}}^{(n)}(\boldsymbol{\beta})$ as $n \to \infty$, where, setting $\varphi_{\boldsymbol{\theta}}(\cdot) := -\dot{f}_{\boldsymbol{\theta}}(\cdot)/f_{\boldsymbol{\theta}}(\cdot)$ and $\mathcal{I}(\boldsymbol{\theta}) := \int_{-\infty}^{\infty} \varphi_{\boldsymbol{\theta}}^2(x) f_{\boldsymbol{\theta}}(x) dx, \mathcal{I}(\boldsymbol{\theta})\mathbf{I}_K$ is the information matrix and

$$
\boldsymbol{\Delta}_{\boldsymbol{\theta}}^{(n)}(\boldsymbol{\beta}) = n^{-1/2} \left(\mathbb{K}^{(n)} \right)' \sum_{i=1}^n \varphi_{\boldsymbol{\theta}} \left(Z_i^{(n)}(\boldsymbol{\beta}) \right) \mathbf{c}_i^{(n)} \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \mathcal{I}(\boldsymbol{\theta})\mathbf{I}_K), \tag{3.2}
$$

the central sequence.

The LAN property is stated here under stable distributions, but actually holds, with information matrix $\mathcal{I}_f(\boldsymbol{\theta})\mathbf{I}_K$ and central sequence $\boldsymbol{\Delta}_f^{(n)}(\boldsymbol{\beta})$, for a much broader class of densities f. A sufficient condition is the mean-square differentiability of $f^{1/2}$ (see, e.g., van der Vaart (2000)). The information matrix and the central sequence then take the same form as in Proposition 3.1, but with score φ of the form $\varphi_f = -2s f^{-1/2}$, where s stands for the mean square derivative of $f^{1/2}$.

Local Asymptotic Normality is the key result allowing us to study the asymptotic behavior of rank-based tests under stable densities via the powerful machinery of the asymptotic theory of statistical experiments associated with the name of Le Cam (as described, for example, in Le Cam and Yang (2000), van der Vaart (2000) or Lehman and Romano (2005)). The application of that theory, however, here is not as straightforward as usual, due to the fact that the central sequence $\boldsymbol{\Delta}_{\boldsymbol{\theta}}^{(n)}(\boldsymbol{\beta})$ and the information quantity $\mathcal{I}(\boldsymbol{\theta})$, for general values of α and β, cannot be expressed under closed form.

3.2 Rank-based tests when the intercept is a nuisance parameter.

Consider the problem of testing the null hypothesis $\mathcal{H}^{(n)}(\boldsymbol{\beta}_0)$ under which $\boldsymbol{\beta} = \boldsymbol{\beta}_0$ in (3.1), where $\boldsymbol{\beta}_0$ is some fixed real K-vector, while the constant term a and the density of $\epsilon_i^{(n)}$, hence that of $a + \epsilon_i^{(n)}$ remain completely unspecified. Denote by $R_i^{(n)}$ the rank of $Z_i^{(n)}(\boldsymbol{\beta}_0)$ among $Z_1^{(n)}(\boldsymbol{\beta}_0), \ldots, Z_n^{(n)}(\boldsymbol{\beta}_0)$; note that this rank does not depend on the unspecified intercept a.

The idea of using ranks in that context is quite natural, particularly so when heavy tails are suspected. The vector $\mathbf{R}^{(n)} = \mathbf{R}^{(n)}(\boldsymbol{\beta}_0) = (R_1^{(n)}, \ldots, R_n^{(n)})$ is not only distribution-free under the null hypothesis $\mathcal{H}^{(n)}(\boldsymbol{\beta}_0)$, it is maximal invariant under a generating group of $\mathcal{H}^{(n)}(\boldsymbol{\beta}_0)$. Distribution-freeness makes it possible to construct rank-based tests with exact finite-sample critical values, while maximal invariance allows achieving parametric efficiency at some selected reference density (see Hallin and Werker (2003)).

The rank-based statistics we are considering are of the form

$$\underset{\sim}{\boldsymbol{\Delta}}_J^{(n)}(\boldsymbol{\beta}_0) := n^{-\frac{1}{2}} \left(\mathbb{K}^{(n)} \right)' \sum_{i=1}^n J \left(\frac{R_i^{(n)}}{n+1} \right) \mathbf{c}_i^{(n)}, \tag{3.3}$$

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

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

The vector $\underset{\sim}{\boldsymbol{\Delta}}_J^{(n)}(\boldsymbol{\beta}_0)$ has the same structure as the central sequence (3.2). Actually, letting

$$\boldsymbol{\Delta}_J^{(n)}(\boldsymbol{\beta}_0) := n^{-\frac{1}{2}} \left(\mathbb{K}^{(n)} \right)' \sum_{i=1}^n J \left(F \left(Z_i^{(n)}(\boldsymbol{\beta}_0) \right) \right) \mathbf{c}_i^{(n)},$$

where F is the unspecified distribution function of the $\epsilon_i^{(n)}$'s, the classical Hájek Projection Theorem for linear rank statistics (see Hájek and Sidák (1967) or Puri and Sen (1985)) implies that, under Assumptions (A1), (A2) and (B),

$$\underset{\sim}{\boldsymbol{\Delta}}_J^{(n)}(\boldsymbol{\beta}_0) - \boldsymbol{\Delta}_J^{(n)}(\boldsymbol{\beta}_0) = o_P(1) \quad \text{as } n \to \infty, \text{ under } \mathcal{H}^{(n)}(\boldsymbol{\beta}_0), \tag{3.4}$$

and that moreover

$$\underset{\sim}{\boldsymbol{\Delta}}_J^{(n)}(\boldsymbol{\beta}_0) \xrightarrow{\mathcal{L}} \mathcal{N}\left(\mathbf{0}, \mathcal{J}(J)\mathbf{I}_K\right) \tag{3.5}$$

with $\mathcal{J}(J) := \int_0^1 J^2(u)du$. Clearly, (3.4) and the fact that central sequences are only defined up to $o_P(1)$'s imply that, for $J(u) = J_f(u) := \varphi_f(F^{-1}(u))$, $\underset{\sim}{\boldsymbol{\Delta}}_J^{(n)}(\boldsymbol{\beta}_0)$ is a rank-based version of the central sequence $\boldsymbol{\Delta}_f^{(n)}(\boldsymbol{\beta}_0)$.

It also follows that, still under $\mathcal{H}^{(n)}(\boldsymbol{\beta}_0)$, the test statistic

$$\underset{\sim}{Q}_J^{(n)}(\boldsymbol{\beta}_0) := \mathcal{J}(J)^{-1}(\underset{\sim}{\boldsymbol{\Delta}}_J^{(n)}(\boldsymbol{\beta}_0))'(\underset{\sim}{\boldsymbol{\Delta}}_J^{(n)}(\boldsymbol{\beta}_0)) \tag{3.6}$$

is asymptotically (central) chi-square with K degrees of freedom; if moreover $J = J_f$ and the actual density of the $\epsilon_i^{(n)}$'s is f, then

$$\begin{aligned} \underset{\sim}{Q}_f^{(n)}(\boldsymbol{\beta}_0) &:= \underset{\sim}{Q}_{J_f}^{(n)}(\boldsymbol{\beta}_0) = \mathcal{J}(J)^{-1} (\boldsymbol{\Delta}_f^{(n)}(\boldsymbol{\beta}_0))'(\boldsymbol{\Delta}_f^{(n)}(\boldsymbol{\beta}_0)) + o_P(1) \tag{3.7} \\ &=: Q_f^{(n)}(\boldsymbol{\beta}_0) + o_P(1), \end{aligned}$$

where $Q_f^{(n)}(\boldsymbol{\beta}_0)$ is the test statistic yielding, under error density f, the *locally asymptotically maximin* test for $\mathcal{H}^{(n)}(\boldsymbol{\beta}_0)$ (see, e.g. Section 11.9 of Le Cam (1986))—a property the rank test based on $\underset{\sim}{Q}\,_f^{(n)}(\boldsymbol{\beta}_0)$, in view of (3.7), automatically inherits, with the substantial additional advantage, though, that, unlike $Q_f^{(n)}(\boldsymbol{\beta}_0)$, it is distribution-free under the null, irrespective of the actual underlying density, be it a stable density.

Finally, consider a local alternative of the form $\boldsymbol{\beta}_0 + n^{-1/2}\boldsymbol{\tau}$. A standard application of Le Cam's third lemma (see van der Vaart (2000)) yields the following result (see the Appendix for a proof).

Proposition 3.2 *Set $\mathcal{J}(J,\boldsymbol{\theta}) := \int_0^1 J(u)\varphi_{\boldsymbol{\theta}}(F_{\boldsymbol{\theta}}^{-1}(u))du$, and let Assumptions (A1), (A2) and (B) hold. Then, under $\mathcal{H}_{\boldsymbol{\theta}}^{(n)}(\boldsymbol{\beta}_0 + n^{-\frac{1}{2}}\boldsymbol{\tau})$, as $n \to \infty$,*

(i) $\underset{\sim}{\boldsymbol{\Delta}}\,_J^{(n)}(\boldsymbol{\beta}_0) \overset{\mathcal{L}}{\longrightarrow} \mathcal{N}\left(\mathcal{J}(J,\boldsymbol{\theta})\boldsymbol{\tau}, \mathcal{J}(J)\mathbf{I}_K\right)$, and

(ii) $\underset{\sim}{Q}\,^{(n)}(\boldsymbol{\beta}_0)$ is asymptotically noncentral chi-square, with K degrees of freedom and noncentrality parameter $\boldsymbol{\tau}'\boldsymbol{\tau}\mathcal{J}^2(J,\boldsymbol{\theta})/\mathcal{J}(J)$.

Here again, the results are stated for stable densities, but hold for all densities f for which the model is locally asymptotically normal.

3.3 Local powers and asymptotic relative efficiencies.

The validity of the tests (3.6) extends to any underlying error density irrespective of the choice of score $J(\cdot)$. Moreover, under stable density $f_{\boldsymbol{\theta}}$, such a test, with score function

$$J_{\boldsymbol{\theta}}(x) = -f_{\boldsymbol{\theta}}'(F_{\boldsymbol{\theta}}^{-1}(x))/f_{\boldsymbol{\theta}}(F_{\boldsymbol{\theta}}^{-1}(x)), \tag{3.8}$$

will be asymptotically optimal. The applicability of this optimality result is, however, hampered by the (theoretical and practical) difficulties implied by the fact that the parameters $\boldsymbol{\theta}$ of stable distributions in applications are unknown, and are not easily estimated. It may be more interesting, therefore, to dispose of valid tests whose power is known to be "not too bad" for a broad range of $\boldsymbol{\theta}$ values, and hence do not require estimation of $\boldsymbol{\theta}$. The objective of this section is to describe some of such tests and to compare their powers, by means of their asymptotic relative efficiency (ARE).

Some of the most common rank-based tests are those reaching optimality under the Gaussian, the logistic and the double-exponential distributions, respectively; they are known as the van der Waerden, Wilcoxon and Laplace tests, and based on the score-generating functions

$$J_1(x) = \Phi^{-1}(x), \; J_2(x) = \frac{\pi}{\sqrt{3}}(2x-1), \; \text{and } J_3(x) = \sqrt{2}\,\text{sign}(\text{F}^{-1}(\text{x})), \tag{3.9}$$

where Φ denotes, as usual, the cdf of the standardized normal distribution, and F that of the double-exponential distribution with variance 1. One of course may like to choose the scores $J_{\boldsymbol{\theta}}$ associated (see (3.8)) with the stable density $f_{\boldsymbol{\theta}}$. For instance, setting $\boldsymbol{\theta} = (1,0,1,0)$ yields the Cauchy score-generating function

$$J_4(x) = \sin(2\pi(x - 1/2)). \tag{3.10}$$

The only other stable distribution for which the score-generating function can be written in closed form is the Lévy distribution (i.e. $\boldsymbol{\theta} = (.5, 1, 1, 0)$), with

$$J_5(x) = \sqrt{2}\left(\Phi^{-1}((x+1)/2)\right)^2\left(3 - 2\sqrt{2}\left(\Phi^{-1}((x+1)/2)^2\right).$$

For all other choices of tail indices, however, a numerical derivation of $J_{\boldsymbol{\theta}}$ is possible, characterizing new tests with good asymptotic properties under strongly skewed and/or heavy-tailed errors.

Recall that asymptotic relative efficiency (ARE) between two testing procedures is the limit, if it exists, of the ratio of the sample sizes they need in order to achieve comparable asymptotic performances. Obtaining such AREs is in general a difficult task unless the model under study is LAN. In the present case, denoting by $\text{ARE}_{\boldsymbol{\theta}}(J/\tilde{J})$ the asymptotic relative efficiency, under stable density $f_{\boldsymbol{\theta}}$, of the rank-based test based on the score-generating function J with respect to the rank-based test based on the score-generating function $\tilde{J}$, the following result holds.

Proposition 3.3 *Let J and $\tilde{J}$ be two score-generating functions satisfying Assumption (B). Then,*

$$ARE_{\boldsymbol{\theta}}(J/\tilde{J}) = \mathcal{J}^2(J, \boldsymbol{\theta})\mathcal{J}(\tilde{J})/\mathcal{J}^2(\tilde{J}, \boldsymbol{\theta})\mathcal{J}(J). \tag{3.11}$$

Numerical evaluations of (3.11) is all but trivial, but can be obtained via inversion of the characteristic function (2.2) and other techniques such as the Fast Fourier Transform or the evaluation of complicated integral transforms such as those given in Nolan (1997). Details are discussed in the Appendix.

Figure 1 provides plots of ARE values for the Wilcoxon, Laplace and Cauchy tests with respect to the van der Waerden ones as functions of the tail index α, for a choice of values of the skewness parameter β ranging from 0 to 1. Note that these AREs are symmetric in β, so that there is no need to consider negative values of the skewness parameter.

Turning to the optimal scores for stable densities with tail parameters $\alpha = 1.6, 1.7, 1.8$, and 1.9, and skewness parameters $|\beta| = 0, 0.2$, and 0.4, the AREs of the corresponding rank tests, still with respect to van der Waerden's, are provided in Table 1. More complete tables as well as Mathematica and R programs performing the various numerical evaluations can be found at the address http://homepages.ulb.ac.be/~yvswan/.

Figure 1 reveals that for $|\beta|$ close to 1, van der Waerden tests are uniformly most powerful among the four tests under study. On the other hand, when the errors are symmetric, the three other contenders appear to be more powerful than van der Waerden, except for tail parameters that are very close to the Gaussian (at $\alpha \approx 1.97$), where van der Waerden regains leadership. Also note that, near the symmetric case $|\beta| \approx 0$, the Cauchy and the Laplace are doing extremely well when α is close to 1, with the Cauchy test being best, as expected. Moreover, for $\alpha \to 0$, AREs appear to become (a) arbitrarily close to zero in favor of van der Waerden tests as $|\beta| \to 1$ and (b) arbitrarily large in favor of Cauchy and Laplace tests as $\beta \to 0$. Interestingly (b) does not hold for the AREs of Wilcoxon with respect to van der Waerden. Actually, it can be shown that the latter is uniformly less than $6/\pi$, a limit which is attained under symmetric α-stable densities with $\alpha \approx 0.01$. Finally note that, for fixed $\alpha > 0.5$, all three ARE curves are strictly increasing with β. This also holds true for the first two when α nears 0, whereas, in the Cauchy case, some oscillation seems to occur.

As for Table 1, it shows that the optimal tests considered there outperform van der Waerden's under the stable density they were devised for, which is not a surprise, but also, quite interestingly, under alternative values of α and β (in the range of values considered there). The lack of sensitivity to variations of the skewness parameter β is particularly remarkable.
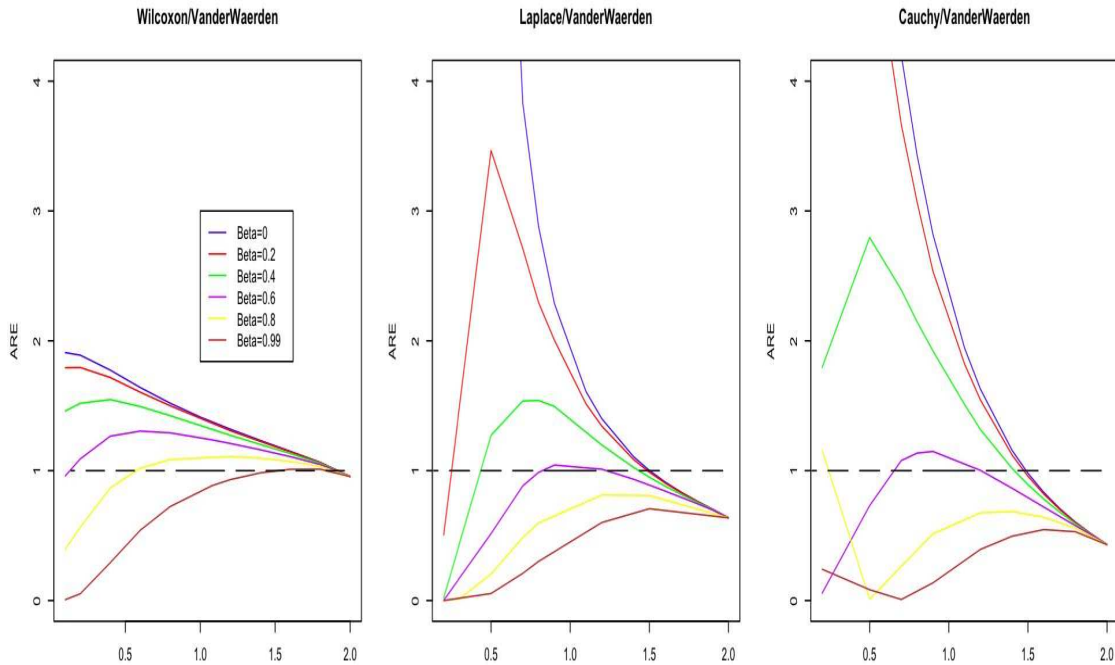
Figure 1: AREs of Wilcoxon, Laplace and Cauchy with respect to van der Waerden as functions of the tail index α, for various values of the skewness parameter β.

$\|\beta\|$	$\alpha =1.6$			$\alpha =1.7$			$\alpha =1.8$			$\alpha =1.9$		
	0	0.2	0.4	0	0.2	0.4	0	0.2	0.4	0	0.2	0.4
$\alpha =1.6$												
0	1.2127	1.2045	1.1787	1.1332	1.1269	1.1075	1.0446	1.0407	1.0288	0.9444	0.9429	0.9386
0.2	1.2043	1.2129	1.2039	1.1277	1.1333	1.1256	1.0416	1.0447	1.0396	0.9433	0.9445	0.9428
0.4	1.1779	1.2033	1.2135	1.1100	1.12811	1.1337	1.0320	1.0425	1.0450	0.9396	0.9438	0.9450
$\alpha =1.7$												
0	1.2010	1.1954	1.1772	1.1442	1.1394	1.1241	1.0756	1.0721	1.0615	0.9893	0.9879	0.9834
0.2	1.1942	1.2011	1.1962	1.1393	1.1444	1.1392	1.0727	1.0757	1.0714	0.9882	0.9894	0.9876
0.4	1.1731	1.1925	1.2017	1.1236	1.1387	1.1448	1.0631	1.0730	1.0759	0.9845	0.9886	0.9897
$\alpha =1.8$												
0	1.1645	1.1614	1.1511	1.1313	1.1284	1.1188	1.0878	1.0852	1.0771	1.0240	1.0226	1.0182
0.2	1.1600	1.1647	1.1628	1.1276	1.1315	1.1291	1.0852	1.0879	1.0852	1.0228	1.0240	1.0222
0.4	1.1465	1.1587	1.1654	1.1161	1.1267	1.1319	1.0767	1.0849	1.0881	1.0190	1.0231	1.0242
$\alpha =1.9$												
0	1.1005	1.0994	1.0957	1.0878	1.0867	1.0831	1.0704	1.0693	1.0656	1.0405	1.0394	1.0361
0.2	1.0988	1.1008	1.1006	1.0862	1.0880	1.0876	1.0689	1.0706	1.0698	1.0394	1.0405	1.0394
0.4	1.0937	1.0988	1.1019	1.0813	1.0860	1.0887	1.0643	1.0686	1.0709	1.0360	1.0393	1.0406

Table 1: AREs for tests based on stable scores with respect to van der Waerden's. Rows correspond to scores, columns to the (stable) densities under which AREs are computed. For instance, row 1 contains the AREs with respect to van der Waerden of the test based on stable scores for $\alpha = 1.6$, $\beta = 0$, under stable densities with tail parameter $\alpha = 1.6$ and skewness β ranging from 0 through 0.4.

4 Monte Carlo simulations.

This section is devoted to the validation, through Monte Carlo simulations, of the results obtained in the previous sections. Our aim here is, on the one hand, to verify the ARE rankings between the proposed rank-based tests and, on the other hand, to compare such tests with the classical Student t-test whose behavior is well known to be (very) complicated to derive when confronted with an α-stable universe (see, for instance, Efron (1969), Logan *et al.* (1973), or Giné *et al.* (1997)).

We generated $N = 2500$ samples from the regression models

$$Y_i^{(l)} = ((l/20))c_i + \epsilon_i, \quad i = 1, \dots, n = 100, \ \ l = 0, \ 1, \ 2, \ 3, \tag{4.1}$$

where the the ϵ_i's are i.i.d. with centered alpha-stable distribution. The regression constants c_i ($i = 1, \dots, 100$) (the same ones across the 2500 replications) were drawn from the uniform distribution on $[-5, 5]$.

Throughout, we performed the various tests at nominal level 5% for the null hypothesis under which the slope β of the model is equal to 0. Observations $Y_i^{(0)}$ thus are generated under the null, $Y_i^{(1)}$, $Y_i^{(2)}$ and $Y_i^{(3)}$ under increasing alternatives of the form $\beta = l/20$, $l = 1, 2$ and 3. Critical values were computed from asymptotic distributions. We recorded, for each situation (the null and the three different alternatives) the rejection frequencies of the different tests over the $N = 2500$ replications. Results, for different (alpha-stable) densities, are collected in Table 4; more extensive evaluations ($\beta = l$, $l = 0$ through 20) are presented, for the Student and van der Waerden tests, in Figure 4.

Inspection of Table 4 confirms that all rank-based tests appear to satisfy the 5% probability level constraint. They are conservative in their original versions (particularly so for van der Waerden scores), hence slightly biased. This (hence, also their powers) could be improved, though, by considering exact finite-sample critical values; such values easily can be obtained by sampling the uniform distribution of ranks over the $n!$ permutations of $(1, \dots, n)$. Rejection frequency rankings are consistent with the corresponding ARE values.

A more intriguing fact is the apparent asymptotic validity-robustness of the Student t test. When the tail index is close to 2, this test seems to reach the 5 % nominal size. This cannot be true: indeed, Giné et al. (1997) showed that the t statistic is asymptotically normal *if and only if* the observations are in the domain of attraction of the Gaussian distribution (actually, they show this for the location problem, but their result is very likely to hold also in the present case). The QQ plots of Figure 3 further confirm the impression that t tests under stable densities are increasingly conservative, hence increasingly biased as α decreases. We could not find in the literature any theoretical confirmation of that fact—see, however, Efron (1969) and Jensen (1979) for some hints. The QQ plots in Figure 3 were obtained from 10,000 replications of i.i.d. samples ($n = 1,000$) of stable observations for various values of α and $\beta = 0$; the same regressors were used throughout.

Now, even if Student tests were valid at usual probability levels, this does not imply that they are providing any recommendable testing procedure in this context. The bar charts in Figure 4 and the power curves in Figure 4, where their rejection frequencies are compared to those of the corresponding van der Waerden tests, show that, although Student tests apparently have the right probability level under the null, their powers are miserable—unless, of course, α is close to two.

test	density	l				density	l			
		0	1	2	3		0	1	2	3
ϕ_{vdW}		.0416	.1728	.3968	.5824		.0420	.2324	.6116	.8728
ϕ_{W}		.0488	.2600	.5712	.7724		.0484	.3176	.7700	.9564
ϕ_{L}	$\alpha = .5$	.0500	.5992	.9032	.9780	$\alpha = .85$	.0476	.4600	.9084	.9908
ϕ_{C}	$\beta = 0$	.0496	.5304	.8576	.9500	$\beta = 0$	.0472	.4304	.8744	.9740
$\phi_{1.6;0}$		.0532	.2916	.6180	.8120		.0516	.3568	.8224	.9720
ϕ_{t}		.0164	.0244	.0240	.0204		.0288	.0344	.0516	.0872
ϕ_{vdW}		.0416	.2004	.4204	.6164		.0408	.2484	.6452	.8836
ϕ_{W}		.0500	.2784	.5784	.7752		.0428	.3388	.7752	.9580
ϕ_{L}	$\alpha = .5$	.0484	.3236	.6980	.8812	$\alpha = .85$	.0472	.3520	.8136	.9716
ϕ_{C}	$\beta = .4$	.0480	.1856	.3956	.5480	$\beta = .4$	.0492	.1932	.5124	.7632
$\phi_{1.6;0}$		.0508	.3124	.6444	.8244		.0476	.3744	.8300	.9772
ϕ_{t}		.0196	.0224	.0196	.0212		.0360	.0420	.0528	.0764
ϕ_{vdW}		.0396	.3472	.6668	.8176		.0424	.3488	.8020	.9604
ϕ_{W}		.0448	.2732	.5992	.7684		.0476	.3216	.7748	.9524
ϕ_{L}	$\alpha = .5$	.0448	.1028	.2224	.4036	$\alpha = .85$	0496	.1784	.5248	.8188
ϕ_{C}	$\beta = .99$	.0420	.1880	.2120	.1776	$\beta = .99$	.0480	.0500	.0544	.0780
$\phi_{1.6;0}$		.0428	.2280	.5396	.7336		.0488	.2884	.7444	.9452
ϕ_{t}		.0136	.0224	.0192	.0252		.0328	.0376	.0424	.0688
ϕ_{vdW}		.0424	.3964	.9208	.9968		.0340	.4488	.9556	.9996
ϕ_{W}		.0512	.4580	.9540	.9992		.0416	.4848	.9660	.9996
ϕ_{L}	$\alpha = 1.6$	.0516	.3724	.9004	.9972	$\alpha = 1.8$	.0428	.3740	.9028	.9984
ϕ_{C}	$\beta = 0$	.0488	.2788	.7400	.9624	$\beta = 0$	.0440	.2392	.7044	.9520
$\phi_{1.6;0}$		.0580	.4864	.9624	.9996		.0432	.4880	.9680	.9996
ϕ_{t}		.0436	.2700	.6948	.8700		.0468	.4052	.8720	.9692
ϕ_{vdW}		.0396	.3972	.9208	.9988		.0364	.4436	.9600	1.000
ϕ_{W}		.0440	.4512	.9548	1.000		.0444	.4860	.9724	1.000
ϕ_{L}	$\alpha = 1.6$	.0492	.3568	.8952	.9956	$\alpha = 1.8$	.0508	.3832	.9024	1.000
ϕ_{C}	$\beta = .4$	.0552	.2164	.6476	.9228	$\beta = .4$	.0536	.2120	.6616	.9312
$\phi_{1.6;0}$		.0460	.4676	.9628	1.000		.0468	.4944	.9752	1.000
ϕ_{t}		.0464	.2836	.6848	.8748		.0468	.4064	.8844	.9664
ϕ_{vdW}		.0392	.4404	.9504	.9992		.0372	.4408	.9728	1.000
ϕ_{W}		.0492	.4584	.9532	.9992		.0400	.4624	.9768	1.000
ϕ_{L}	$\alpha = 1.6$	.0524	.3332	.8684	.9948	$\alpha = 1.8$	.0480	.3440	.8976	.9976
ϕ_{C}	$\beta = .99$	.0500	.1352	.4172	.7512	$\beta = .99$	.0464	.1736	.5440	.8708
$\phi_{1.6;0}$		.0552	.4392	.9472	.9988		.0408	.4608	.9676	1.000
ϕ_{t}		.0440	.2824	.7120	.8664		.0496	.3916	.8800	.9696
ϕ_{vdW}		.0488	.5028	.9836	1.000					
ϕ_{W}		.0544	.5048	.9816	1.000					
ϕ_{L}	$\alpha = 2$	.0512	.3728	.9032	.9968					
ϕ_{C}		.0564	.2096	.6284	.9420					
$\phi_{1.6;0}$		.0604	.4812	.9740	1.000					
ϕ_{t}		.0580	.5540	.9888	1.000					

Table 2: Rejection frequencies (out of $2,500$ replications), under the null ($l = 0$) and under alternatives ($l = 1, 2, 3$), of the van der Waerden test ϕ_{vdW}, the Wilcoxon test ϕ_{W}, the Laplace test (the sign test) ϕ_{L}, the Cauchy test ϕ_{C}, the test $\phi_{1.6;/0}$ which is optimal at the stable distribution with $\alpha = 1.6$ and $\beta = 0$, and the Student test ϕ_{t}.
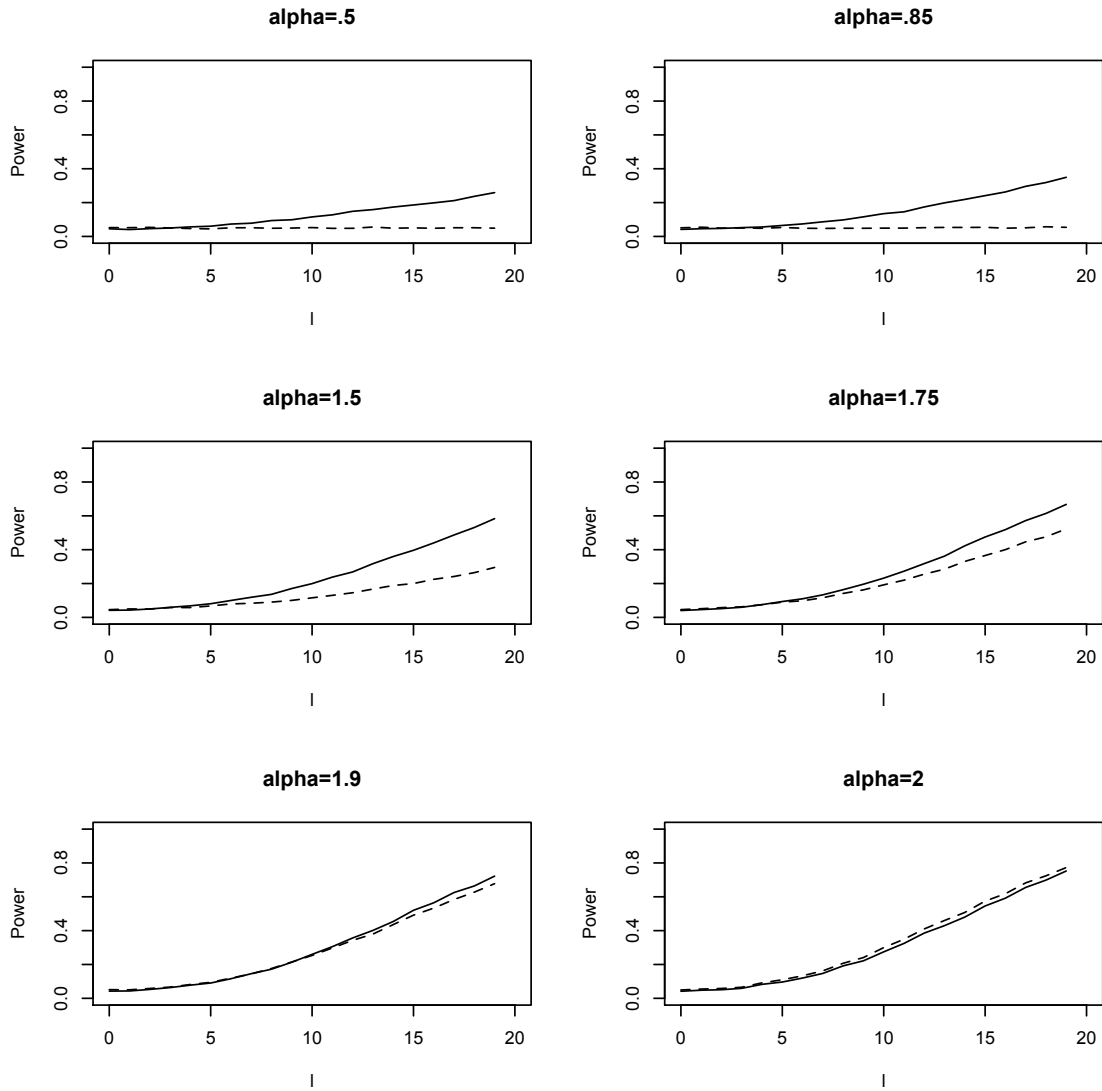
Figure 2: Power curves of the van der Waerden (solid line) and Student (dotted line) tests computed from 10,000 replications for various symmetric stable errors. Sample size is $n = 100$ and regression constants are drawn from the uniform distribution on $[-5, 5]$.
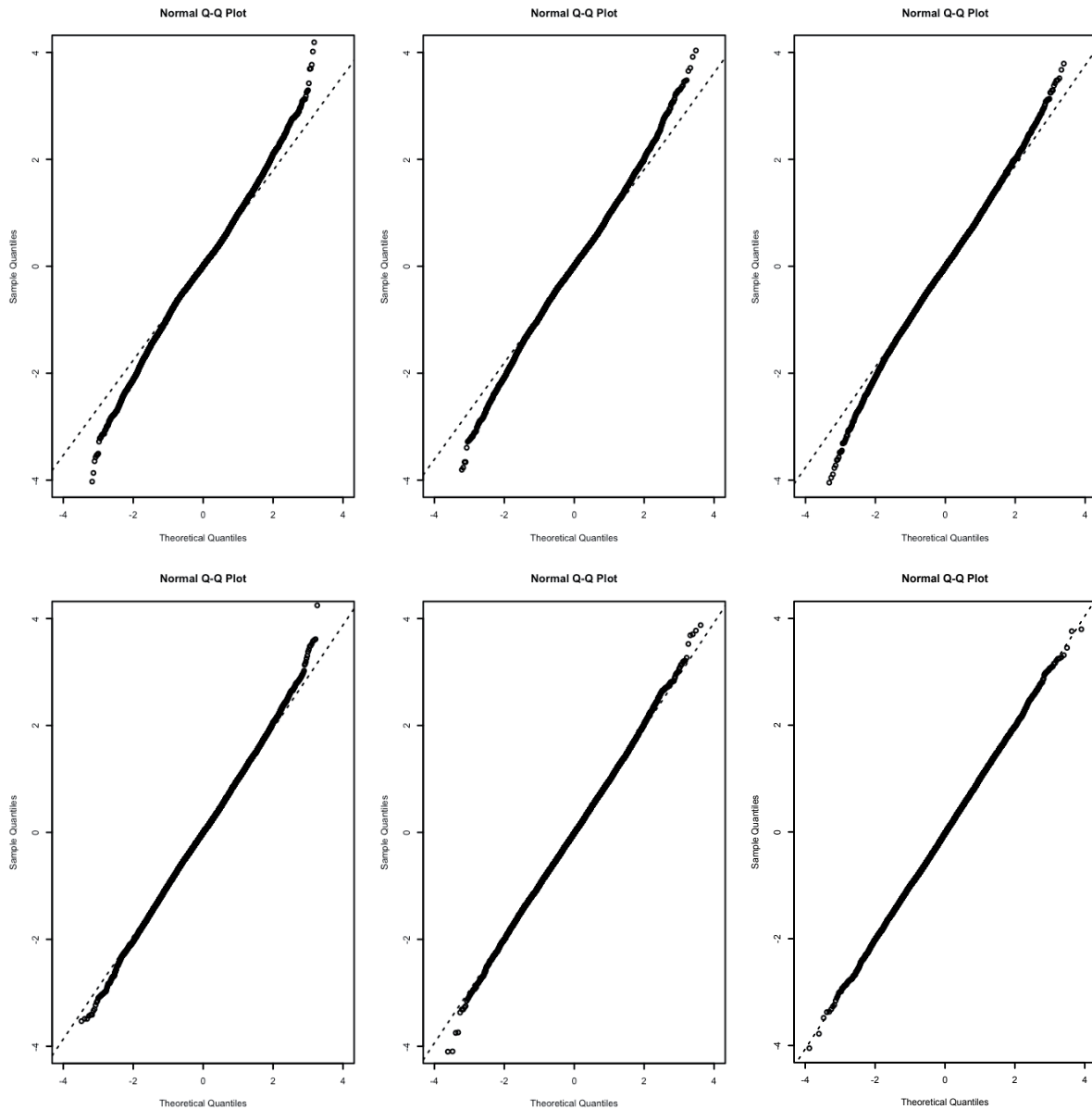
Figure 3: Normal QQ plots of the Student statistic obtained from $10,000$ replications under the null hypothesis of no slope and various symmetric stable distributions: $\alpha = .5, .8, 1.2, 1.5, 1.8,$ and 2, respectively from the top-left corner to the bottom-right corner. The sample size is $n = 1000$ and the regression constants are filtered (with a GARCH(1,1)) S&P500 daily returns from 2004 to 2008 .
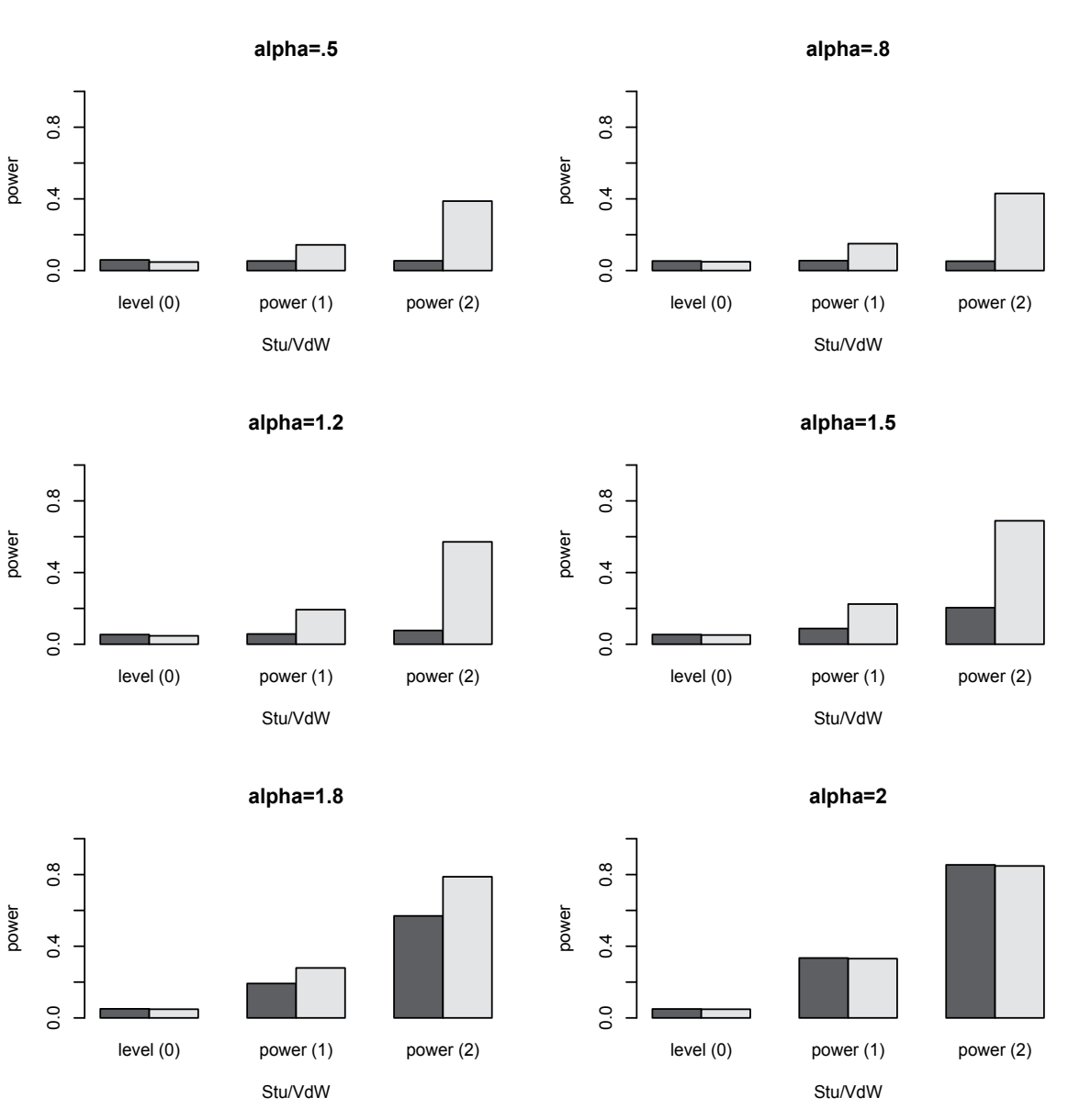
Figure 4: Rejection frequencies of the Student test (black) and the van der Waerden test (white) obtained from $10,000$ replications under the null hypothesis of no slope (0) and local alternatives ((1) and (2)) for various symmetric stable distributions: $\alpha = .5, .8, 1.2, 1.5, 1.8,$ and 2, respectively, from the top-left corner to the bottom-right corner. The sample size is $n = 1000$ and the regression constants are filtered (with a GARCH(1,1)) S&P500 daily returns from 2004 to 2008 .

5 Conclusions and discussion.

This paper provides a study of the asymptotic properties of rank-based tests for the general linear model with stable errors. Being distribution-free, rank-based tests are validity-robust. That is, for any values of the tail index α and the skewness parameter β, exact finite-sample critical values can be computed, whereas the test statistics are asymptotically chi-square under the null, uniformly in α and β.

Using the LAN property of the model, we were able to obtain the asymptotic distribution of our test statistics under local alternatives. We compare them using asymptotic relative efficiencies. Besides the usual rank-based statistics (van der Waerden, Wilcoxon and Laplace), we also show how to construct new ones that are optimal under some prespecified stable density (while remaining valid under any density). ARE calculations and Monte-Carlo experiments show that the traditional parametric tests (essentially the Student test, the validity of which remains an unsolved conjecture) are severely outperformed by the rank-based ones—particularly for small values of the tail index α.

Rank-based methods also can be considered for estimation purposes, hopefully with the same good performances as in testing problems; such estimation methods are the subject of ongoing research.

6 Appendix

6.1 Appendix A: Proofs

Proof of Lemma 2.1. Consider an admissible $\boldsymbol{\theta}$ with tail parameter $\alpha < 2$. From (2.3), we get

$$c_{\boldsymbol{\theta}} = \lim_{x \to \infty} \frac{1 - F_{\boldsymbol{\theta}}(x)}{x^{-\alpha}} = \lim_{x \to \infty} \frac{-f_{\boldsymbol{\theta}}(x)}{-\alpha x^{-\alpha-1}} = \lim_{x \to \infty} \frac{-\dot{f}_{\boldsymbol{\theta}}(x)}{\alpha(\alpha+1)x^{-\alpha-2}},$$

where the last two equalities hold via Hospital's rule. Hence, if $1 - F_{\boldsymbol{\theta}}(x) = O(x^{-\alpha})$ for large x, then $f_{\boldsymbol{\theta}}(x) = O(x^{-\alpha-1})$ and $\dot{f}_{\boldsymbol{\theta}}(x) = O(x^{-\alpha-2})$. The assertion follows. $\square$

Proof of Proposition 3.1. Theorem 12.2.1 in Lehmann and Romano (2005) entails that the sequence of models is ULAN (hence, a fortiori, LAN) if

(i) $\boldsymbol{\theta} \mapsto f_{\boldsymbol{\theta}}(x)$ is continuously differentiable, with gradient vector $\nabla_{\boldsymbol{\theta}} f_{\boldsymbol{\theta}}(x)$, for almost every x;

(ii) the Fisher information matrix $\Gamma(\boldsymbol{\theta}) := \mathrm{E}_{\boldsymbol{\theta}}\left[(\nabla_{\boldsymbol{\theta}} \log f_{\boldsymbol{\theta}}(X_i))(\nabla_{\boldsymbol{\theta}} \log f_{\boldsymbol{\theta}}(X_i))'\right]$ is finite, and the mapping $\boldsymbol{\theta} \mapsto \Gamma(\boldsymbol{\theta})$ is continuous.

In our context, (i) is trivially satisfied. As for (ii), one easily sees that it suffices to show that $\int_M^\infty (\dot{f}_{\alpha,\beta}(x)/f_{\alpha,\beta}(x))^2 f(x)dx < \infty$, for M large. This last assertion results directly from Lemma 2.1. This concludes the proof. $\square$

Proof of Proposition 3.2. (i) Since $\boldsymbol{\Delta}_{\boldsymbol{\theta}}^{(n)}(\boldsymbol{\beta})$ and $\boldsymbol{\Delta}_J^{(n)}(\boldsymbol{\beta})$ both are sums of i.i.d. terms, the Central Limit Theorem yields

$$\left(\begin{array}{c} \boldsymbol{\Delta}_{J,\boldsymbol{\theta}}^{(n)}(\boldsymbol{\beta}) \\ \boldsymbol{\tau}' \boldsymbol{\Delta}_{\boldsymbol{\theta}}^{(n)}(\boldsymbol{\beta}) \end{array} \right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(\left(\begin{array}{c} \mathbf{0} \\ 0 \end{array} \right), \left(\begin{array}{cc} \sigma_1^2 \mathbf{I}_K & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{12}' & \sigma_2^2 \end{array} \right) \right), \tag{6.1}$$

under $\mathcal{H}_{\boldsymbol{\theta}}^{(n)}(\boldsymbol{\beta})$, where

$$\sigma_1^2 := \mathcal{J}(J) = \int_0^1 J^2(u)du, \qquad \sigma_2^2 = \boldsymbol{\tau}'\boldsymbol{\tau}\mathcal{I}(\boldsymbol{\theta}),$$

and

$$\boldsymbol{\Sigma}_{12} := \left(\int_0^1 J(u)\phi_{\boldsymbol{\theta}}(F_{\boldsymbol{\theta}}^{-1}(u))du\right)\boldsymbol{\tau}. \tag{6.2}$$

Le Cam's Third Lemma together with the asymptotic equivalence (3.7) immediately yields the desired result. As for part (ii) of the proposition, it is a direct consequence of part (i). □

6.2 Appendix B: Numerical calculations with stable densities

For the sake of completeness, we provide here the various integral formulas for stable densities (in Zolotarev's parameterization) that have been used throughout. Assuming that $\alpha \neq 1$, let (recall the notation $\zeta = \zeta(\alpha, \beta) := -\beta\tan(\pi\alpha/2)$)

$$\theta_0 := \theta_0(\alpha, \beta) = -\frac{1}{\alpha}\arctan(\zeta), \qquad c_1(\alpha, \beta) := \begin{cases} \frac{1}{\pi}(\frac{\pi}{2} - \theta_0) & \text{if } \alpha < 1 \\ 1 & \text{if } \alpha > 1, \end{cases}$$

and

$$\begin{aligned} g(\theta; x, \alpha, \beta) &= (x - \zeta)^{\alpha/(\alpha-1)}(\cos\alpha\theta_0)^{1/(\alpha-1)} \\ &\times \left(\frac{\cos\theta}{\sin\alpha(\theta + \theta_0)}\right)^{\alpha/(\alpha-1)}\frac{\cos(\alpha\theta_0 + (\alpha - 1)\theta)}{\cos\theta}. \end{aligned}$$

With that notation, one obtains (see Nolan 1997)

(i) $f_{\alpha,\beta}(x) = \frac{\alpha}{\pi|\alpha - 1|}(x - \zeta)^{-1}\int_{-\theta_0}^{\pi/2} g(\theta; x, \alpha, \beta)e^{-g(\theta;x,\alpha,\beta)}d\theta$ for $x > \zeta$,

(ii) $f_{\alpha,\beta}(\zeta) = \Gamma(1 + 1/\alpha)\cos\theta_0/\pi(1 + \zeta^2)^{1/2\alpha}$ and, by symmetry,

(iii) $f_{\alpha,\beta}(x) = f_{\alpha,-\beta}(-x)$ for $x < \zeta$;

hence,

(iv) $F_{\alpha,\beta}(x) = c_1(\alpha, \beta) + \frac{\text{sign}(1 - \alpha)}{\pi}\int_{-\theta_0}^{\pi/2} e^{-g(\theta;x,\alpha,\beta)}d\theta$ for $x > \zeta$,

(v) $F_{\alpha,\beta}(x) = 1 - F_{\alpha,\beta}(-x)$ for $x < \zeta$, and

(vi) $F_{\alpha,\beta}(\zeta) = \frac{1}{\pi}(\frac{\pi}{2} - \theta_0)$.

Integral expressions for the derivatives of stable densities readily follow. Indeed, interchanging derivatives and integrals in (i) yields, for given $x > \zeta$,

$$\begin{aligned} \dot{f}_{\alpha,\beta}(x) &= \frac{1}{(\alpha - 1)(x - \zeta)} \\ &\times \left(f_{\alpha,\beta}(x) - \frac{\alpha^2}{\pi|\alpha - 1|(x - \zeta)}\int_{-\theta_0}^{\pi/2} g(\theta; x, \alpha, \beta)^2 e^{-g(\theta;x,\alpha,\beta)}d\theta\right). \end{aligned} \tag{6.3}$$

Similarly, for $x < \zeta$, we get $\dot{f}_{\alpha,\beta}(x) = -\dot{f}_{\alpha,-\beta}(-x)$, and, taking the limit as $x \to \zeta$,

$$\dot{f}_{\alpha,\beta}(\zeta) = \Gamma(1 + \frac{2}{\alpha})\sin(2\theta_0)/2\pi(1 + \zeta^2)^{1/\alpha}.$$

These expressions allow us to compute

$$\mathcal{J}(J, \boldsymbol{\theta}) = \int_0^1 J(u)\varphi_{\boldsymbol{\theta}}(F_{\boldsymbol{\theta}}^{-1}(u))du = \int_{-\infty}^{\infty} J(F_{\boldsymbol{\theta}}(v))\dot{f}_{\boldsymbol{\theta}}(v)dv$$

for any closed-form score function J.

On the one hand, the integrands associated with Wilcoxon and Cauchy scores are sufficiently well-behaved (at least, when α is "not too close to one", i.e. $|\alpha - 1| \geq 0.05$) for straightforward numerical evaluation. The corresponding computation for Laplace scores J_{L} is even easier, since one can show that, in that case,

$$\mathcal{J}(J_{\mathrm{L}}, \boldsymbol{\theta}) = 2\sqrt{2}f_{\alpha,\beta}(F_{\alpha,\beta}^{-1}(0)).$$

On the other hand, computations for van der Waerden scores are more delicate due to the oscillating nature of the integrands. A computational remedy to this numerical instability consists in restricting the domain of integration to a finite interval in order for numerical integration algorithms to converge. Our numerical experience indicates that, setting

$$\tilde{\dot{f}}_{\boldsymbol{\theta}}(z) = \dot{f}_{\boldsymbol{\theta}}(z)1_{\{\dot{f}_{\boldsymbol{\theta}}(z) \geq 10^{-8}\}}$$

suffices for stabilizing the integral, whereas the loss incurred by this modification into $\tilde{\dot{f}}_{\boldsymbol{\theta}}$ of $\dot{f}_{\boldsymbol{\theta}}$ appears to be negligible.

The cross-correlation quantities for AREs for general stable scores are more complicated to derive for stable scores with general skewness and tail parameter. Indeed, computing

$$\mathcal{J}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = \int_0^1 \varphi_{\boldsymbol{\theta}_1}(F_{\boldsymbol{\theta}_1}^{-1}(u))\varphi_{\boldsymbol{\theta}_2}(F_{\boldsymbol{\theta}_2}^{-1}(u))du$$

requires integrating the quantile function of a stable distribution. Despite the rather surprisingly robust character of the integrals in (iv), numerical inversion is the only way to obtain the corresponding quantile functions; our numerical evaluations of the corresponding AREs therefore rely on the integral approximations

$$\mathcal{J}^{(n)}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = \frac{1}{n+1}\sum_{j=1}^{n}\varphi_{\boldsymbol{\theta}_1}\left(F_{\boldsymbol{\theta}_1}^{-1}\left(\frac{j}{n+1}\right)\right)\varphi_{\boldsymbol{\theta}_2}\left(F_{\boldsymbol{\theta}_2}^{-1}\left(\frac{j}{n+1}\right)\right). \tag{6.4}$$

Taking $n \geq 10^4$ in (6.4) yields an absolute error which is less than 10^{-3}.

The programs performing those evaluations have been written in Mathematica 5, and are available at the url: http://homepages.ulb.ac.be/~yvswan/. These programs make use of Mathematica's powerful numerical evaluation abilities to compute the different integrals reliably, although some of the computations are rather time-consuming. For more advanced results on computational issues with stable densities, tails and quantiles, we refer to Rimmer and Nolan (2005).

References

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

[2] Benktus, V., Bing-Yi J., Shao, S.M., and Zhou, W. (2007). Limiting distributions of the non-central t-statistic and their applications to the power of t-tests under non-normality. *Bernoulli* **13** (2), 346-364.

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

[4] Borak, Sz., Härdle, W., and Weron, R. (2005). Stable Distributions, in: Cizek, P., Härdle, W. and Weron, R. (Eds.), *Statistical Tools for Finance and Insurance*, Springer-Verlag, Berlin, 21-44.

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

[6] Efron, B. (1969). Student's t-test under symmetry conditions. *Journal of the American Statistical Association* **64**, 1278-1302.

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

[8] Fama, E. (1965a). Behavior of Stock Market Prices. *Journal of Business* **38**, 34-105.

[9] Fama, E. (1965b). Portfolio Analysis in a Stable Paretian Market. *Management Science* **11**, 404-419.

[10] Fama, E. and Roll, R. (1968). Some Properties of Symmetric Stable Distributions. *Journal of the American Statistical Association* **63** (323), 817836.

[11] Fama, E. and Roll, R. (1971). Parameter Estimates for Symmetric Stable Distributions. *Journal of the American Statistical Association* **66** (334), 331338

[12] Feller W. (1971). *An Introduction to Probability and its Applications*, Vol. 2, 2nd ed., Wiley Publications in Statistics.

[13] Garcia, R., Renault, E. and Veredas, D. (2010). Estimation of stable distributions with indirect inference. To appear in *Journal of Econometrics*.

[14] Giné, E., Götze, F. and Mason, D. (1997). When is the Student t-statistic asymptotically standard normal? *Annals of Probability* **25**, 1514-1531.

[15] Hájek, J. and Šidák, S. (1967). *Theory of Rank Tests*. Academic Press, New York.

[16] Hallin, M. and Werker, B.J.M. (2003). Semiparametric efficiency, distribution-freeness, and invariance. *Bernoulli* **9**, 137-165.

[17] Jensen, D. R. (1979). Linear Models without Moments. *Biometrika* **66** (3), 611-617

[18] Le Cam, L. (1986) *Asymptotic Methods in Statistical Decision Theory*, Springer-Verlag, New York.

[19] Le Cam, L., and Yang, G. L. (2000). *Asymptotics in Statistics*, 2nd edition. Springer-Verlag, New York.

[20] Lehmann, E. L. and Romano, J. P. (2005). *Testing Statistical Hypotheses*, Springer-Verlag, New York.

[21] Logan, B. F., Mallows, C. L., Rice S. O. and Shepp L. (1973). Limit distributions of self-normalized sums. *Annals of Probability* **1**, 788-809.

[22] Mandelbrot, B. (1963), The Variation of Speculative Prices, *Journal of Business* **36**, 394-419.

[23] McCulloch, J. H. (1986). Simple consistent estimators of stable distribution parameters. *Communications in Statistics-Simulations* **15**, 1109-1136.

[24] Mittnik, S., Paolella, M.S. and Rachev, S.T. (2000). Diagnosis and treating the fat tails in financial returns data. *Journal of Empirical Finance* **7**, 389-416.

[25] Mittnik, S., Rachev, S.T. and Paolella, M.S. (1998). Stable Paretian Modeling in Finance: Some Empirical and Theoretical Aspects, in: Adler, R., Feldman, R. and Taqqu, M.S. (Eds.), *A Practical Guide to Heavy Tails*, Birkhäuser, Boston, 79-110.

[26] Nolan, J. P. (1997). Numerical calculation of stable densities and distribution functions. *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] Nolan, J.P. (2010). *Stable Distributions - Models for Heavy Tailed Data*, Birkhäuser, Boston.

[29] Puri, M. L. and Sen, P. K. (1985). *Nonparametric Methods in General Linear Models*, Wiley, New York.

[30] Rimmer, R. H. and Nolan, J. P. (2005). Stable Distributions in Mathematica. *Mathematica Journal* **9**, 776-789.

[31] Samorodnitsky, G., and Taqqu, M. S. (1994). *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance* . Chapman and Hall, New York.

[32] van der Vaart, A. (2000). *Asymptotic Statistics*, Cambridge University Press, Cambridge.

[33] Zolotarev, V. M. (1986) *One-dimensional stable distributions*, Amer. Math. Soc. Transl. of Math. Monographs, Vol. 65. Amer. Math. Soc., Providence, RI.

[34] Zolotarev, V. M. (1995), On representation of densities of stable laws by special functions, *Theory of Probability and its Applications* **39**, 354-362.