Centre interuniversitaire de recherche
en économie quantitative

CIREQ

Cahier 13-2007

PROPERTIES AND ESTIMATION OF ASYMMETRIC EXPONENTIAL POWER DISTRIBUTION

Dongming ZHU and
Victoria ZINDE-WALSH

CIREQ

Le **Centre interuniversitaire de recherche en économie quantitative (CIREQ)** regroupe des chercheurs dans les domaines de l'économétrie, la théorie de la décision, la macroéconomie et les marchés financiers, la microéconomie appliquée et l'économie expérimentale ainsi que l'économie de l'environnement et des ressources naturelles. Ils proviennent principalement des universités de Montréal, McGill et Concordia. Le CIREQ offre un milieu dynamique de recherche en économie quantitative grâce au grand nombre d'activités qu'il organise (séminaires, ateliers, colloques) et de collaborateurs qu'il reçoit chaque année.

*The **Center for Interuniversity Research in Quantitative Economics (CIREQ)** regroups researchers in the fields of econometrics, decision theory, macroeconomics and financial markets, applied microeconomics and experimental economics, and environmental and natural resources economics. They come mainly from the Université de Montréal, McGill University and Concordia University. CIREQ offers a dynamic environment of research in quantitative economics thanks to the large number of activities that it organizes (seminars, workshops, conferences) and to the visitors it receives every year.*

Cahier 13-2007

PROPERTIES AND ESTIMATION OF ASYMMETRIC EXPONENTIAL POWER DISTRIBUTION

Dongming ZHU and Victoria ZINDE-WALSH

CIREQ, Université de Montréal
C.P. 6128, succursale Centre-ville
Montréal (Québec) H3C 3J7
Canada

téléphone : (514) 343-6557
télécopieur : (514) 343-5831
cireq@umontreal.ca
http://www.cireq.umontreal.ca

Dépôt légal - Bibliothèque nationale du Canada, 2007, ISSN 0821-4441
Dépôt légal - Bibliothèque et Archives nationales du Québec, 2007
ISBN-13 : 978-2-89382-554-0

Properties and Estimation of Asymmetric Exponential Power Distribution

Dongming Zhu
Shenzhen Graduate School of Business, Peking University

Victoria Zinde-Walsh*
Department of Economics, McGill University and CIREQ

October 23, 2007

Abstract

The new distribution class, Asymmetric Exponential Power Distribution (AEPD), proposed in this paper generalizes the class of Skewed Exponential Power Distributions (SEPD) in a way that in addition to skewness introduces different decay rates of density in the left and right tails. Our parametrization provides an interpretable role for each parameter. We derive moments and moment-based measures: skewness, kurtosis, expected shortfall. It is demonstrated that a maximum entropy property holds for the AEPD distributions. We establish consistency, asymptotic normality and efficiency of the maximum likelihood estimators over a large part of the parameter space by dealing with the problems created by non-smooth likelihood function and derive explicit analytical expressions of the asymptotic covariance matrix; where the results apply to the SEPD class they enlarge on the current literature. Finally, we give a convenient stochastic representation of the distribution; our Monte Carlo study illustrates the theoretical results.

JEL classification C13, C16

Key words: asymmetric distributions, maximum likelihood estimation.

1 Introduction

Observed characteristics of many financial data series have motivated exploration of classes of distributions that can accommodate properties such as fat-tailedness and skewness while nesting distributions typically used in estimation such as the normal (and skew-normal). An important desired property of

*Corresponding author. Address: Department of Economics, 855 Sherbrooke Street West, Montreal, PQ, H3A 2T7, Canada. Tel. (514) 398 4834; email victoria.zinde-walsh@mcgill.ca
The support of the Social Sciences and Humanities Research Council of Canada (SSHRC) and the *Fonds québecois de la recherche sur la société et la culture* (FRQSC) is gratefully acknowledged.

any such class is that it permits maximum likelihood estimation of all parameters. Obtaining closed-form expressions for the moments of interest, such as the mean, variance, skewness and kurtosis, as well as components of the information matrix provides useful interpretable features of the distributions in the class. For applications in risk management one may in addition be interested in closed-form expressions for value-at-risk and expected shortfall of asset/portfolio returns. Classes of non-symmetric distributions that nest the skew-normal were constructed by Azzalini (1986). Other classes of distributions with the desired properties of accommodating heavy tails and skewness, the Skewed Exponential Power Distribution (SEPD) classes, were proposed in Fernandez et al (1995), Theodossiou (2000) and Komunjer (2007); they all generalize the generalized error distribution (GED) class[1]. Many financial applications of the GED as well as its skew extensions have been considered in Hsieh (1989), Nelson (1991), Duan (1999), Rachev and Mittinik (2000), Theodossiou (2000), Ayebo and Kozubowski (2004), Komunjer (2007), Christoffersen et al (2005) and others. Especially in applications to option pricing, the GED and its skew extensions are preferred to Student-t distributions because it is found that Student-t distributions are not suited to model continuously compounded returns (see Duan (1999) and Theodossiou (2000)). Since all moments of the GED exist, the moments of exponential transformations of GED random variables, needed to price options, can be evaluated.

Ayebo and Kozubowski (2004) presented basic properties of the SEPD of Fernandez et al (1995), derived maximum likelihood estimators of scale and skewness parameters given other parameters, and discussed its applications in finance. Komunjer (2007) explored moments (also see Theodossiou (2000)) as well as measures such as value at risk and expected shortfall useful in financial applications. DiCiccio and Monti (2004) studied properties of MLEs of the Azzalini's (1986) SEPD, and obtained results for the information matrix (not in closed form) and for inferential properties of MLE.

However, for some applications in finance and risk management, the skew extensions may not be rich enough to capture all the asymmetry of distributions of asset returns, particularly asymmetry in the tails. For example, it is found especially for portfolios such as *S&P500* and *NASDAQ* that *ex post* innovations from estimated *GARCH* models (even with a leverage effect) are not normally distributed—the QQ plot of *ex post* innovations typically shows that the fit in the upper tail is good but the lower tail is heavier than that of the normal distribution (see Figure 6 of Bradley and Taqqu (2003) for *NASDAQ*, Figure 4.2 of Christoffersen (2003) for *S&P500*). To capture the asymmetry in the tails, this paper extends the SEPD to a fully asymmetric exponential power distribution (AEPD) where heavy-tailedness itself may be asymmetric with different tail exponents on different sides of the distribution.

We demonstrate that the AEPD class has desired properties: interpretable parameters to represent location, scale, and shape, closed-form expressions for

[1] The GED class was proposed first by Subbotin (1923); Box and Tiao (1973) called such a distribution the Exponential Power Distribution (EPD). It is also called the Generalized Power Distribution or the Generalized Laplace Distribution.

the moments as well as for value at risk and expected shortfall. A maximum entropy property is shown to hold and a stochastic representation of the AEPD is given. We develop asymptotic properties of the MLE (consistency and asymptotic normality) and obtain fully closed-form expressions for the information matrix for all parameters. Thus we also provide new theoretical results such as closed-form expressions for the asymptotic covariance matrix and consistency and asymptotic normality of MLE for some SEPD classes, expanding on results currently available in the literature. Comparing the AEPD with Azzalini's (1986) SEPD class, both classes have continuous but non-differentiable densities; the latter density however involves an integral (normal cdf). Also, the AEPD has more flexible tail behavior and analytical expressions for mode and moments; while for Azzalini's (1986) SEPD, the left tail is always thinner than the right one, its odd moments involve infinite series expansions, and it is not possible to find an analytic expression for the mode. In addition, note that DiCiccio and Monti (2004) were not able to provide closed form expressions for the information matrix (nor complete proofs of asymptotics for MLE) for Azzalini's (1986) SEPD.

The paper is organized as follows. Section 2 explains the relation between EPD, SEPD and AEPD classes highlighting the main features of the new AEPD class. The interpretation of parameters is provided in Section 3. Section 4 gives basic properties of the AEPD such as analytical expressions of cdf, quantiles, moments and expected shortfall. In Section 5 we establish consistency and asymptotic normality of the MLE and Section 6 provides some finite sample Monte Carlo results. Technical results and proofs are collected in the appendices A, B and C. Figures and tables are in Appendix D.

2 The relation between EPD, SEPD and AEPD

The density function of the EPD (or GED) is usually defined as:

$$f_{EP}(x \mid p, \mu, \sigma) = \frac{1}{\sigma} K_{EP}(p) \exp\left(-\frac{1}{p} \left| \frac{x-\mu}{\sigma} \right|^p \right), \tag{1}$$

where $\mu \in R$ and $\sigma > 0$ are the location and scale parameters respectively, $p > 0$ is the shape parameter, and $K_{EP}(p)$ is the normalizing constant, $K_{EP}(p) \equiv 1/[2p^{1/p}\Gamma(1+1/p)]$. If X is a random variable with the EPD density, then the location parameter $\mu = E(X) = med(X)$, the median of X; the scale parameter $\sigma = (E\,|X-\mu|^p)^{1/p}$, which is the L_p-norm deviation, has an interpretation similar to that of the standard deviation of the normal distribution. When the shape parameter p gets smaller and smaller, the EPD becomes more and more heavy-tailed and leptokurtic. With $p = 2$, $p = 1$, and $p \to +\infty$, the EPD reduces to the normal, Laplace and uniform distributions, respectively.

So far, there are two different methods to extend the EPD to a skewed exponential power distribution (SEPD). Azzalini (1986) first proposed a family of SEPD based on the fact that if $f(\cdot)$ is a density symmetric about 0 and $\Pi(\cdot)$ an absolutely continuous distribution function such that its pdf $\Pi'(\cdot)$ is symmetric about 0, then $2\Pi(\lambda x)f(x)$ is a density for any real λ. Taking $f = f_{EP}$ and Π = normal cdf or EPD's cdf, we get Azzalini's SEPD class. Fernandez et al (1995) extended the EPD class to *another* family of SEPD by using a two-piece method, in which an additional skew parameter γ is introduced (also see Kotz et al (2001), p271).

By a method similar to that of Fernandez et al (1995), Theodossiou (2000) and Komunjer (2007), respectively, constructed seemingly different classes of SEPD, which are actually reparametrizations of that of Fernandez et al (1995). However, Komunjer's (2007) asymmetry parameter α is interestingly interpreted as the probability such that the location parameter is exactly the $\alpha-$quantile of the SEPD. Noting the interpretable nature of the parameters this paper follows a similar method to construct the AEPD.

The AEPD density has the following form:

$$f_{AEP}(x \mid \beta) = \begin{cases} (\frac{\alpha}{\alpha^*})\frac{1}{\sigma}K_{EP}(p_1)\exp\left(-\frac{1}{p_1}\left|\frac{x-\mu}{2\alpha^*\sigma}\right|^{p_1}\right), & \text{if } x \le \mu; \\ \\ (\frac{1-\alpha}{1-\alpha^*})\frac{1}{\sigma}K_{EP}(p_2)\exp\left(-\frac{1}{p_2}\left|\frac{x-\mu}{2(1-\alpha^*)\sigma}\right|^{p_2}\right), & \text{if } x > \mu, \end{cases} \tag{2}$$

where $\beta = (\alpha, p_1, p_2, \mu, \sigma)^T$ is the parameter vector, $\mu \in R$ and $\sigma > 0$ still represent location and scale, respectively, $\alpha \in (0,1)$ is the skewness parameter, $p_1 > 0$ and $p_2 > 0$ are the left and right tail parameters respectively, $K_{EP}(p)$ is the same as in (1), and α^* is defined as

$$\alpha^* = \alpha K_{EP}(p_1)/[\alpha K_{EP}(p_1) + (1-\alpha)K_{EP}(p_2)]. \tag{3}$$

Note that

$$\frac{\alpha}{\alpha^*}K_{EP}(p_1) = \frac{1-\alpha}{1-\alpha^*}K_{EP}(p_2) = \alpha K_{EP}(p_1) + (1-\alpha)K_{EP}(p_2) \equiv B. \tag{4}$$

The AEPD density function is still continuous at every point and unimodal with mode at μ. The parameter α^* in the AEPD density provides scale adjustments respectively to the left and right parts of the density so as to ensure continuity of the density under changes of shape parameters (α, p_1, p_2). If $p_1 = p_2 = p$, implying $\alpha^* = \alpha$, the AEPD reduce to a new version of SEPD:

$$f_{SEP}(x \mid \beta) = \begin{cases} \frac{1}{\sigma}K_{EP}(p)\exp\left(-\frac{1}{p}\left|\frac{x-\mu}{2\alpha\sigma}\right|^{p}\right), & \text{if } x \le \mu; \\ \\ \frac{1}{\sigma}K_{EP}(p)\exp\left(-\frac{1}{p}\left|\frac{x-\mu}{2(1-\alpha)\sigma}\right|^{p}\right), & \text{if } x > \mu, \end{cases} \tag{5}$$

which is equivalent to those of Fernandez et al (1995), Theodossiou (2000) and Komunjer (2007). This new version of SEPD provides new interesting interpretations for scale and skewness in terms of L_p distances. The skewness parameter $\alpha \in (0,1)$ plays the same role as the parameter γ of Fernandez et al (1995). By reparametrization, $\alpha = \gamma^2/(1+\gamma^2)$ and $\sigma = (2/p)^{1/p}(\gamma + 1/\gamma)\sigma'/2$, the SEPD (5) will become that of Fernandez et al (1995); a re-scaling of the density leads to Komunjer's (2007); letting $\alpha = (1+\lambda)/2$, $\sigma = \theta\sigma' p^{-1/p}$ and $\mu = \mu' - \delta\sigma'$, the density will be that (i.e., $f(y \mid \mu', \sigma', p, \lambda)$ in Eq (10)) of Theodossiou (2000), where θ and δ are given in Equations (12) and (13) of Theodossiou (2000). With $\alpha = 1/2$, the SEPD (5) reduces to the EPD (1). The skewed Laplace distribution and skewed normal distribution are special cases of the SEPD, respectively, with $p = 1$ and $p = 2$. The SEPD density is skewed to the right for $\alpha < 1/2$ and to the left for $\alpha > 1/2$.

A convenient reparametrization of (2) is obtained by rescaling,

$$f_{AEP}(x \mid \theta) = \begin{cases} \frac{1}{\sigma} \exp\left(-\frac{1}{p_1}\left|\frac{x-\mu}{2\alpha\sigma K_{EP}(p_1)}\right|^{p_1}\right), & \text{if } x \le \mu; \\ \frac{1}{\sigma} \exp\left(-\frac{1}{p_2}\left|\frac{x-\mu}{2(1-\alpha)\sigma K_{EP}(p_2)}\right|^{p_2}\right), & \text{if } x > \mu, \end{cases} \tag{6}$$

where $\theta = (\alpha, p_1, p_2, \mu, \sigma)^T$. From the rescaled AEPD density (6), we can clearly observe the effects of the shape parameters on the distribution. The density in the form (6) is used in deriving a closed form expression for the information matrix of the maximum likelihood estimator (MLE).

3 Interpretation of parameters of AEPD

The main tools that are used for interpretation are various L_r space related distance measures. Define for $r > 0$,

$$d_L(r) \equiv [E(|X-\mu|^r \mid X \le \mu)]^{1/r},$$

$$d_R(r) \equiv [E(|X-\mu|^r \mid X > \mu)]^{1/r},$$

respectively called the $L_r - norm$ deviation (or distance) conditional on $X \le \mu$ and the $L_r - norm$ deviation conditional on $X > \mu$. The total conditional deviation (or distance) is

$$d(r) = d_L(r) + d_R(r);$$

the L_r-norm deviation $\|X-\mu\|_r = (E|X-\mu|^r)^{1/r}$.

Suppose now that random variable X has the AEPD density defined in (2) with shape parameters (α, p_1, p_2), location μ and scale σ.

Proposition 1 *The following relations hold:*

(a) $P(X \le \mu) = \alpha$; also $d_L(p_1) = 2\alpha^*\sigma$, and $d_R(p_2) = 2(1-\alpha^*)\sigma$, where α^* is defined in (3);

(b) $\sigma = \frac{1}{2}[d_L(p_1) + d_R(p_2)]$;

(c) there is a positive function $r^*(c \mid p)$ depending on parameter p and increasing in its argument, c, such that

$$\alpha = \frac{d_L(r^*(c \mid p_1))}{d_L(r^*(c \mid p_1)) + d_R(r^*(c \mid p_2))};\ \forall\ c > \max\{lb(p_1), lb(p_2)\}$$

where $lb(p) \equiv [2\Gamma(1+1/p)]^{-1} \exp\{\frac{1}{p}\Psi(1/p)\}$ *and* $\Psi(x) \equiv \Gamma'(x)/\Gamma(x)$ *is digamma function.*

(d) $d_L(r) = 2\alpha^*\sigma M(p_1, r) = 2\alpha\sigma\xi(p_1, r)/B$ *and* $d_R(r) = 2(1-\alpha^*)\sigma M(p_2, r)$ $= 2(1-\alpha)\sigma\xi(p_2, r)/B$, *where* $M(p, r) \equiv p^{1/2}\{\Gamma((r+1)/p)/\Gamma(1/p)\}$; $\xi(p, r) \equiv K_{EP}(p)M(p, r)$ *is strictly increasing in* r *and decreasing in* p, B *and* $K_{EP}(p)$ *are defined above.*

Proof. See Appendix A.

From part (a) the location μ is the $\alpha-$quantile of the r.v. X and the scale σ is related to the left and right L_p conditional deviations by the parameter α^* (for the SEPD $\alpha^* = \alpha$). Part (b) represents the scale σ via an average of the left and right conditional deviations. It follows from (a) that the ratio of the left conditional deviation to the total is α^* (for SEPD just α). Part (c) gives an interpretation of α with two adjusted order functions $r^*(c \mid p_i)$, $i = 1, 2$; in the SEPD case ($p_1 = p_2$) the left and right conditional deviations enter with a same order thus then $\alpha = d_L(r)/d(r)$ for any $r > 0$.

Part (d) allows us to investigate the effect of shape parameters a, p_1, p_2. These shape parameters have a common effect on both $d_L(r)$ and $d_R(r)$ through $B = \alpha K(p_1)+(1-\alpha)K(p_2)$, which represents a scale adjustment effect. Ignoring the common effect, α has the same effect on the AEPD as it does on the SEPD, but the left and right tail parameters, $p_1$and p_2, respectively control the left and right L_r-deviation, $d_L(r)$ and $d_R(r)$. Since $\xi(p, r)$ is a strictly decreasing function of p for any given r, a smaller p_1 (or p_2) leads to a larger left (or right) L_r-deviation, thus AEPD with a smaller p_1 (or p_2) has a heavier left (or right) tail.

The effect of p_1 (or p_2) on the left (or right) tail can be measured by a generalized kurtosis index $kur_L(r)$ (or $kur_R(r)$) for $r > 0$, called the left (or right) generalized kurtosis (similar to Mineo (1989) who defined generalized kurtosis as $\frac{E|X-\mu|^{2p}}{(E|X-\mu|^p)^2}$ and showed that for EPD it is $p+1$). The left and right (generalized) kurtoses are defined as

$$kur_L(r) \equiv [d_L(2r)/d_L(r)]^{2r},$$
$$kur_R(r) \equiv [d_R(2r)/d_R(r)]^{2r}.$$

With $r = 2$ we get the usual definition of kurtosis.

Proposition 2 *For the AEPD the left and right (generalized) kurtosis can be expressed as follows:*

$$kur_L(r) = \Gamma(\frac{1}{p_1})\Gamma(\frac{2r+1}{p_1})/\Gamma^2(\frac{r+1}{p_1}), \quad (7)$$

$$kur_R(r) = \Gamma(\frac{1}{p_2})\Gamma(\frac{2r+1}{p_2})/\Gamma^2(\frac{r+1}{p_2}); \quad (8)$$

they are strictly decreasing respectively in p_1 *and* p_2 *for any given* $r > 0$, *and strictly increasing in* r *given* $p_1,\ p_2 > 0$.

Proof. See Appendix A.

From the expressions for $kur_L(r)$ and $kur_R(r)$ of the AEPD, the heaviness of the left (or right) tail is controlled by only p_1(or p_2). If $p_1 < p_2$, then $kur_L(r) > kur_R(r)$, implying that the left tail is heavier than the right. When $p_i < 2$ $(i = 1, 2)$, the AEPD is more heavy-tailed than the normal distribution. The left (or right) tail parameter p_1 (or p_2) is directly related to the left (or right) generalized kurtosis by the relation: $kur_L(p_1) = p_1 + 1$ (or $kur_R(p_2) = p_2 + 1$). Further results about kurtosis via moments are in the next section.

The plots 1-3 in Appendix D (Figure 1) plot the AEPD densities of the form (6) with $\mu = 0$ and $\sigma = 1$ for combinations of shape parameters (α, p_1, p_2). Plot 1 shows that for given p_1 and p_2 the density curve shifts to the right with α decreasing but its mode does not change; Plot 2 shows how p_2 controls only the right tail — heavier and heavier for smaller and smaller p_2. The effect of skewness parameter and tail parameters on tails is compared in Plot 3. Although a smaller α leads to a fatter right tail, this influence eventually is dominated by the effect of a smaller p_2.

4 Basic Properties of the AEPD

4.1 Cumulative distribution, quantile function and moments

All the formulae in this section follow straightforwardly from results for the classical EPD (summarized in (III) in appendix A).

Suppose that X is a random variable with the standard AEPD density ($\mu = 0$, $\sigma = 1$). Denote $a \wedge b \equiv \min\{a, b\}$, $a \vee b \equiv \max\{a, b\}$, by $G(x; \gamma)$ the gamma cdf:

$$G(x; \gamma) \equiv (\Gamma(\gamma))^{-1} \int_0^x z^{\gamma-1} \exp(-z) dz, \tag{9}$$

and by $G^{-1}(x; \gamma)$ the inverse function of $G(x; \gamma)$. Then for the standard AEPD, the cdf can be expressed via $G(\cdot; \cdot)$:

$$F_{AEP}(x \mid \alpha, p_1, p_2) = \begin{cases} \alpha \left[1 - G\left(\frac{1}{p_1}(\frac{|x|}{2\alpha^*})^{p_1}; \frac{1}{p_1}\right)\right], & \text{if } x \leq 0 \\ \alpha + (1-\alpha) G\left(\frac{1}{p_2}(\frac{|x|}{2(1-\alpha^*)})^{p_2}; \frac{1}{p_2}\right), & \text{if } x > 0 \end{cases} \tag{10}$$

and the quantile function is expressed via $G^{-1}(\cdot; \cdot)$

$$F_{AEP}^{-1}(\upsilon \mid \alpha, p_1, p_2) = \begin{cases} -2\alpha^* \left[p_1 G^{-1}\left(1 - \frac{\upsilon}{\alpha}; \frac{1}{p_1}\right)\right]^{1/p_1}, & \text{if } \upsilon \leq \alpha \\ 2(1-\alpha^*) \left[p_2 G^{-1}\left(1 - \frac{1-\upsilon}{1-\alpha}; \frac{1}{p_2}\right)\right]^{1/p_2}, & \text{if } \upsilon > \alpha \end{cases} \tag{11}$$

where $\upsilon \in [0, 1]$.

Note that, for any measurable function $h(X)$ of the standard AEPD random variable X, we have

$$E[h(X)] = \alpha E[h(X) \mid X \leq 0] + (1-\alpha)E[h(X) \mid X > 0],$$

implying that all unconditional moments can be expressed as a weighted sum of two conditional moments. Therefore, we first give the conditional moments of the standard AEPD r.v. X. From expression for the absolute moment of EPD (32), we get for any real $r > -1$,

$$E(|X|^r \mid X < 0) = [2\alpha^*]^r E\,|Z_{p_1}|^r = B^{-r}\alpha^r H_r(p_1), \tag{12}$$

$$E(|X|^r \mid X > 0) = [2(1-\alpha^*)]^r E\,|Z_{p_2}|^r = B^{-r}(1-\alpha)^r H_r(p_2). \tag{13}$$

where Z_p is a random variable that has the standard EPD density ($\mu = 0, \sigma_p = 1$ in (1)) with power index p, B is defined in (4), $H_r(p) \equiv p^r\Gamma(\frac{1+r}{p})/\Gamma^{1+r}(\frac{1}{p})$. For any non-negative integer k, the kth right-conditional moment, $E(X^k \mid X > 0)$, has the same expression as in (13), while the kth left-conditional moment, $E(X^k \mid X < 0)$, has an expression slightly different from (12):

$$E(X^k \mid X < 0) = [-2\alpha^*]^k E\,|Z_{p_1}|^k = B^{-k}(-\alpha)^k H_k(p_1).$$

Thus, the kth moment of the standard AEPD r.v. X equals

$$E(X^k) = B^{-k}[(-1)^k\alpha^{1+k}H_k(p_1) + (1-\alpha)^{1+k}H_k(p_2)],\ k = 1, 2, 3, ..., \tag{14}$$

and its r-absolute moment is expressed as

$$E(|X|^r) = B^{-r}[\alpha^{1+r}H_r(p_1) + (1-\alpha)^{1+r}H_r(p_2)],\ r > -1. \tag{15}$$

In particular, the mean and variance of the standard AEPD r.v. X are given as follows:

$$E(X) = \frac{1}{B}[(1-\alpha)^2\frac{p_2\Gamma(2/p_2)}{\Gamma^2(1/p_2)} - \alpha^2\frac{p_1\Gamma(2/p_1)}{\Gamma^2(1/p_1)}], \tag{16}$$

$$\begin{aligned} Var(X) \quad = \quad & \frac{1}{B^2}\{(1-\alpha)^3\frac{p_2^2\Gamma(3/p_2)}{\Gamma^3(1/p_2)} + \alpha^3\frac{p_1^2\Gamma(3/p_1)}{\Gamma^3(1/p_1)} \\ & -[(1-\alpha)^2\frac{p_2\Gamma(2/p_2)}{\Gamma^2(1/p_2)} - \alpha^2\frac{p_1\Gamma(2/p_1)}{\Gamma^2(1/p_1)}]^2\}. \end{aligned} \tag{17}$$

We see that all moments can be expressed simply and conveniently in terms of gamma function. In the case of the SEPD $p_1 = p_2 = p$ and we get simplified expressions for moments:

$$E(X^k) = (2p^{1/p})^k[(-1)^k\alpha^{1+k} + (1-\alpha)^{1+k}]\Gamma((1+k)/p)/\Gamma(1/p), \tag{18}$$

$$E(|X|^r) = (2p^{1/p})^r[\alpha^{1+r} + (1-\alpha)^{1+r}]\Gamma((1+r)/p)/\Gamma(1/p), \tag{19}$$

where $k = 1, 2, ...,$and $r > -1$. These provide an advantage over Azzalini's (1986) SEPD class where the expressions for the odd moments involve infinite series expansions; (18) is a reparametrization of formulae of Fernandez et al (1995) and Komunjer (2007).

4.2 Value at Risk and Expected Shortfall

Value at risk (VaR) for return on a portfolio or an asset is defined as the $\upsilon-$quantile of the distribution of returns with a negative value corresponding to a loss. Here the quantile function $F^{-1}_{AEP}(\upsilon \mid \alpha, p_1, p_2)$ of (11) provides VaR at υ for the historical distribution of returns in the AEPD class, i.e., $VaR_{AEP}(\upsilon) \equiv F^{-1}_{AEP}(\upsilon \mid \alpha, p_1, p_2)$. The Expected Shortfall (ES) of a standard AEPD random variable X,

$$ES_{AEP}(q) \equiv E(-X \mid X < q),$$

also called Conditional Value at Risk (CVaR) represents the negative expected return (or loss) conditional on it being below the threshold q. It can be expressed in terms of the gamma CDFs with parameters $1/p_1$, $2/p_1$, $1/p_2$, and $2/p_2$:

$$ES_{AEP}(q) = \begin{cases} 2\alpha^* C(p_1) \left[\dfrac{1-G\left(\frac{1}{p_1}\left|\frac{q}{2\alpha^*}\right|^{p_1};\, 2/p_1\right)}{1-G\left(\frac{1}{p_1}\left|\frac{q}{2\alpha^*}\right|^{p_1};\, 1/p_1\right)} \right], & q \le 0; \\ \dfrac{2\alpha\alpha^* C(p_1) - 2(1-\alpha)(1-\alpha^*)C(p_2)G\left(\frac{1}{p_2}(\frac{|q|}{2(1-\alpha^*)})^{p_2};\, 2/p_2\right)}{\alpha + (1-\alpha)G\left(\frac{1}{p_2}(\frac{|q|}{2(1-\alpha^*)})^{p_2};\, 1/p_2\right)}, & q > 0, \end{cases} \tag{20}$$

where $C(p) \equiv p^{1/p}\Gamma(2/p)/\Gamma(1/p)$, $G(x;\gamma)$ is the gamma cdf given in (9). Recall that $G^{-1}(x;\gamma)$ is the inverse function of $G(x;\gamma)$. For $q = VaR_{AEP}(\upsilon)$, the ES as a function of confidence level υ, denoted by $ES^*_{AEP}(\upsilon)$, can be expressed as follows: $ES^*_{AEP}(\upsilon) =$

$$\begin{cases} \frac{2}{\upsilon}\alpha\alpha^* C(p_1)\left\{1 - G\left[G^{-1}\left(\frac{\alpha-\upsilon}{\alpha}; \frac{1}{p_1}\right); 2/p_1\right]\right\}, & \upsilon \le \alpha \\ \frac{2}{\upsilon}\left\{\alpha\alpha^* C(p_1) - (1-\alpha)(1-\alpha^*)C(p_2)G\left[G^{-1}\left(\frac{\upsilon-\alpha}{1-\alpha}; \frac{1}{p_2}\right); 2/p_2\right]\right\}, & \upsilon > \alpha \end{cases}$$

In practice ES is often used in the following form:

$$E(q - X \mid X < q) = q + E(-X \mid X < q), \tag{21}$$

which is the average loss when an asset return falls below q; the expression follows from $ES_{AEP}(q)$ or $ES^*_{AEP}(\upsilon)$.

4.3 Maximum entropy property

In a distribution class maximum entropy is achieved by a distribution that encodes information in the least biased way without giving any preferential measure weight to any part of the distribution (other than what is required by the distribution class itself). Here we consider a class of absolutely continuous distributions with specific shape (moment) constraints on the left and right deviations and show that the AEPD as defined in (6) has the maximum entropy property in that class.

Specifically consider for parameters $\theta = (\alpha, p_1, p_2, \mu, \sigma)$ an absolute deviation function of $x \in R$ scaled differently on two sides of μ :

$$y(x) = L(x;\theta) + R(x;\theta),$$

with

$$L(x;\theta) = \frac{\Gamma(1+1/p_1)\,|x-\mu|}{\alpha\sigma} 1(x<\mu),\ \ R(x;\theta) = \frac{\Gamma(1+1/p_2)\,|x-\mu|}{(1-\alpha)\sigma} 1(x>\mu).$$

Define a class $\Omega(\alpha, p_1, p_2, \mu, \sigma)$ of absolutely continuous distributions having densities $p(x)$ with support $(-\infty, +\infty)$ that satisfy the following moment constraints on the left and right deviations of $y(x)$:

$$\begin{aligned} \|L(x;\theta)\|_{p_1} &= \left(\int y(x)^{p_1} 1(x<\mu) p(x) dx\right)^{1/p_1} = \left(\tfrac{\alpha}{p_1}\right)^{1/p_1}; \\ \|R(x;\theta)\|_{p_2} &= \left(\int y(x)^{p_2} 1(x>\mu) p(x) dx\right)^{1/p_2} = \left(\tfrac{1-\alpha}{p_2}\right)^{1/p_2}. \end{aligned} \tag{22}$$

This class allows for the location μ, scale σ and three shape parameters α, p_1, p_2 that produce different effects: when $p_1 = p_2$ parameter α alone governs which of the sides gets a larger weight, when p_1, p_2 differ the smaller imparts a heavier tail to its side regardless of a. Thus such a class for fixed values of the parameters gives rise to distributions that could fit required properties for shape in terms of the left/right conditional deviations.

Proposition 3 *The AEPD distribution in (6) has maximum entropy in the class* $\Omega(\alpha, p_1, p_2, \mu, \sigma)$.

Proof. See Appendix A.

5 Asymptotic properties of the Maximum Likelihood Estimator

Since AEPD generalizes the EPD and SEPD classes, we note the asymptotic results available for the latter two classes. The MLE for the EPD parameters and its properties are investigated in Agrò (1995) where the information matrix $I(\beta)$ and the covariance matrix are derived; for $p > 2$ consistency, asymptotic normality and efficiency of MLE are proved; other theoretical results for the MLE are available when p is known. Ayebo and Kozubowski (2004) focused on estimators of scale σ and skewness parameter α in the SEPD by assuming that location μ and tail parameter p are known; they gave the expressions for the MLEs of σ and α, showed that they are consistent, asymptotically normal and efficient and provided the asymptotic covariance matrix for this subset of parameters. DiCiccio and Monti (2004) investigated properties of the MLE of all parameters for the Azzalini's SEPD class, but they did not give a closed-form expression for information matrix and did not provide a rigorous proof of asymptotics for the MLEs which is needed due to the non-smoothness of the log-likelihood function. Here we establish consistency, asymptotic normality and efficiency for MLE of all parameters in the AEPD class (which nests EPD and SEPD) with $p_1 > 1$ and $p_2 > 1$, and provide a closed-form asymptotic covariance matrix of the MLE.

Suppose that the true density $f(y \mid \theta_0)$ with $\theta_0 = (\alpha_0, p_{01}, p_{02}, \mu_0, \sigma_0)$ belongs to the AEPD class (given in (6)) with parameter vector θ in a parameter space $\Theta \subset \Xi \equiv \{\theta \mid \theta = (\alpha, p_1, p_2, \mu, \sigma),\ \sigma,\ p_1,\ p_2 > 0,\ \alpha \in (0,1),\ \mu \in R\}$, where Θ is assumed to be a compact set and θ_0 to be an interior point of Θ. For a random sample $y = (y_1, y_2, ..., y_T)$, the log-likelihood function $l_T(\theta \mid y) \equiv \sum_{t=1}^{T} \ln f(y_t \mid \theta)$ is given as follows:

$$\begin{aligned} l_T(\theta \mid y) = -T \ln \sigma - \textstyle\sum_{t=1}^{T} \left(\frac{\Gamma(1+1/p_1)(\mu - y_t)}{\alpha\sigma} \right)^{p_1} 1(y_t \le \mu) \\ - \textstyle\sum_{t=1}^{T} \left(\frac{\Gamma(1+1/p_2)(y_t - \mu)}{(1-\alpha)\sigma} \right)^{p_2} 1(y_t > \mu). \end{aligned}$$

Note that the AEPD does not satisfy the regularity conditions under which the ML estimator has the usual $\sqrt{T}-$asymptotics, because of a non-differentiable likelihood function. However, we nonetheless establish consistency of the MLE by using Theorem 2.5 in Newey and McFadden (1994, p 2131) and under certain parameter restrictions establish the usual asymptotic normality for the AEPD's MLE by using Theorem 3 as well as its corollary in Huber (1967).

Proposition 4 *(Consistency of MLE). The MLE $\widehat{\theta}$ of θ_0 is consistent, i.e., $\widehat{\theta} \to^p \theta_0$.*

Proof. See Appendix C.

Proposition 5 *(Asymptotic normality of MLE). Suppose that $p_{01} > 1$ and $p_{02} > 1$. Then the MLE $\widehat{\theta}$ of θ_0 is asymptotically normal, i.e.,*

$$\sqrt{T}(\widehat{\theta} - \theta_0) \to^d N(0, I^{-1}(\theta_0)),$$

where $I(\theta_0)$ is the Fisher information matrix:

$$I(\theta_0) \equiv E[(\nabla_\theta \ln f(Y_t \mid \theta_0))(\nabla_\theta \ln f(Y_t \mid \theta_0))'];$$

it can be consistently estimated by $I(\widehat{\theta})$. The closed form expression for $I(\theta_0)$ is provided in Appendix B.

Proof. See Appendices B and C.

The information matrix for the MLE of the SEPD is also given in Appendix B; to our knowledge these results were not available in the literature so far. From the proof in the Appendix, the information matrix equality $I(\theta_0) = -H(\theta_0)$ holds only for $p_{01} > 1/2$ and $p_{02} > 1/2$, because $E[\frac{\partial \ln f}{\partial \mu}]^2$ as an element of $I(\theta_0)$ may not exist or may be negative for some points (p_{01}, p_{02}) in $(0, 1/2] \times (0, 1/2]$. Since $I(\theta)$ is continuous for all $\theta \in \Xi$ satisfying $p_1 > 1/2$ and $p_2 > 1/2$, it follows from the consistency of $\widehat{\theta}$ that $I(\widehat{\theta})$ is a consistent estimator of $I(\theta_0)$. The restriction of $p_{01} > 1$ and $p_{02} > 1$ is due to the requirement involved the estimation of location parameter. The restriction is not an impediment in most applications. Even for the GARCH option pricing model with GED conditional distribution in Duan (1999), this restriction is imposed in order to ensure the existence of the expected simple return. If μ_0 is known, then the usual $\sqrt{T}-$asymptotics hold for the MLEs of other parameters $(\alpha_0, p_{01}, p_{02}, \sigma_0)$ without any restrictions. When location parameter μ_0 can be consistently estimated by a nonparametric method (see Andrews et al. (1972) and Bickel (2002)), the MLEs of other parameters are still consistent, asymptotically normal but may not be efficient.

6 Performance of MLE in Simulation

A stochastic representation of a distribution is important to simulation studies. For given values of parameters, p_1, p_2 and α $(0 < \alpha < 1,\ p_i > 0,\ i = 1, 2)$, we can generate standard AEPD random numbers by the following method: first, generate three random numbers U, W_1 and W_2, where U is drawn from standard uniform distribution $U(0, 1)$ and W_i $(i = 1, 2)$ is from the gamma distribution with shape parameter $1/p_i$ and pdf $f_{W_i}(w) = \Gamma(1/p_i)^{-1} w^{1/p_i - 1} \exp(-w)$; second, define a random variable Y :

$$\begin{aligned} Y &= \frac{\alpha}{\Gamma(1+1/p_1)} W_1^{1/p_1} \left[\frac{sign(U-\alpha)-1}{2}\right] \\ &\quad + \frac{1-\alpha}{\Gamma(1+1/p_2)} W_2^{1/p_2} \left[\frac{sign(U-\alpha)+1}{2}\right] \end{aligned} \tag{23}$$

where $sign(x) = +1$ if $x > 0$, -1 if $x < 0$, and 0 if $x = 0$. It is straightforward to show that random variable Y has the density (6) of standard AEPD (location $\mu = 0$, scale $\sigma = 1$). An alternative method is the inverse method, i.e., using $Y = F_{AEP}^{-1}(U)$ to generate standard AEPD random numbers, where U is a standard uniform random variable and F_{AEP} is the standard AEPD cdf. However, this method is very time-consuming, while the method given in (23) allows us to generate AEPD random numbers more quickly in Matlab.

To assess the asymptotic properties of the MLE in finite samples, following Agrò (1995), a numerical investigation of bias and variance of MLEs was made using sample sizes of T = 500, 1000, 2000, 4000, 8000. We choose $\mu_0 = 0$, $\sigma_0 = 1$ and various different true values of (α, p_1, p_2): $\alpha = 0.3$, 0.5 and $p_i = 0.7$, 1, 1.5, 2.5 $(i = 1, 2)$. To save space, here we only report the cases of $\alpha = 0.3$ and $p_2 = 1$, 1.5. For each set of true values of parameters and every sample size, $N = 2000$ replications are drawn from the AEPD with the set of parameter values, and then $N = 2000$ ML estimates $\widehat{\theta}^i$ $(i = 1, 2, ..., N)$ are obtained using these samples. So, we can estimate the means and standard deviations of the MLEs of parameters, denoted respectively by $M(\widehat{\theta})$ and $STD(\widehat{\theta})$,

$$M(\widehat{\theta}) = \frac{1}{N}\sum_{i=1}^{N} \widehat{\theta}^i, \quad STD(\widehat{\theta}) = \left(\frac{1}{N}\sum_{i=1}^{N} \left[\widehat{\theta}^i - M(\widehat{\theta})\right]^2\right)^{1/2},$$

and compare these estimated standard deviations with their theoretical values which are taken from the square root of the diagonal elements of Cramer-Rao bound (i.e., $I^{-1}(\widehat{\theta})/T$). Simulation results are presented in the Table (see Appendix D). All entries labeled "Mean of MLEs" report $M(\widehat{\theta})$, and those in "STD Ratio" rows are the ratios of simulated standard deviations $STD(\widehat{\theta})$ to the theoretical ones from $I^{-1}(\widehat{\theta})/T$.

From our simulation studies, we can see that the estimates $\widehat{\theta}$ of all parameters seem asymptotically unbiased for all given true values, and that their variance seems to be approaching the Cramer-Rao bound for the cases of $p_{01} > 1$ and $p_{02} > 1$ (see the cases in which $(p_{01}, p_{02}) = (1.5, 1.5)$ and $(2.5, 1.5)$). However, for the cases of $p_{01} \leq 1$ or $p_{02} \leq 1$, the behavior of the variance appears to be problematic. More specifically, although the estimates of scale σ appear always efficient in all cases, there are significantly large ratios of standard deviation for the other parameters, especially for μ and α; but the larger the values of the tail parameters, p_1 and p_2, the more efficient the estimates of α and μ appear to be. Other observed phenomena are that (1), because of fewer observations on the left side, estimates of the left tail parameter p_1 have slower convergence and lower efficiency than those of the right tail p_2 (see the cases in which $p_{01} = p_{02} = 1$ or 1.5); (2), in general, the MLE is more efficient in the cases with larger tail parameter p_1 or p_2 than for those with smaller tail parameters. Finally, we want to point out that for a small sample, say a size less than 500, the likelihood function may not have any maximum point. This problem still exists for the GED and is discussed in detail in Agrò (1995).

References

[1] Abramowitz, M. and Stegun, I. A. (1970). Handbook of Mathematical Functions, National Bureau of Standard, Applied Math. Series 55, U.S. Department of Commerce.

[2] Agrò, G. (1995). Maximum likelihood estimation for the exponential power function parameters, *Comm. Simulation Comput.* 24(2), 523-536.

[3] Andrews, D.F., Bickel, P.J., Hampel, F.R., Huber, P.J., Rogers, W.H., and Tukey, J.W. (1972). *Robust Estimates of Location,* Princeton University Press, Princeton.

[4] Arnold, B.C. and Groeneveld, R.A. (1995). Measuring skewness with respect to the mode. *Am. Statistitian*, 49, 34-38.

[5] Artin, E. (1964). *The Gamma Function*, Holt, Rinehart and Winston, Inc.

[6] Ayebo, A. and Kozubowski, T.J. (2004). An asymmetric generalization of Gaussian and Laplace laws. *Journal of Probability and Statistical Science,* 1(2), 187-210.

[7] Azzalini, A. (1986). Further results on a class of distributions which includes the normal ones, *Statistica*, 46,199-208.

[8] Bickel, D.R. (2002). Robust estimators of the mode and skewness of continuous data, *Comput. Statist. Data Anal.* 39, 153-163.

[9] Box, G.E.P & Tiao, G.C. (1973). *Bayesian inference in statistical analysis.* Addison Wesley ed.

[10] Bradley, B.O. and Taqqu, M.S. (2003). Financial risk and heavy tails. *Handbook of Heavy Tailed Distributions in Finance*, edited by S. T. Rachev.

[11] Christoffersen, P. F. (2003). Elements of Financial Risk Management. Academic Press.

[12] Christoffersen, P. F., Jacobs, K. and Wang, Y. (2005). Volatility components: affine restrictions and non-normal innovations. Working paper, McGill University.

[13] DiCiccio, T and Monti, A. C. (2004). Inferential aspects of the skew exponential power distribution. *Journal of the American Statistical Association*, 99(466), 439-450.

[14] Duan, J-C. (1999). Conditionally fat-tailed distributions and the volatility smile in options. Working paper, University of Toronto.

[15] Farrell, O. J. and Ross, B. (1963). Solved Problems: Gamma and Beta Functions, Legendre Polynomials, Bessel Functions. The Macmillan Company, Now York.

[16] Fernandez, C., Osiewalski, J. and Steel, M.F.J. (1995), *Modeling and inference with $v-$distributions, J. Amer. Statist. Assoc. Vol 90, No. 432, 1331-1340.*

[17] Gradshteyn, I.S. and I.M.Ryzhik (1994). Tables of integrals, Series and Products, Academic Press.

[18] Hsieh, D. (1989). Modeling heteroskedasticity in daily foreign exchange rates. *Journal of Business and Economics Statistics,* 7, 307-317.

[19] Huber, P. J. (1967). The Behavior of Maximum Likelihood Estimates Under Nonstandard Conditions, in: L.M. LeCam and J. Neyman, eds., *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Berkeley: University of California Press.*

[20] Kagan, A. M., Linnik, Yu. V. and Rao, C. R. (1973). *Characterization Problems in Mathematical Statistics*, Wiley, New York.

[21] Komunjer, I. (2007). Asymmetric power distribution: Theory and applications to risk measurement, *Journal of Applied Econometrics,* 22, 891-921.

[22] Kotz, S., Kozubowski, T.J. and Podgorski, K. (2001), *The Laplace Distribution and Generalizations: A Revisit with Applications to Communications, Economics, Engineering, and Finance, Birkhauser Boston.*

[23] Loève, M. (1977). Probability theory I, 4th edition.

[24] Mineo, A. (1989). The norm-p estimation of location, scale and simple linear regression parameters, *Lecture notes in statistics, Statistical modelling,* No. 57, Springer-Verlag, 222-233.

[25] Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, 59 (2), 347-370.

[26] Newey, W.K. and McFadden, D. (1994). Large Sample Estimation and Hypothesis Testing, in: R. Engle and D. McFadden, eds., *Handbook of Econometrics, Vol. 4,* Amsterdam: North-Holland.

[27] Rachev, S. T. and Mittinik, S. (2000). Stable Paretian Models in Finance. Wiley.

[28] Subbotin, M. T. (1923). On the law of frequency of error. *Matematicheskii Sbornik,* 31, 296-301.

[29] Theodossiou, P. (2000). Skewed Generalized Error Distribution of Financial Assets and Option Pricing. SSRN working paper.

7 Appendix A

In the proofs we make extensive use of several results.

(I). The following integral result (see Gradshteyn and I.M.Ryzhik,1994. #3.478) is useful to derive moments of the EPD.

$$\int_0^\infty x^{v-1}\exp(-\mu x^p)dx = \frac{1}{p}\mu^{-v/p}\Gamma(\frac{v}{p}),\ \text{for } \mu > 0, v > 0,\ p > 0. \tag{24}$$

Derivation of the normalizing constant of the AEPD density is based on applying the integral (24) to $v = 1$ and $\mu = [p(2\alpha\sigma)^p]^{-1}$.

(II). Properties of the gamma function $\Gamma(x)$. The Gamma function $\Gamma(x)$ on the domain of the definition $(0, +\infty)$ has the following basic properties: (i) $\Gamma(x+1) = x\Gamma(x)$, (ii) $\Gamma(x)$ is strictly convex, and $\Gamma(x) \to +\infty$ as $x \to 0^+$ or $+\infty$ (see Farrell & Ross, 1963, p22), (iii) $\Gamma(x)$ has derivatives of arbitrarily high order (see Artin, 1964, p16). Define $\Psi(x) \equiv \Gamma'(x)/\Gamma(x)$, called *digamma function*, then

$$\Psi(x) = -C - \frac{1}{x} + \sum_{i=1}^{\infty}(\frac{1}{i} - \frac{1}{x+i}), \tag{25}$$

$$\frac{d^{k-1}\Psi(x)}{dx^{k-1}} = \sum_{i=0}^{\infty}\frac{(-1)^k(k-1)}{(x+i)^k},\ \text{for } k \geq 2, \tag{26}$$

where C is Euler's constant. More properties and details can be found in Abramowitz and Stegun (1970, p255-263), Artin (1964, p16-26) and Farrell & Ross (1963). By differentiating both sides of $\Gamma(x+1) = x\Gamma(x)$ and $\Psi(x) \equiv \Gamma'(x)/\Gamma(x)$, we easily get the following equalities:

$$\Gamma'(x+1)/\Gamma(x) = 1 + x\Psi(x), \tag{27}$$

$$\Gamma''(x)/\Gamma(x) = \Psi'(x) + \Psi^2(x), \tag{28}$$

$$\Gamma''(x+1)/\Gamma(x) = 2\Psi(x) + x\Psi^2(x) + x\Psi'(x). \tag{29}$$

(III). Properties of the EPD (see Box & Tiao, 1973; Kotz, et al., 2001). The following results for the EPD can be obtained by straightforward calculations.

Suppose that Z_p has the standard EPD density ($\mu = 0, \sigma_p = 1$), which is defined in (1). Then its cdf and quantile function are expressed as

$$F_{EP}(x \mid p) = \frac{1}{2}\left[1 + sign(x)G\left(\frac{1}{p}|x|^p; \frac{1}{p}\right)\right], \tag{30}$$

$$F_{EP}^{-1}(\upsilon \mid p) = sign(2\upsilon - 1)\left[pG^{-1}\left(|2\upsilon - 1|; \frac{1}{p}\right)\right]^{1/p}, \tag{31}$$

where $G(x;\gamma)$ is the gamma cdf, and $G^{-1}(x;\gamma)$ is the inverse function of $G(x;\gamma)$. By using the change of variable and (24), the absolute moment of Z_p is given by

$$E(|Z_p|^\delta) = p^{\delta/p}\Gamma(\frac{\delta+1}{p})/\Gamma(1/p) \equiv [M(p,\delta)]^\delta, \quad \delta > -1, \tag{32}$$

where $M(p,\delta)$ is defined in (d) of Proposition 1. The expected shortfall of Z_p, $ES_{EP}(x \mid p) \equiv E(-Z_p \mid Z_p < x)$, is given as follows:

$$ES_{EP}(x \mid p) = p^{1/p}\frac{\Gamma(2/p)}{\Gamma(1/p)}\left[\frac{1 - G(\frac{1}{p}|x|^p; 2/p)}{1 + sign(x)G(\frac{1}{p}|x|^p; 1/p)}\right]. \tag{33}$$

Taking $x = VaR_{EP}(\upsilon) \equiv F_{EP}^{-1}(\upsilon \mid p)$, the ES as a function of confidence level υ, $ES_{EP}^*(\upsilon \mid p)$, can be expressed as

$$ES_{EP}^*(\upsilon \mid p) = p^{1/p}\frac{\Gamma(2/p)}{\Gamma(1/p)}\frac{1}{2\upsilon}\left\{1 - G\left[G^{-1}\left(|2\upsilon - 1|; \frac{1}{p}\right); 2/p\right]\right\}. \tag{34}$$

Proof of Proposition 1.

The result $P(X \leq \mu) = \alpha$ follows directly from (10). Proofs of all other equalities in (a), (b) and (d) of Proposition 1 boil down to calculations of $d_L(r)$ and $d_R(r)$. Consider $d_R(r)$. For the standard AEPD ($\mu = 0$, $\sigma = 1$), by using the change of variable and (32) or (24), we have

$$\begin{aligned}
d_R(r) &= \{E[|X|^r \mid X>0]\}^{1/r} \\
&= \left\{\int_0^\infty x^r \frac{1-\alpha}{1-\alpha^*} K_{EP}(p_2) \exp\left(-\frac{1}{p_2}\left|\frac{x}{2(1-\alpha^*)}\right|^{p_2}\right) \frac{1}{1-\alpha} dx\right\}^{1/r} \\
&= 2(1-\alpha^*)\left\{2\int_0^\infty z^r K_{EP}(p_2)\exp\left(-\frac{1}{p_2} z^{p_2}\right) dz\right\}^{1/r} \\
&= 2(1-\alpha^*)\{E(|Z_{p_2}|^r)\}^{1/r} = 2(1-\alpha^*)M(p_2, r),
\end{aligned}$$

where the last equality follows from (32).

To prove that $\xi(p,r) \equiv K_{EP}(p)M(p,r)$ is strictly increasing in r, we evaluate the derivative of $\ln M(p,r)$ with respect to r and show $\frac{\partial \ln M(p,r)}{\partial r} > 0$. Note that

$$\ln M(p,r) = \frac{1}{p}\ln p + \frac{1}{r}[\ln\Gamma((r+1)/p) - \ln\Gamma(1/p)],$$

and

$$\frac{\partial \ln M(p,r)}{\partial r} = \frac{1}{pr}\Psi(\frac{r+1}{p}) - \frac{1}{r^2}[\log\Gamma((r+1)/p) - \log\Gamma(1/p)].$$

By the mean value theorem, we have

$$\frac{\partial \ln M(p,r)}{\partial r} = \frac{1}{pr}[\Psi(\frac{r+1}{p}) - \Psi(\frac{\varepsilon r+1}{p})], \text{ where } 0 < \varepsilon < 1.$$

Since $\Psi'(x)$ is positive for any $x > 0$ (see Abramowitz and Stegun,1970, p260, 6.4.10), implying that $\Psi(x)$ is strictly increasing, it follows that $\frac{\partial \ln M(p,r)}{\partial r} > 0$ for any $r > 0$ and $p > 0$.

Now we show that $\xi(p,r)$ is strictly decreasing in p for any given $r > 0$. Similarly, we evaluate derivative of $\ln\xi(p,r)$ with respect to p,

$$\frac{\partial \ln\xi(p,r)}{\partial p} = \frac{1}{p^2}\Psi(1+1/p) + \frac{1}{pr}\left[\frac{1}{p}\Psi(\frac{1}{p}) - \frac{r+1}{p}\Psi(\frac{r+1}{p})\right],$$

and note that the second part of the above expression, denoted by $h(p,r)$,

$$\begin{aligned}
h(p,r) &\equiv \frac{1}{pr}\left[\frac{1}{p}\Psi(\frac{1}{p}) - \frac{r+1}{p}\Psi(\frac{r+1}{p})\right] \\
&= \frac{C}{p^2} + \sum_{i=1}^{\infty}\frac{1}{pr}\left[g_i(\frac{1}{p}) - g_i(\frac{r+1}{p})\right], \qquad (35)
\end{aligned}$$

where $g_i(x) \equiv x/i - x/(i+x)$ and C is Euler's constant. The expression (35) for $h(p,r)$ is due to (25). Using the mean value theorem and the facts $g_i'(x) > 0$ and $g_i''(x) > 0$ for any $x > 0$ and all $i \geq 1$, we can prove that $h_i(p,r) \equiv$

$\frac{1}{pr}\left[g_i(\frac{1}{p}) - g_i(\frac{r+1}{p})\right]$ is strictly decreasing in r for every $i \geq 1$; so $h(p,r)$ is also a decreasing function of r. Therefore, it follows that for $r > 0$

$$\begin{aligned}\frac{\partial \ln \xi(p,r)}{\partial p} &< \frac{1}{p^2}\Psi(1+1/p) + \lim_{r\to 0^+}\frac{1}{pr}\left[\frac{1}{p}\Psi(\frac{1}{p}) - \frac{r+1}{p}\Psi(\frac{r+1}{p})\right] \\ &= \frac{1}{p^2}\left[\Psi(1+\frac{1}{p}) - \frac{1}{p}\Psi'(\frac{1}{p}) - \Psi(\frac{1}{p})\right],\ \forall r > 0 \\ &= \frac{1}{p^2}\left[p - \frac{1}{p}\Psi'(\frac{1}{p})\right] = -\left(\frac{1}{p}\right)^3\sum_{i=1}^{\infty}\frac{1}{(i+1/p)^2} < 0,\end{aligned}$$

where we use the equality $\Psi(1+x) = \Psi(x) + 1/x$ from (27), and the expression (26) for $\Psi'(x)$ in the last equality.

To prove Proposition 1-(c), we define an increasing function $r = r^*(c \mid p) \equiv \xi^{-1}(c \mid p)$ for a given p. Note that $\xi(p,r) \downarrow lb(p)$ as $r \downarrow 0$ and $\xi(p,r) \uparrow +\infty$ as $r \uparrow +\infty$ (this follows by using Equality 6.1.20 in Abramowitz and Stegun,1970, p256). When $c > lb(p)$, $r = \xi^{-1}(c \mid p) > 0$ and thus $r^*(c \mid p_1) > 0$ and $r^*(c \mid p_2) > 0$ for any $c > \max\{lb(p_1), lb(p_2)\}$. Combining definition of $r^*(c \mid p)$ with equalities in Proposition 1-(d), we get Proposition 1-(c).■

Proof of Proposition 2.

The expressions for $kur_L(r)$ and $kur_R(r)$ in Proposition 2 are easily obtained by straightforward calculations using equalities in (d) of Proposition 1. Here we prove only that $k(r,p) \equiv \Gamma(\frac{1}{p})\Gamma(\frac{2r+1}{p})/\Gamma^2(\frac{r+1}{p})$ is strictly decreasing in p and increasing in r. The second point is easily shown by the fact that

$$\frac{\partial \ln k(r,p)}{\partial r} = \frac{2}{p}[\Psi(\frac{2r+1}{p}) - \Psi(\frac{r+1}{p})],\ p > 0, r > 0$$

and that $\Psi(\cdot)$ is strictly increasing (see (26), $\Psi'(x) > 0$), implying $\frac{\partial \ln k(r,p)}{\partial r} > 0$ for any $r > 0$ and $p > 0$. The first point is obtained by noting that

$$\frac{\partial \ln k(r,p)}{\partial p} = \frac{1}{p}[-\rho(\frac{2r+1}{p}) - \rho(\frac{1}{p}) + 2\rho(\frac{r+1}{p})],$$

where $\rho(x) = x\Psi(x)$ is strictly convex in $(0,+\infty)$ (it follows from (26) that $\rho''(x) = 2\sum_{i=0}^{\infty} i/(i+x)^3 > 0$ for any $x > 0$), implying $\frac{\partial \ln k(r,p)}{\partial p} < 0$ for any $p > 0$ and $r > 0$.■

Proof of Proposition 3.

The entropy of a distribution with density $f(x;\theta)$ is by definition

$$H(f) \equiv -\int_{-\infty}^{+\infty} f(x;\theta)\ln f(x;\theta)dx.$$

A straightforward calculation shows that for AEPD

$$H(f) = \ln\sigma + \frac{\alpha}{p_1} + \frac{1-\alpha}{p_2}.$$

By Theorem 13.2.1 of Kagan et al. (1973, page 409) (also see Proposition 2.4.6 in Kotz et al. (2001, p51)), among all continuous distributions whose densities p have support $(-\infty, +\infty)$ and that satisfy the constraints:

$$\int_{-\infty}^{+\infty} [L(x;\theta)]^{p_1} p(x)dx = \frac{\alpha}{p_1}; \quad \int_{-\infty}^{+\infty} [R(x;\theta)]^{p_2} p(x)dx = \frac{1-\alpha}{p_2}, \tag{36}$$

the maximum entropy is attained by distributions with the density of the form

$$p_{ME}(x) = \exp\{-\lambda_0 - \lambda_1[L(x;\theta)]^{p_1} - \lambda_2[R(x;\theta)]^{p_2}\}$$

(and only by them), if there exist constants λ_0, λ_1 and λ_2 such that the density $p_{ME}(x) > 0$ for all $x \in (-\infty, +\infty)$ and satisfies the conditions in (36). In fact, we need only to show that there exists a unique set of constants such that $p_{ME}(x)$ is exactly the form (6) of the AEPD density. From the conditions in (36) and $\int_{-\infty}^{+\infty} p_{ME}(x)dx = 1$, by changing the variable in an integral, some straightforward calculations show that

$$\frac{\alpha\sigma}{\lambda_1^{1/p_1}} + \frac{(1-\alpha)\sigma}{\lambda_2^{1/p_2}} = \frac{\sigma}{\lambda_1^{1+1/p_1}} = \frac{\sigma}{\lambda_2^{1+1/p_2}} = \exp(\lambda_0),$$

implying

$$\alpha\lambda_1 + (1-\alpha)\lambda_2 = 1, \quad \lambda_1^{1+1/p_1} = \lambda_2^{1+1/p_2}. \tag{37}$$

Obviously, both the equations in (37) uniquely determine a set of $(\lambda_1, \lambda_2) = (1, 1)$ because λ_2 as a function of λ_1 is strictly decreasing by the first equation in (37) and increasing by the second, and thus $\lambda_0 = \ln\sigma$.■

8 Appendix B

Appendix B is devoted to deriving a closed-form expression for the information matrix and to verifying the information matrix equality under certain conditions. Expectations are always taken with respect to the true underlying distribution $f(y;\theta_0)$, where $\theta_0 = (\alpha_0, p_{01}, p_{02}, \mu_0, \sigma_0)$.

Suppose that y_t $(t = 1, 2, ...T)$ are i.i.d. observations from the AEPD whose density $f(y_t;\theta)$ with $\theta \in \Xi$ is defined in (6). Let

$$\begin{aligned} L(y_t;\theta) &\equiv \frac{\Gamma(1+1/p_1)\,|\mu - y_t|}{\alpha\sigma} 1(y_t < \mu), \\ R(y_t;\theta) &\equiv \frac{\Gamma(1+1/p_2)\,|y_t - \mu|}{(1-\alpha)\sigma} 1(y_t > \mu). \end{aligned}$$

Then the log-density function $\ln f(y_t;\theta) = -\ln\sigma - [L(y_t;\theta)]^{p_1} - [R(y_t;\theta)]^{p_2}$, and the score (vector) for observation t, $\frac{\partial}{\partial\theta}\ln f(y_t;\theta)$, is given by

$$\begin{aligned}
\frac{\partial \ln f}{\partial \alpha} &= \frac{p_1}{\alpha}[L(y_t;\theta)]^{p_1} - \frac{p_2}{1-\alpha}[R(y_t;\theta)]^{p_2}, \qquad (38)\\
\frac{\partial \ln f}{\partial p_1} &= \left[\frac{1}{p_1}\Psi(1+1/p_1) - \ln L(y_t;\theta)\right][L(y_t;\theta)]^{p_1},\\
\frac{\partial \ln f}{\partial p_2} &= \left[\frac{1}{p_2}\Psi(1+1/p_2) - \ln R(y_t;\theta)\right][R(y_t;\theta)]^{p_2},\\
\frac{\partial \ln f}{\partial \mu} &= -\frac{\Gamma(1/p_1)}{\alpha\sigma}[L(y_t;\theta)]^{p_1-1} + \frac{\Gamma(1/p_2)}{(1-\alpha)\sigma}[R(y_t;\theta)]^{p_2-1},\\
\frac{\partial \ln f}{\partial \sigma} &= \frac{p_1}{\sigma}[L(y_t;\theta)]^{p_1} + \frac{p_2}{\sigma}[R(y_t;\theta)]^{p_2} - \frac{1}{\sigma},
\end{aligned}$$

where for $x = 0$ and $p > 0$ set $x^p \ln x = 0$. To derive the information matrix $I(\theta_0) \equiv E[\frac{\partial}{\partial\theta}\ln f(y_t,\theta_0)\frac{\partial}{\partial\theta'}\ln f(y_t;\theta_0)]$ and the Hessian $H(\theta_0) \equiv E[\frac{\partial^2}{\partial\theta\partial\theta'}\ln f(y_t;\theta_0)]$ and to verify the information matrix equality $I(\theta_0) = -H(\theta_0)$, we first give the following Lemmas.

Lemma 6 *For any real number $r > -1$ and integer $m = 0, 1, 2$, we have*

$$E[L(y_t;\theta_0)]^r[\ln L(y_t;\theta_0)]^m 1(y_t < \mu_0) = \frac{\alpha_0}{p_{01}^{m+1}}\frac{\Gamma^{(m)}((1+r)/p_{01})}{\Gamma(1+1/p_{01})}, \qquad (39)$$

$$E[R(y_t;\theta_0)]^r[\ln R(y_t;\theta_0)]^m 1(y_t > \mu_0) = \frac{1-\alpha_0}{p_{02}^{m+1}}\frac{\Gamma^{(m)}((1+r)/p_{02})}{\Gamma(1+1/p_{02})}, \qquad (40)$$

where $\Gamma^{(m)}(\cdot)$ is the mth order derivative of gamma function $\Gamma(\cdot)$ and $\Gamma^{(0)}(\cdot)$ means $\Gamma(\cdot)$.

Proof:[2] Both the equalities (39) and (40) can be proved in the same manner. So here we only show equality (39). Denote by EL the expectation of the left hand side of (39) and note that

$$\begin{aligned}
EL &= \int_{-\infty}^{\mu}[L(y;\theta)]^r[\ln L(y;\theta)]^m f(y;\theta)dy\\
&= \int_{-\infty}^{\mu}[L(y;\theta)]^r[\ln L(y;\theta)]^m \frac{1}{\sigma}\exp\{-[L(y;\theta)]^{p_1}\}dy.
\end{aligned}$$

Then a change of variable $x = [L(y;\theta)]^{p_1}$ results in

$$\begin{aligned}
EL &= \frac{\alpha}{p_1^{m+1}\Gamma(1+1/p_1)}\int_0^{+\infty} x^{(1+r)/p_1-1}(\ln x)^m \exp(-x)dx\\
&= \frac{\alpha}{p_1^{m+1}\Gamma(1+1/p_1)}\Gamma^{(m)}((1+r)/p_1),
\end{aligned}$$

where we used the expression for derivatives of gamma function (see Farrell and Ross, 1963, p22).■

[2] For simplicity, we omit the subscript "0" on the true parameters in all the following proofs.

Lemma 7 *The score vector for observation t,* $\frac{\partial}{\partial\theta}\ln f(y_t;\theta)$*, satisfies the equation*

$$E[\frac{\partial}{\partial\theta}\ln f(y_t;\theta_0)] = 0. \tag{41}$$

Proof. It is very easy to verify this by using (39), (40) and equalities (27)-(29) in appendix A-(II).

(i).

$$\begin{aligned} E[\frac{\partial \ln f}{\partial \alpha}] &= \frac{p_1}{\alpha}E[L(y_t;\theta)]^{p_1} - \frac{p_2}{1-\alpha}E[R(y_t;\theta)]^{p_2} \\ &= \frac{p_1}{\alpha}\frac{\alpha\Gamma(1+1/p_1)}{p_1\Gamma(1+1/p_1)} - \frac{p_2}{1-\alpha}\frac{(1-\alpha)\Gamma(1+1/p_2)}{p_2\Gamma(1+1/p_2)} \\ &= 1-1=0; \end{aligned}$$

(ii).

$$\begin{aligned} E[\frac{\partial \ln f}{\partial p_1}] &= \frac{1}{p_1}\Psi(1+1/p_1)E[L(y_t;\theta)]^{p_1} - E[L(y_t;\theta)]^{p_1}\ln L(y_t;\theta) \\ &= \frac{1}{p_1}\Psi(1+1/p_1)\frac{\alpha\Gamma(1+1/p_1)}{p_1\Gamma(1+1/p_1)} - \frac{\alpha\Gamma'(1+1/p_1)}{p_1^2\Gamma(1+1/p_1)} \\ &= \frac{\alpha}{p_1^2}\Psi(1+1/p_1) - \frac{\alpha}{p_1^2}\Psi(1+1/p_1) = 0; \end{aligned}$$

(iii). Similarly, we have

$$E[\frac{\partial \ln f}{\partial p_2}] = \frac{1}{p_2}\Psi(1+1/p_2)E[R(y_t;\theta)]^{p_2} - E[R(y_t;\theta)]^{p_2}\ln R(y_t;\theta) = 0;$$

(iv).

$$\begin{aligned} E[\frac{\partial \ln f}{\partial \mu}] &= -\frac{\Gamma(1/p_1)}{\alpha\sigma}E[L(y_t;\theta)]^{p_1-1} + \frac{\Gamma(1/p_2)}{(1-\alpha)\sigma}E[R(y_t;\theta)]^{p_2-1} \\ &= -\frac{\Gamma(1/p_1)}{\alpha\sigma}\frac{\alpha\Gamma(1)}{p_1\Gamma(1+1/p_1)} - \frac{\Gamma(1/p_2)}{(1-\alpha)\sigma}\frac{(1-\alpha)\Gamma(1)}{p_2\Gamma(1+1/p_2)} \\ &= -\frac{1}{\sigma}+\frac{1}{\sigma}=0; \end{aligned}$$

(v).

$$\begin{aligned} E[\frac{\partial \ln f}{\partial \sigma}] &= -\frac{1}{\sigma} + \frac{p_1}{\sigma}E[L(y_t;\theta)]^{p_1} + \frac{p_2}{\sigma}E[R(y_t;\theta)]^{p_2} \\ &= -\frac{1}{\sigma} + \frac{p_1}{\sigma}\frac{\alpha\Gamma(1+1/p_1)}{p_1\Gamma(1+1/p_1)} + \frac{p_2}{\sigma}\frac{(1-\alpha)\Gamma(1+1/p_2)}{p_2\Gamma(1+1/p_2)} \\ &= -\frac{1}{\sigma}+\frac{\alpha}{\sigma}+\frac{1-\alpha}{\sigma} = 0.\blacksquare \end{aligned}$$

Proposition 8 *The information matrix equality $I(\theta_0) = -H(\theta_0)$ holds for $p_{01} > 1/2$ and $p_{02} > 1/2$. The elements of the Fisher information matrix, denoted by ϕ_{ij},*

$$\phi_{ij} \equiv E[\partial \ln f(y_t;\theta_0)/\partial\theta_i] \cdot [\partial \ln f(y_t;\theta_0)/\partial\theta_j], \tag{42}$$

where $\phi_{ij} = \phi_{ji}$ and θ_j represents the jth element of parameter vector $\theta = (\alpha, p_1, p_2, \mu, \sigma)^T$, are given as follows[3]*:*

$$\begin{aligned}
\phi_{11} &= \frac{p_1+1}{\alpha} + \frac{p_2+1}{1-\alpha}, \quad \phi_{12} = -\frac{1}{p_1}, \quad \phi_{13} = \frac{1}{p_2}, \\
\phi_{14} &= -\frac{1}{\sigma}\left(\frac{p_1}{\alpha} + \frac{p_2}{1-\alpha}\right), \quad \phi_{15} = \frac{p_1 - p_2}{\sigma}, \\
\phi_{22} &= \frac{\alpha}{p_1^3}(1+1/p_1)\Psi'(1+1/p_1), \quad \phi_{23} = 0, \\
\phi_{24} &= \frac{1}{\sigma p_1}[\Psi(2) - \Psi(1+1/p_1)], \quad \phi_{25} = -\frac{\alpha}{\sigma p_1}, \\
\phi_{33} &= \frac{1-\alpha}{p_2^3}(1+1/p_2)\Psi'(1+1/p_2), \quad \phi_{34} = -\frac{1}{\sigma p_2}[\Psi(2) - \Psi(1+1/p_2)], \\
\phi_{35} &= -\frac{1-\alpha}{\sigma p_2}, \quad \phi_{44} = \frac{1}{\sigma^2}\left[\frac{\Gamma(1/p_1)\Gamma(2-1/p_1)}{\alpha} + \frac{\Gamma(1/p_2)\Gamma(2-1/p_2)}{(1-\alpha)}\right], \\
\phi_{45} &= \frac{1}{\sigma^2}(p_2 - p_1), \quad \phi_{55} = \frac{\alpha p_1 + (1-\alpha)p_2}{\sigma^2},
\end{aligned}$$

where all the expressions for ϕ_{ij} above are evaluated at the true values $(\alpha_0, p_{01}, p_{02}, \mu_0, \sigma_0)$ of parameters.

Proof. We derive expressions for $E[\partial \ln f(y_t;\theta)/\partial\theta_i] \cdot [\partial \ln f(y_t;\theta)/\partial\theta_j]$ and $E[\partial^2 \ln f(y_t;\theta)/\partial\theta_i\partial\theta_j]$ separately and then verify

$$E\left[\frac{\partial \ln f(y_t;\theta)}{\partial\theta_i} \cdot \frac{\partial \ln f(y_t;\theta)}{\partial\theta_j}\right] = -E\left[\frac{\partial^2 \ln f(y_t;\theta)}{\partial\theta_i\partial\theta_j}\right], \quad i,\ j = 1, 2, ...5.$$

In the proof we use the fact that $1(y_t < \mu)1(y_t > \mu) = 0$ and make repeated use of Lemmas above and properties in Appendix A-(II).

ϕ_{11} is given by

$$\begin{aligned}
E\left[\frac{\partial \ln f}{\partial\alpha}\right]^2 &= \left(\frac{p_1}{\alpha}\right)^2 E[L(y_t;\theta)]^{2p_1} + \left(\frac{p_2}{1-\alpha}\right)^2 E[R(y_t;\theta)]^{2p_2} \\
&= \left(\frac{p_1}{\alpha}\right)^2 \frac{\alpha\Gamma(2+1/p_1)}{p_1\Gamma(1+1/p_1)} + \left(\frac{p_2}{1-\alpha}\right)^2 \frac{(1-\alpha)\Gamma(2+1/p_2)}{p_2\Gamma(1+1/p_2)} \\
&= \frac{1+p_1}{\alpha} + \frac{1+p_2}{1-\alpha}; \text{ also}
\end{aligned}$$

[3] By using $\Gamma(x)\Gamma(1-x) = \pi/\sin(\pi x)$ (see Artin, 1964, p26; or Farrell and Ross, 1963, p39), the element of ϕ_{44} can also be expressed as

$$\phi_{44} = \frac{\pi}{\sigma^2}\left[\frac{1-1/p_1}{\alpha\sin(\pi/p_1)} + \frac{1-1/p_2}{(1-\alpha)\sin(\pi/p_2)}\right].$$

$$
\begin{aligned}
E[\frac{\partial^2 \ln f}{\partial \alpha^2}] &= -\frac{p_1(1+p_1)}{\alpha^2}E[L(y_t;\theta)]^{p_1} - \frac{p_2(1+p_2)}{(1-\alpha)^2}E[R(y_t;\theta)]^{p_2} \\
&= -\frac{p_1(1+p_1)}{\alpha^2}\frac{\alpha\Gamma(1+1/p_1)}{p_1\Gamma(1+1/p_1)} - \frac{p_2(1+p_2)}{(1-\alpha)^2}\frac{(1-\alpha)\Gamma(1+1/p_2)}{p_2\Gamma(1+1/p_2)} \\
&= -\frac{1+p_1}{\alpha} - \frac{1+p_2}{1-\alpha}.
\end{aligned}
$$

ϕ_{22} is given by

$$
\begin{aligned}
E[\frac{\partial \ln f}{\partial p_1}]^2 &= \frac{1}{p_1^2}\Psi^2(1+1/p_1)E[L(y_t;\theta)]^{2p_1} + E[L(y_t;\theta)]^{2p_1}[\ln L(y_t;\theta)]^2 \\
&\quad -\frac{2}{p_1}\Psi(1+1/p_1)E\left\{[L(y_t;\theta)]^{2p_1}\ln L(y_t;\theta)\right\} \\
&= \frac{1}{p_1^2}\Psi^2(1+1/p_1)\frac{\alpha\Gamma(2+1/p_1)}{p_1\Gamma(1+1/p_1)} + \frac{\alpha\Gamma''(2+1/p_1)}{p_1^3\Gamma(1+1/p_1)} \\
&\quad -\frac{2}{p_1}\Psi(1+1/p_1)\frac{\alpha\Gamma'(2+1/p_1)}{p_1^2\Gamma(1+1/p_1)} \\
&= \frac{\alpha}{p_1^3}(1+\frac{1}{p_1})\Psi'(1+1/p_1); \text{ and}
\end{aligned}
$$

$$
\begin{aligned}
E[\frac{\partial^2 \ln f}{\partial p_1^2}] &= -\left[\frac{1}{p_1^3}\Psi'(1+1/p_1) + \frac{1}{p_1^2}\Psi^2(1+1/p_1)\right]E[L(y_t;\theta)]^{p_1} \\
&\quad +\frac{2}{p_1}\Psi(1+1/p_1)E\left\{[L(y_t;\theta)]^{p_1}\left[1-\ln L(y_t;\theta)\right]\ln L(y_t;\theta)\right\} \\
&= -\left[\frac{1}{p_1^3}\Psi'(1+1/p_1) + \frac{1}{p_1^2}\Psi^2(1+1/p_1)\right]\frac{\alpha\Gamma(1+1/p_1)}{p_1\Gamma(1+1/p_1)} \\
&\quad +\frac{2}{p_1}\Psi(1+1/p_1)\frac{\alpha\Gamma'(1+1/p_1)}{p_1^2\Gamma(1+1/p_1)} - \frac{\alpha\Gamma''(1+1/p_1)}{p_1^3\Gamma(1+1/p_1)} \\
&= -\frac{\alpha}{p_1^3}(1+\frac{1}{p_1})\Psi'(1+1/p_1).
\end{aligned}
$$

ϕ_{44} is given by

$$
\begin{aligned}
E[\frac{\partial \ln f}{\partial \mu}]^2 &= \left[\frac{\Gamma(1/p_1)}{\alpha\sigma}\right]^2 E[L(y_t;\theta)]^{2(p_1-1)} + \left[\frac{\Gamma(1/p_2)}{(1-\alpha)\sigma}\right]^2 E[R(y_t;\theta)]^{2(p_2-1)} \\
&= \left[\frac{\Gamma(1/p_1)}{\alpha\sigma}\right]^2\frac{\alpha\Gamma(2-1/p_1)}{p_1\Gamma(1+1/p_1)} + \left[\frac{\Gamma(1/p_2)}{(1-\alpha)\sigma}\right]^2\frac{(1-\alpha)\Gamma(2-1/p_2)}{p_2\Gamma(1+1/p_2)} \\
&= \frac{1}{\sigma^2}\left[\frac{\Gamma(1/p_1)\Gamma(2-1/p_1)}{\alpha} + \frac{\Gamma(1/p_2)\Gamma(2-1/p_2)}{1-\alpha}\right]; \text{ also}
\end{aligned}
$$

$$
\begin{aligned}
& E[\frac{\partial^2 \ln f}{\partial \mu^2}] = -\frac{\Gamma(1/p_1)}{\alpha\sigma}\frac{(p_1-1)\Gamma(1+1/p_1)}{\alpha\sigma}E[L(y_t;\theta)]^{p_1-2} \\
& -\frac{\Gamma(1/p_2)}{(1-\alpha)\sigma}\frac{(p_2-1)\Gamma(1+1/p_2)}{(1-\alpha)\sigma}E[R(y_t;\theta)]^{p_2-2} \\
= & -[\frac{\Gamma(1/p_1)}{\alpha\sigma}]^2\frac{p_1-1}{p_1}\frac{\alpha\Gamma(1-1/p_1)}{p_1\Gamma(1+1/p_1)} - [\frac{\Gamma(1/p_2)}{(1-\alpha)\sigma}]^2\frac{p_2-1}{p_2}\frac{(1-\alpha)\Gamma(1-1/p_2)}{p_2\Gamma(1+1/p_2)} \\
= & -\frac{1}{\sigma^2}\left[\frac{\Gamma(1/p_1)(1-1/p_1)\Gamma(1-1/p_1)}{\alpha} + \frac{\Gamma(1/p_2)(1-1/p_2)\Gamma(1-1/p_2)}{1-\alpha}\right] \\
= & -\frac{\pi}{\sigma^2}\left[\frac{1-1/p_1}{\alpha\sin(\pi/p_1)} + \frac{1-1/p_2}{(1-\alpha)\sin(\pi/p_2)}\right],
\end{aligned}
$$

where we used the formula $\Gamma(x)\Gamma(1-x) = \pi/\sin(\pi x)$ in the last equality.

ϕ_{55} is given by

$$
\begin{aligned}
E[\frac{\partial \ln f}{\partial \sigma}]^2 &= -\frac{1}{\sigma^2} + \left(\frac{p_1}{\sigma}\right)^2 E[L(y_t;\theta)]^{2p_1} + \left(\frac{p_2}{\sigma}\right)^2 E[R(y_t;\theta)]^{2p_2} \\
&= -\frac{1}{\sigma^2} + \left(\frac{p_1}{\sigma}\right)^2 \frac{\alpha\Gamma(2+1/p_1)}{p_1\Gamma(1+1/p_1)} + \left(\frac{p_2}{\sigma}\right)^2 \frac{(1-\alpha)\Gamma(2+1/p_2)}{p_2\Gamma(1+1/p_2)} \\
&= \frac{\alpha p_1 + (1-\alpha)p_2}{\sigma^2};\ \text{also,}
\end{aligned}
$$

$$
\begin{aligned}
E[\frac{\partial^2 \ln f}{\partial \sigma^2}] &= \frac{1}{\sigma^2} - \frac{p_1(1+p_1)}{\sigma^2}E[L(y_t;\theta)]^{p_1} - \frac{p_2(1+p_2)}{\sigma^2}E[R(y_t;\theta)]^{2p_2} \\
&= \frac{1}{\sigma^2} - \frac{p_1(1+p_1)}{\sigma^2}\frac{\alpha\Gamma(1+1/p_1)}{p_1\Gamma(1+1/p_1)} - \frac{p_2(1+p_2)}{\sigma^2}\frac{(1-\alpha)\Gamma(1+1/p_2)}{p_2\Gamma(1+1/p_2)} \\
&= -\frac{\alpha p_1 + (1-\alpha)p_2}{\sigma^2}.
\end{aligned}
$$

ϕ_{12} is given by

$$
\begin{aligned}
E[\frac{\partial \ln f}{\partial \alpha}\frac{\partial \ln f}{\partial p_1}] &= \frac{1}{\alpha}\Psi(1+1/p_1)E[L(y_t;\theta)]^{2p_1} - \frac{p_1}{\alpha}E[L(y_t;\theta)]^{2p_1}\ln L(y_t;\theta) \\
&= \frac{1}{\alpha}\Psi(1+1/p_1)\frac{\alpha\Gamma(2+1/p_1)}{p_1\Gamma(1+1/p_1)} - \frac{p_1}{\alpha}\frac{\alpha\Gamma'(2+1/p_1)}{p_1^2\Gamma(1+1/p_1)} = -1/p_1;
\end{aligned}
$$

also

$$
E[\frac{\partial^2 \ln f}{\partial \alpha \partial p_1}] = \frac{1}{\alpha}E[L(y_t;\theta)]^{p_1} - \frac{p_1}{\alpha}E\left[\frac{\partial \ln f}{\partial p_1}\right] = \frac{1}{\alpha}\frac{\alpha\Gamma(1+1/p_1)}{p_1\Gamma(1+1/p_1)} - 0 = 1/p_1.
$$

ϕ_{14} is given by

$$
\begin{aligned}
E[\frac{\partial \ln f}{\partial \alpha}\frac{\partial \ln f}{\partial \mu}] &= -\frac{p_1\Gamma(1/p_1)}{\alpha^2\sigma}E[L(y_t;\theta)]^{2p_1-1} - \frac{p_2\Gamma(1/p_2)}{(1-\alpha)^2\sigma}E[R(y_t;\theta)]^{2p_2-1} \\
&= -\frac{p_1\Gamma(1/p_1)}{\alpha^2\sigma}\frac{\alpha\Gamma(2)}{p_1\Gamma(1+1/p_1)} - \frac{p_2\Gamma(1/p_2)}{(1-\alpha)^2\sigma}\frac{(1-\alpha)\Gamma(2)}{p_2\Gamma(1+1/p_2)} \\
&= -\frac{1}{\sigma}\left(\frac{p_1}{\alpha} + \frac{p_2}{1-\alpha}\right);\ \text{also}
\end{aligned}
$$

$$\begin{aligned}
E[\frac{\partial^2 \ln f}{\partial\alpha\partial\mu}] &= \frac{p_1}{\alpha}\frac{\Gamma(1/p_1)}{\alpha\sigma}E[L(y_t;\theta)]^{p_1-1} + \frac{p_2}{1-\alpha}\frac{\Gamma(1/p_2)}{(1-\alpha)\sigma}E[R(y_t;\theta)]^{p_2-1} \\
&= \frac{p_1\Gamma(1/p_1)}{\alpha^2\sigma}\frac{\alpha\Gamma(1)}{p_1\Gamma(1+1/p_1)} + \frac{p_2\Gamma(1/p_2)}{(1-\alpha)^2\sigma}\frac{(1-\alpha)\Gamma(1)}{p_2\Gamma(1+1/p_2)} \\
&= \frac{1}{\sigma}\left(\frac{p_1}{\alpha} + \frac{p_2}{1-\alpha}\right).
\end{aligned}$$

ϕ_{15} is given by

$$\begin{aligned}
E[\frac{\partial \ln f}{\partial\alpha}\frac{\partial \ln f}{\partial\sigma}] &= -\frac{1}{\sigma}E[\frac{\partial \ln f}{\partial\alpha}] + \frac{p_1^2}{\alpha\sigma}E[L(y_t;\theta)]^{2p_1} - \frac{p_2^2}{(1-\alpha)\sigma}E[R(y_t;\theta)]^{2p_2} \\
&= 0 + \frac{p_1^2}{\alpha\sigma}\frac{\alpha\Gamma(2+1/p_1)}{p_1\Gamma(1+1/p_1)} - \frac{p_2^2}{(1-\alpha)\sigma}\frac{(1-\alpha)\Gamma(2+1/p_2)}{p_2\Gamma(1+1/p_2)} \\
&= \frac{1}{\sigma}(p_1-p_2);\ \text{also}
\end{aligned}$$

$$\begin{aligned}
E[\frac{\partial^2 \ln f}{\partial\alpha\partial\sigma}] &= -\frac{p_1^2}{\alpha\sigma}E[L(y_t;\theta)]^{p_1} + \frac{p_2^2}{(1-\alpha)\sigma}E[R(y_t;\theta)]^{p_2} \\
&= -\frac{p_1^2}{\alpha\sigma}\frac{\alpha\Gamma(1+1/p_1)}{p_1\Gamma(1+1/p_1)} - \frac{p_2^2}{(1-\alpha)\sigma}\frac{(1-\alpha)\Gamma(1+1/p_2)}{p_2\Gamma(1+1/p_2)} \\
&= -\frac{1}{\sigma}(p_1-p_2).
\end{aligned}$$

Note that $\frac{\partial^2 \ln f}{\partial p_1 \partial p_2} = 0$ and $\frac{\partial \ln f}{\partial p_1}\frac{\partial \ln f}{\partial p_2} = 0$ because of $1(y_t < \mu)1(y_t > \mu) = 0$. Then we have

$$\phi_{23} = E[\frac{\partial \ln f}{\partial p_1}\frac{\partial \ln f}{\partial p_2}] = -E[\frac{\partial^2 \ln f}{\partial p_1\partial p_2}] = 0.$$

ϕ_{24} is given by

$$\begin{aligned}
&E[\frac{\partial \ln f}{\partial p_1}\frac{\partial \ln f}{\partial\mu}] = -\frac{\Gamma(1/p_1)}{\alpha\sigma}\frac{\Psi(1+1/p_1)}{p_1}E[L(y_t;\theta)]^{2p_1-1} \\
&+\frac{\Gamma(1/p_1)}{\alpha\sigma}E\left\{[L(y_t;\theta)]^{2p_1-1}\ln L(y_t;\theta)\right\} \\
= &-\frac{\Gamma(1/p_1)}{\alpha\sigma}\frac{\Psi(1+1/p_1)}{p_1}\frac{\alpha\Gamma(2)}{p_1\Gamma(1+1/p_1)} + \frac{\Gamma(1/p_1)}{\alpha\sigma}\frac{\alpha\Gamma'(2)}{p_1^2\Gamma(1+1/p_1)} \\
= &\frac{1}{\sigma p_1}[\Psi(2) - \Psi(1+1/p_1)];\ \text{also}
\end{aligned}$$

$$\begin{aligned}
&E[\frac{\partial^2 \ln f}{\partial p_1\partial\mu}] = \frac{\Gamma(1/p_1)}{p_1\alpha\sigma}\left[\Psi(1+\frac{1}{p_1}) - 1\right]E[L(y_t;\theta)]^{p_1-1} \\
&-\frac{\Gamma(1/p_1)}{\alpha\sigma}E\left\{[L(y_t;\theta)]^{p_1-1}\ln L(y_t;\theta)\right\} \\
= &\frac{\Gamma(1/p_1)}{p_1\alpha\sigma}\left[\Psi(1+\frac{1}{p_1}) - 1\right]\frac{\alpha\Gamma(1)}{p_1\Gamma(1+1/p_1)} - \frac{\Gamma(1/p_1)}{\alpha\sigma}\frac{\alpha\Gamma'(1)}{p_1^2\Gamma(1+1/p_1)} \\
= &-\frac{1}{\sigma p_1}[\Psi(2) - \Psi(1+1/p_1)].
\end{aligned}$$

ϕ_{25} is given by

$$\begin{aligned}
E[\frac{\partial \ln f}{\partial p_1}\frac{\partial \ln f}{\partial \sigma}] &= -\frac{1}{\sigma}E[\frac{\partial \ln f}{\partial p_1}] + \frac{\Psi(1+1/p_1)}{\sigma}E[L(y_t;\theta)]^{2p_1} \\
&\quad -\frac{p_1}{\sigma}E\left\{[L(y_t;\theta)]^{2p_1}\ln L(y_t;\theta)\right\} \\
&= 0 + \frac{\Psi(1+1/p_1)}{\sigma}\frac{\alpha\Gamma(2+1/p_1)}{p_1\Gamma(1+1/p_1)} - \frac{p_1}{\sigma}\frac{\alpha\Gamma'(2+1/p_1)}{p_1^2\Gamma(1+1/p_1)} \\
&= -\frac{\alpha}{\sigma p_1};\ \text{also}
\end{aligned}$$

$$E[\frac{\partial^2 \ln f}{\partial p_1 \partial \sigma}] = \frac{1}{\sigma}E[L(y_t;\theta)]^{p_1} - \frac{p_1}{\sigma}E[\frac{\partial \ln f}{\partial p_1}] = \frac{1}{\sigma}\frac{\alpha\Gamma(1+1/p_1)}{p_1\Gamma(1+1/p_1)} - 0 = \frac{\alpha}{\sigma p_1}.$$

ϕ_{45} is given by

$$\begin{aligned}
E[\frac{\partial \ln f}{\partial \mu}\frac{\partial \ln f}{\partial \sigma}] &= -\frac{p_1\Gamma(1/p_1)}{\alpha\sigma^2}E[L(y_t;\theta)]^{2p_1-1} + \frac{p_2\Gamma(1/p_2)}{(1-\alpha)\sigma^2}E[R(y_t;\theta)]^{2p_2-1} \\
&= -\frac{p_1\Gamma(1/p_1)}{\alpha\sigma^2}\frac{\alpha\Gamma(2)}{p_1\Gamma(1+1/p_1)} + \frac{p_2\Gamma(1/p_2)}{(1-\alpha)\sigma^2}\frac{(1-\alpha)\Gamma(2)}{p_2\Gamma(1+1/p_2)} \\
&= \frac{1}{\sigma^2}(-p_1+p_2);\ \text{also}
\end{aligned}$$

$$\begin{aligned}
E[\frac{\partial^2 \ln f}{\partial \mu \partial \sigma}] &= \frac{p_1}{\sigma}\frac{\Gamma(1/p_1)}{\alpha\sigma}E[L(y_t;\theta)]^{p_1-1} - \frac{p_2}{\sigma}\frac{\Gamma(1/p_2)}{(1-\alpha)\sigma}E[R(y_t;\theta)]^{p_2-1} \\
&= \frac{p_1\Gamma(1/p_1)}{\alpha\sigma^2}\frac{\alpha\Gamma(1)}{p_1\Gamma(1+1/p_1)} - \frac{p_2\Gamma(1/p_2)}{(1-\alpha)\sigma^2}\frac{(1-\alpha)\Gamma(1)}{p_2\Gamma(1+1/p_2)} \\
&= \frac{1}{\sigma^2}(p_1-p_2).
\end{aligned}$$

By the symmetry of p_1 and p_2, we similarly have

$$\begin{aligned}
\phi_{33} &= E[\frac{\partial \ln f}{\partial p_2}]^2 = -E[\frac{\partial^2 \ln f}{\partial p_2^2}] = \frac{1-\alpha}{p_2^3}(1+\frac{1}{p_2})\Psi'(1+1/p_2); \\
\phi_{13} &= E[\frac{\partial \ln f}{\partial p_2}\frac{\partial \ln f}{\partial \alpha}] = -E[\frac{\partial^2 \ln f}{\partial p_2 \partial \alpha}] = 1/p_2; \\
\phi_{34} &= E[\frac{\partial \ln f}{\partial p_2}\frac{\partial \ln f}{\partial \mu}] = -E[\frac{\partial^2 \ln f}{\partial p_2 \partial \mu}] = -\frac{1}{\sigma p_2}[\Psi(2)-\Psi(1+1/p_2)]; \\
\phi_{35} &= E[\frac{\partial \ln f}{\partial p_2}\frac{\partial \ln f}{\partial \sigma}] = -E[\frac{\partial^2 \ln f}{\partial p_2 \partial \sigma}] = -\frac{1-\alpha}{\sigma p_2}.\blacksquare
\end{aligned}$$

Corollary 9 *For the SEPD ($p_1 = p_2 = p$), the component $\frac{\partial \ln f}{\partial p}$ of its score vector is the sum of $\frac{\partial \ln f}{\partial p_1}$ and $\frac{\partial \ln f}{\partial p_2}$ of the AEPD score (38). Thus, by incorporating the terms of ϕ_{ij} involving p_1 and p_2, i.e., $\phi_{12}+\phi_{13}$, $\phi_{22}+\phi_{33}$, $\phi_{24}+\phi_{34}$*

and $\phi_{25} + \phi_{35}$, we can obtain the information matrix for the MLE of the SEPD parameters (α, p, μ, σ) as follows:

$$I(\theta_0) = -H(\theta_0) = \begin{pmatrix} \frac{p+1}{\alpha(1-\alpha)} & 0 & -\frac{p}{\sigma\alpha(1-\alpha)} & 0 \\ 0 & \frac{p+1}{p^4}\Psi'(\frac{p+1}{p}) & 0 & -\frac{1}{\sigma p} \\ -\frac{p}{\sigma\alpha(1-\alpha)} & 0 & \frac{\Gamma(1/p)\Gamma(2-1/p)}{\sigma^2\alpha(1-\alpha)} & 0 \\ 0 & -\frac{1}{\sigma p} & 0 & \frac{p}{\sigma^2} \end{pmatrix}.$$

The asymptotic covariance matrix is given by the inverse of the information matrix,

$$\begin{pmatrix} \frac{\xi\alpha(1-\alpha)}{\xi(p+1)-p^2} & 0 & \frac{p\sigma\alpha(1-\alpha)}{\xi(p+1)-p^2} & 0 \\ 0 & \frac{p^4}{\eta(p+1)-p} & 0 & \frac{\sigma p^2}{\eta(p+1)-p} \\ \frac{p\sigma\alpha(1-\alpha)}{\xi(p+1)-p^2} & 0 & \frac{\sigma^2(p+1)\alpha(1-\alpha)}{\xi(p+1)-p^2} & 0 \\ 0 & \frac{\sigma p^2}{\eta(p+1)-p} & 0 & \frac{\eta}{p}\frac{\sigma^2(p+1)}{\eta(p+1)-p} \end{pmatrix},$$

where $\xi = \Gamma(1/p)\Gamma(2-1/p)$ and $\eta = \Psi'(1+1/p)$. It is easy to show that $\xi(p+1) - p^2 > 0$ and $\eta(p+1) - p > 0$ for $p \geq 1$.

Remark 10 *Although the information matrix equality may hold for all $p_{01} > 0$ and $p_{02} > 0$ except for points $1/n$ (n=2, 3, 4,...) of p_{01} and p_{02}, we restrict p_{01} and p_{02} to satisfy $p_{01} > 1/2$ and $p_{02} > 1/2$. The reason is that (1) $I(\theta_0)$ and $H(\theta_0)$ are undefined and thus discontinuous at some points of $p_{01} \leq 1/2$ and (or) $p_{02} \leq 1/2$, i.e., those points $1/n$ (n=2, 3, 4,...); and (2) the information matrix equality has no significance for p_{01} and p_{02} in intervals $(\frac{1}{2n+1}, \frac{1}{2n})$, $n = 1, 2, 3, ...$, because $E[\frac{\partial \ln f}{\partial \mu}]^2$ is negative when both p_{01} and p_{02} are in these intervals (see the expression of $E[\frac{\partial \ln f}{\partial \mu}]^2$ above and properties of the gamma function). Here we need to point out that the existence of $E[\partial^2 \ln f / \partial \mu^2]$ at $p_{01} = 1$ and (or) $p_{02} = 1$ is due to the fact that $x\Gamma(x) \to 1$ or $\sin(x)/x \to 1$ as $x \to 0$.*

9 Appendix C

Appendix C is devoted to establishing consistency and asymptotic normality of the MLE of all parameters of the AEPD. The results in the following preliminary lemma are used in the proof of Proposition 5.

Lemma 11 *(a). For any $\varepsilon > 0$ there exists a positive constant M_0, that may depend on ε, such that*

$$|\ln x| \leq M_0 \left(1 + x^{-\varepsilon} + x^{\varepsilon}\right), \text{ for any } x > 0. \tag{43}$$

(b). For any (μ^*, q^*) such that $|q^* - q| \leq d$ and $|\mu^* - \mu| \leq d$, the following

inequalities hold:

$$\begin{aligned}
(\mu^* - y)^{q^*} &\leq 2 + (\mu + d - y)^{q+d} + (\mu - d - y)^{q-d}, \text{ if } y < \mu - d; && (44)\\
(\mu^* - y)^{q^*} &\leq 1 + (\mu + d - y)^{q+d}, \quad \text{if } q^* > 0,\ y < \mu - d; && (45)\\
(y - \mu^*)^{q^*} &\leq 2 + (y - \mu + d)^{q+d} + (y - \mu - d)^{q-d}, \text{ if } y > \mu + d; && (46)\\
(y - \mu^*)^{q^*} &\leq 1 + (y - \mu + d)^{q+d}, \quad \text{if } q^* > 0,\ y > \mu + d. && (47)
\end{aligned}$$

(c). Suppose that Y *is an AEPD r.v. with density* $f(y \mid \theta_0)$ *defined in (6), where* $\theta_0 = (\alpha_0, p_{01}, p_{02}, \mu_0, \sigma_0)$*. Then, for any* $\mu \in R$ *and* $r > -1$*, the following inequality holds:*

$$E\left|Y - \mu\right|^r \leq M_1(\mu, r)\Gamma(\frac{1+r}{p_{01}}) + M_2(\mu, r)\Gamma(\frac{1+r}{p_{02}}), \tag{48}$$

where $M_1(\cdot,\cdot)$ *and* $M_2(\cdot,\cdot)$ *are two positive continuous functions.*

Proof. Part (a) is immediate from the facts that for any $\varepsilon > 0$, $x^\varepsilon \left|\ln x\right| \to 0$ as $x \to 0^+$, and $\left|\ln x\right| / x^\varepsilon \to 0$, as $x \to +\infty$. Part (b) is obtained by first considering the cases of $q^* > 0$ and $q^* < 0$ and then dealing with the cases of $|\mu \pm d - y| > 1$ and $|\mu \pm d - y| < 1$. Part (c) can be proved by using the c_r-inequality (see Loève, 1977, p157), $|y - \mu_0|^p \geq 2^{1-p}|y - \mu|^p - |\mu_0 - \mu|^p$ for $p \geq 1$, and then using a change of variable. In fact, for $r > -1$,

$$\begin{aligned}
E\left|Y - \mu\right|^r &= \int_{-\infty}^{\mu_0} |y - \mu|^r \frac{1}{\sigma_0} \exp\left[-C_1(\theta_0)\left|y - \mu_0\right|^{p_{01}}\right] dy \\
&\quad + \int_{\mu_0}^{+\infty} |y - \mu|^r \frac{1}{\sigma_0} \exp\left[-C_2(\theta_0)\left|y - \mu_0\right|^{p_{02}}\right] dy \\
&\leq D_1(\mu) \int_{-\infty}^{\mu} |y - \mu|^r \exp\left[-2^{1-p_{01}} C_1(\theta_0)\left|y - \mu\right|^{p_{01}}\right] dy \\
&\quad + D_2(\mu) \int_{\mu}^{+\infty} |y - \mu|^r \exp\left[-2^{1-p_{02}} C_2(\theta_0)\left|y - \mu\right|^{p_{02}}\right] dy \\
&= \sum_{i=1}^{2} M_i(\mu, r) \int_0^{+\infty} x^{(1+r)/p_{0i} - 1} e^{-x} dx \\
&= \sum_{i=1}^{2} M_i(\mu, r) \Gamma(\frac{1+r}{p_{0i}}),
\end{aligned}$$

where

$$\begin{aligned}
C_1(\theta_0) &\equiv \left[\Gamma(1 + 1/p_{01}) / \left(\alpha_0 \sigma_0\right)\right]^{p_{01}}, \\
C_2(\theta_0) &\equiv \left[\Gamma(1 + 1/p_{02}) / \left((1 - \alpha_0)\sigma_0\right)\right]^{p_{02}}, \\
D_i(\mu) &\equiv \frac{2}{\sigma_0} \exp\left\{C_i(\theta_0)\left|\mu - \mu_0\right|^{p_{0i}}\right\}, \\
M_i(\mu, r) &\equiv \frac{D_i(\mu)}{p_{0i}} \left[2^{1-p_{0i}} C_i(\theta_0)\right]^{-(r+1)/p_{0i}}, \quad i = 1, 2. \blacksquare
\end{aligned}$$

Proof of Proposition 4 (consistency).

The consistency of the MLE $\widehat{\theta}_T$ can be shown by verifying the conditions of Theorem 2.5 in Newey and McFadden (1994, p.2131), which holds under conditions that are primitive and also quite weak. Condition (ii) of Theorem 2.5, compactness of the parameter set, is ensured by considering a compact parameter set Θ. Condition (iii) of Theorem 2.5 requires that the log-likelihood $\ln f(y \mid \theta)$ be continuous at each $\theta \in \Theta$ with probability one. This condition holds by inspection. We only need to check the identification condition and dominance condition (corresponding to conditions (i) and (iv) of Theorem 2.5 respectively).

The identification condition says that if $\theta \neq \theta_0$ then $f(y \mid \theta) \neq f(y \mid \theta_0)$, that is, $\Pr\{f(y \mid \theta) \neq f(y \mid \theta_0)\} > 0$. It is sufficient to show that for any given $\theta \neq \theta_0$ and $\theta \in \Theta$, there exists a set of positive probability, $S(\theta)$, such that

$$\ln f(y \mid \theta) \neq \ln f(y \mid \theta_0), \text{ a.e. } \forall\, y \in S(\theta). \tag{49}$$

The proof will use the fact that any AEPD random variable Y has positive probability on any interval. If $\mu \neq \mu_0$, says, $\mu > \mu_0$, then in interval $(\mu_0, \mu]$ the log-density function $\ln f(y \mid \theta)$ is strictly increasing but $\ln f(y \mid \theta_0)$ decreases strictly, so (49) always holds on $(\mu_0, \mu]$. Now suppose $\mu = \mu_0$. We shall show that (49) is true almost everywhere in $(-\infty, \mu_0]$ (or $(\mu_{0,} + \infty)$) if $p_1 \neq p_{01}$ (or $p_2 \neq p_{02}$). Suppose $p_2 \neq p_{02}$. Then, for $y \in (\mu_0, +\infty)$, $\ln f(y \mid \theta) = -\ln \sigma - C_2(\theta)(y - \mu_0)^{p_2}$ (since $\mu = \mu_0$) and $\ln f(y \mid \theta_0) = -\ln \sigma_0 - C_2(\theta_0)(y - \mu_0)^{p_{02}}$, where $C_2(\theta) = (\Gamma(1 + 1/p_2)/((1 - \alpha)\sigma))^{p_2}$. Since both the log-density functions on $(\mu_0, +\infty)$ are power functions, they intersect at no more than two points on that interval, implying that (49) holds for $S(\theta) = (\mu_0, +\infty)$. Similarly, under the assumptions of $\mu = \mu_0$, $p_1 = p_{01}$ and $p_2 = p_{02}$, it is easy to show that (49) holds if $\alpha \neq \alpha_0$ or $\sigma \neq \sigma_0$ (see Newey and McFadden, p.2126).

The dominance condition of Theorem 2.5, $E[\sup_{\theta \in \Theta} |\ln f(Y \mid \theta)|] < \infty$, can be verified here by the compactness of parameter set Θ and the boundedness of the $\overline{p}$th order absolute moment of a standard AEPD r.v., where $\overline{p}$ is the supremum of p_1 and p_2 in Θ. Since the parameter set Θ is assumed to be compact, so that any continuous function of θ is bounded on Θ, by using the c_r-inequality (see Loève, 1977, p.157), i.e., $|a + b|^r \leq c_r |a|^r + c_r |b|^r$, where $c_r = 1$ or 2^{r-1} according as $0 < r \leq 1$ or $r \geq 1$, we have $|\ln f(Y \mid \theta)| \leq K_1 + K_2 |X|^{\overline{p}}$ for all $\theta \in \Theta$, where K_1 and K_2 are certain positive constants and $X = \sigma_0(Y - \mu_0)$, a standard AEPD r.v. with parameters $(\alpha_0, p_{01}, p_{02})$. Thus, the dominance condition will be satisfied as long as $E[|X|^{\overline{p}}] < \infty$, which is shown to hold in (15).■

Proof of Proposition 5 (asymptotic normality).

The proof of the asymptotic normality result proceeds by verifying the conditions of Theorem 3 as well as its corollary in Huber (1967). Following the notation of Huber (1967), let $\psi(y, \theta) = \frac{\partial \ln f(y,\theta)}{\partial \theta}$, the score vector, and set

$$\lambda(\theta) = E\psi(y, \theta), \quad u(y, \theta, d) \equiv \sup_{\theta^* \in D^*} |\psi(y, \theta^*) - \psi(y, \theta)|, \tag{50}$$

where $D^* \equiv \{\theta^* \mid |\theta^* - \theta| \leq d\}$ and all expectations are always taken with respect to the true underlying distribution $f(y;\theta_0)$ with $\theta_0 = (\alpha_0, p_{01}, p_{02}, \mu_0, \sigma_0)$. Similar to Example 1 of Huber (1967), the condition N-1 (i.e., for each fixed θ, $\psi(y,\theta)$ is measurable and separable: see Assumption (A-1) of Huber (1967)) is immediate; both conditions (N-2) and (N-4), i.e., $\lambda(\theta_0) = 0$ and $E[|\psi(y,\theta_0)|^2] < \infty$, hold immediately from (41) and the fact that ϕ_{ii} in (42) are finite. By the definition of the MLE $\widehat{\theta}$, we have $\sum_{t=1}^T \psi(y_t, \widehat{\theta}) = 0$, implying that Equation (27) of Huber (1967) holds. Since consistency has been proved, the remaining condition of Huber's (1967) Theorem 3 is the condition (N-3): there are strictly positive numbers a, b, c, d_0 such that

$$\begin{aligned} |\lambda(\theta)| &\geq a\,|\theta - \theta_0|\,, \text{ for } |\theta - \theta_0| \leq d_0, && (51)\\ Eu(y,\theta,d) &\leq bd, \text{ for } |\theta - \theta_0| + d \leq d_0,\, d \geq 0, && (52)\\ E[u(y,\theta,d)^2] &\leq cd, \text{ for } |\theta - \theta_0| + d \leq d_0,\, d \geq 0, && (53)\end{aligned}$$

where $|\theta|$ denotes any norm equivalent to the Euclidean norm.

Now we check the condition (52). Separate the location parameter from the other parameters, $\tau = (\alpha, p_1, p_2, \sigma)$, i.e. $\theta = (\tau, \mu)$ and $\theta^* = (\tau^*, \mu^*)$. Then

$$u(y,\theta,d) \leq \sup_{\theta^* \in D^*} |\psi(y,\tau^*,\mu^*) - \psi(y,\tau^*,\mu)| + \sup_{|\tau^*-\tau| \leq d} |\psi(y,\tau^*,\mu) - \psi(y,\tau,\mu)|\,. \tag{54}$$

The condition (52) is easily verified for the second part in (54), because the location μ is fixed and $\psi(y,\tau,\mu)$ as a function of τ is smooth enough. For the first part in (54), note from (38) that each element of $\psi(y,\tau,\mu)$ can be expressed as the following form:

$$\begin{aligned} C(\tau) + [C_{11}(\tau)\,|\mu - y|^{q_1} + C_{12}(\tau)\,|\mu - y|^{q_1} \ln|\mu - y|]\,1(y &< \mu)\\ + [C_{21}(\tau)\,|y - \mu|^{q_2} + C_{22}(\tau)\,|y - \mu|^{q_2} \ln|y - \mu|]\,1(y &> \mu), \end{aligned} \tag{55}$$

where $(q_1, q_2) = (p_1, p_2)$ or $(q_1, q_2) = (p_1 - 1, p_2 - 1)$, $C(\cdot)$ and $C_{ij}(\cdot)$ are continuous functions of $\tau = (\alpha, p_1, p_2, \sigma)$, implying that they are bounded on the compact set Θ. So, we need only to prove that

$$E\left[\sup_{\theta^* \in D^*} \left|(\mu^* - y)^{q_1^*} 1(y < \mu^*) - (\mu - y)^{q_1^*} 1(y < \mu)\right|\right] \leq bd \tag{56}$$

and

$$E\left[\sup_{\theta^* \in D^*} \left|(\mu^* - y)^{q_1^*} \ln|\mu^* - y|\,1(y < \mu^*) - (\mu - y)^{q_1^*} \ln|\mu - y|\,1(y < \mu)\right|\right] \leq bd. \tag{57}$$

Here we show only the condition (57); the condition (56) can be verified similarly; the counterparts involved with "$1(y > \mu)$" can also be shown in the same way. In fact, denoting by $\underline{p}$ the infimum of the components p_i of $\theta \in \Theta$, by the assumption $p_i > 1$ $(i = 1, 2)$, we have $\underline{p} > 1$ and $q_1^* \geq \underline{p} - 1 \equiv \underline{q} > 0$. Taking

$d_0 < \min\{\underline{q}/2, \frac{1}{3}\}$ and noting that $|x^{\underline{q}} \ln x|$ is bounded in $(0,1)$, the condition (57) reduces to

$$E\left[\sup_{\theta^*\in D^*}\left|(\mu^*-y)^{q_1^*}\ln(\mu^*-y)-(\mu-y)^{q_1^*}\ln(\mu-y)\right|\right]1(y<\mu-2d)\le bd. \tag{58}$$

By using the mean-value theorem, (43) and (44), for any (μ^*, q_1^*) such that $|q_1^*-q_1|\le d$ and $|\mu^*-\mu|\le d$, we can show for $y<\mu-2d$ that

$$\begin{aligned}
&\left|(\mu^*-y)^{q_1^*}\ln(\mu^*-y)-(\mu-y)^{q_1^*}\ln(\mu-y)\right| \\
= &\left|(\widetilde{\mu}-y)^{q_1^*-1}\left\{q_1^*\ln(\widetilde{\mu}-y)+1\right\}\right||\mu^*-\mu| \\
\le & d\left[(\widetilde{\mu}-y)^{q_1^*-1}\left|\ln(\widetilde{\mu}-y)\right|+(\widetilde{\mu}-y)^{q_1^*-1}\right] \\
\le & dM_0(\varepsilon)\left[(\widetilde{\mu}-y)^{q_1^*-1}+(\widetilde{\mu}-y)^{q_1^*-1-\varepsilon}+(\widetilde{\mu}-y)^{q_1^*-1+\varepsilon}\right] \\
\le & dM_0\sum_{i=1}^{3}\left[2+(\mu+d-y)^{q_1+d-\delta_i}+(\mu-d-y)^{q_1-d-\delta_i}\right],
\end{aligned} \tag{59}$$

where $\widetilde{\mu}$ is a real number between μ and μ^*, $\delta_1=1$, $\delta_2=1+\varepsilon$ and $\delta_3=1-\varepsilon$. Note that $q_1\pm d-\delta_i>-1$ as long as $\varepsilon<\underline{q}/2$, say $\varepsilon=\underline{q}/4$, because $d\le d_0$ and $q_1\ge\underline{q}$. Then (58) holds immediately from (48) and the assumption of compactness of the parameter space Θ.

To verify the condition (53), it is sufficient to show that

$$E\left[\sup_{\theta^*\in D^*}\left|(\mu^*-y)^{q_1^*}\ln(\mu^*-y)-(\mu-y)^{q_1^*}\ln(\mu-y)\right|\right]^2 1(y<\mu-2d)\le cd. \tag{60}$$

In fact, for any (μ^*, q_1^*) such that $|q_1^*-q_1|\le d$ and $|\mu^*-\mu|\le d$, we have

$$\begin{aligned}
&\left|(\mu^*-y)^{q_1^*}\ln(\mu^*-y)-(\mu-y)^{q_1^*}\ln(\mu-y)\right| \\
\le & M_0(\varepsilon)\sum_{i=1}^{3}\left[1+(\mu+d-y)^{q_1+d-1+\delta_i}\right]1(y<\mu-2d),
\end{aligned} \tag{61}$$

where (45) is used, δ_i are defined in (59), and $q_1+d-1+\delta_i>0$ when $\varepsilon<\underline{q}$. Combining (61) with (59) and using the c_r-inequality (see Loève, 1977, p157) yields

$$\begin{aligned}
&\left[\sup_{\theta^*\in D^*}\left|(\mu^*-y)^{q_1^*}\ln(\mu^*-y)-(\mu-y)^{q_1^*}\ln(\mu-y)\right|\right]^2 1(y<\mu-2d) \\
\le & dK_0\left\{1+\sum_{i=1}^{K}\left[(\mu+d-y)^{\xi_i}+(\mu-d-y)^{\eta_i}\right]1(y<\mu-2d)\right\},
\end{aligned}$$

where K_0 is a positive constant that may depends on ε, K is a positive integer less than 28, ξ_i and η_i are some real numbers greater than -1 when the positive constant ε is small enough, say $\varepsilon = \underline{q}/4$. Thus, the condition (60) follows from (48) and the assumption of compactness of the parameter space Θ.

A sufficient condition for (51) to be true is that $\lambda(\theta)$ has continuous (partial) derivatives in some neighborhood $O(\theta_0, d_0)$ of θ_0, because (51) can be obtained by using this condition and the fact that the Hessian $H(\theta_0)$ is negative definite. Here we show only that $\partial\lambda_4(\theta)/\partial\mu$ is continuous, where $\lambda_4(\theta) = E\left[\partial \ln f(y,\theta)/\partial\mu\right]$; the continuity of other partial derivatives will be easily proved by using Lemma 3.6 of Newey and McFadden (1994, p 2152), the c_r-inequality (Loève, 1977, p.157) and results from (43) to (48). Note that

$$\begin{aligned}
\lambda_4(\theta) &= A_1(\tau)E\left|\mu - y\right|^{p_1-1} 1(y<\mu) + A_2(\tau)E\left|y-\mu\right|^{p_2-1} 1(y>\mu) \\
&= A_1(\tau)\int_0^{+\infty} x^{p_1-1} f(\mu - x;\theta_0)dx + A_2(\tau)\int_0^{+\infty} x^{p_2-1} f(x+\mu;\theta_0)dx \\
&= \int_0^{+\infty} a(x;\theta)dx,
\end{aligned}$$

where A_1 and A_2 are some continuously differentiable functions of $\tau = (\alpha, p_1, p_2, \sigma)$, implying that they are bounded in the compact parameter space Θ; $f(y;\theta_0)$ is the true AEPD density and

$$a(x,\theta) = A_1(\tau)x^{p_1-1}f(\mu - x;\theta_0) + A_2(\tau)x^{p_2-1}f(x+\mu;\theta_0).$$

Let $d_0 > 0$ be small enough that $O(\theta_0, d_0) \equiv \{\theta : |\theta - \theta_0| < d_0\} \subset \Theta$. Then, obviously, $a(x,\theta)$ is continuously differentiable in the neighborhood $O(\theta_0, d_0)$ of θ_0, a.s.; and by the c_r-inequality (see Loève, 1977, p 157) and compactness of the parameter space Θ,

$$\sup_{|\theta-\theta_0|<d_0} \left|\frac{\partial a(x,\theta)}{\partial\mu}\right| \leq \begin{cases} B_0 x^{\underline{p}-1}, & 0 \leq x \leq 1 \\ B_0 x^{2(\overline{p}-1)} \exp(-B_1 x), & x > 1 \end{cases}$$

where B_0 and B_1 are some positive constants that do not depend on θ, $\underline{p} > 1$ and $\overline{p} > 1$ are, respectively, the infimum and supremum of the components p_i of $\theta \in \Theta$. From Lemma 3.6 of Newey and McFadden (1994, p. 2152) it follows that $\lambda_4(\theta)$ is continuously differentiable with respect to μ in the neighborhood $O(\theta_0, d_0)$ of θ_0.■

10 Appendix D

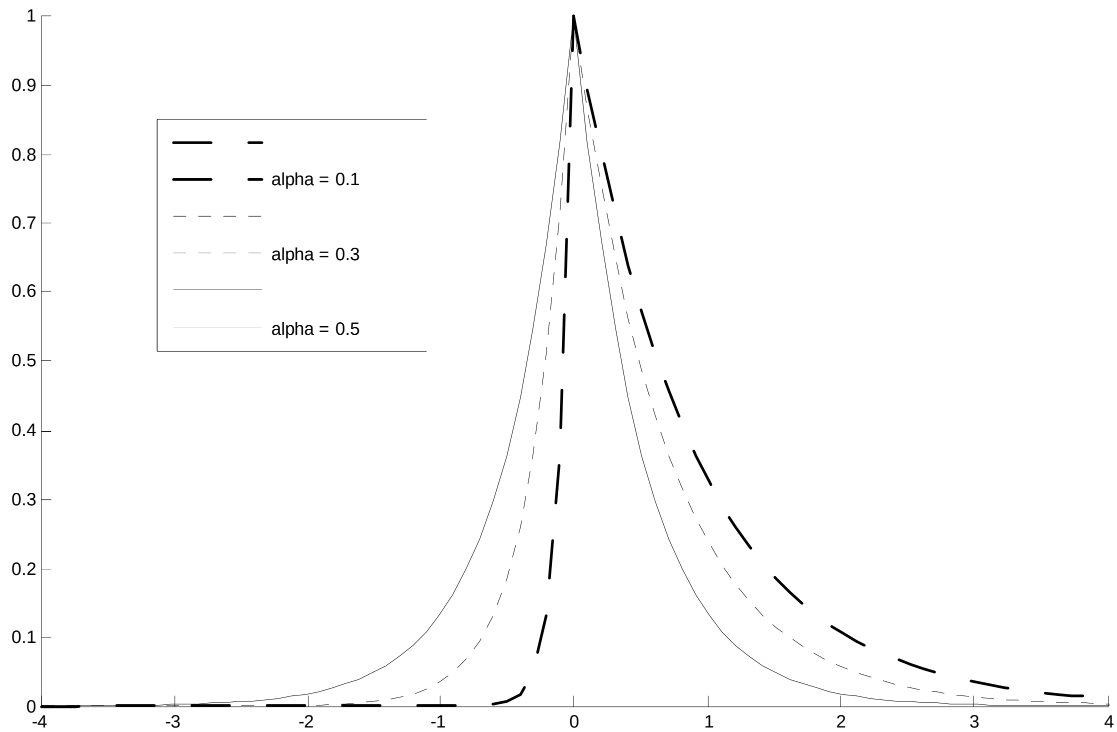

Figure 1: Plot 1. The AEPD densities for $p_1 = p_2 = 1$ and varying α.

Figure 1---plot 2: The AEPD densities for alpha = 0.5, p1 = 1

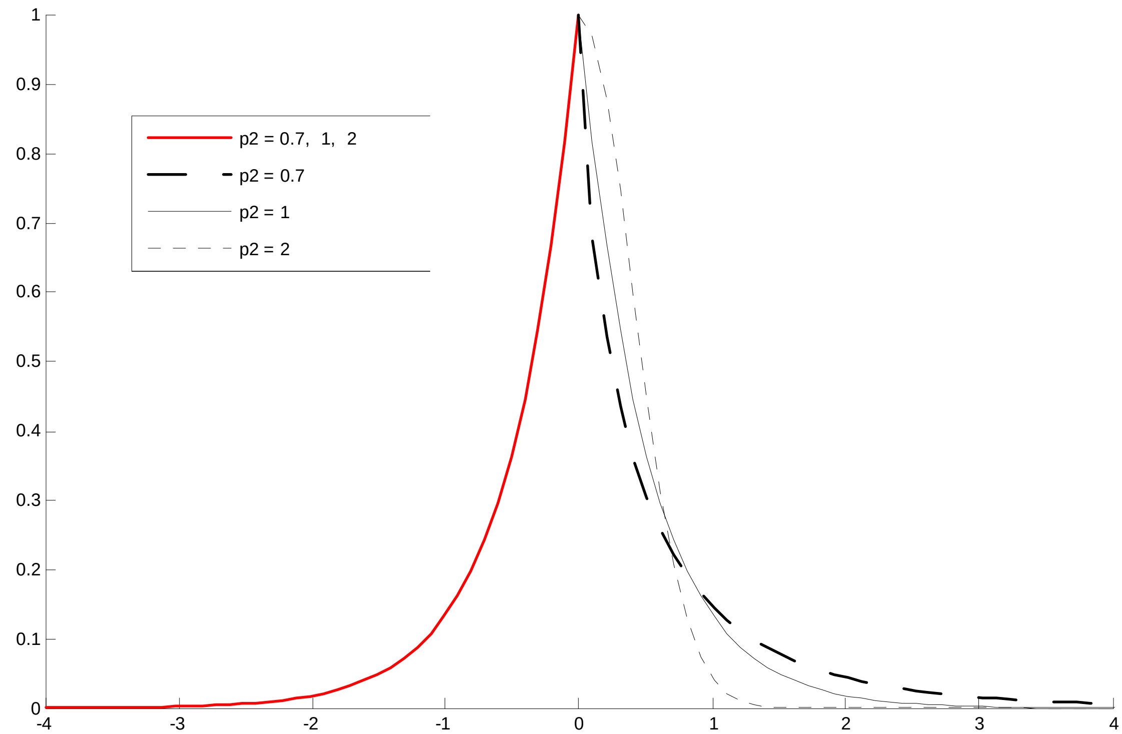

Figure 1---plot 3: The AEPD densities for p1 = 1

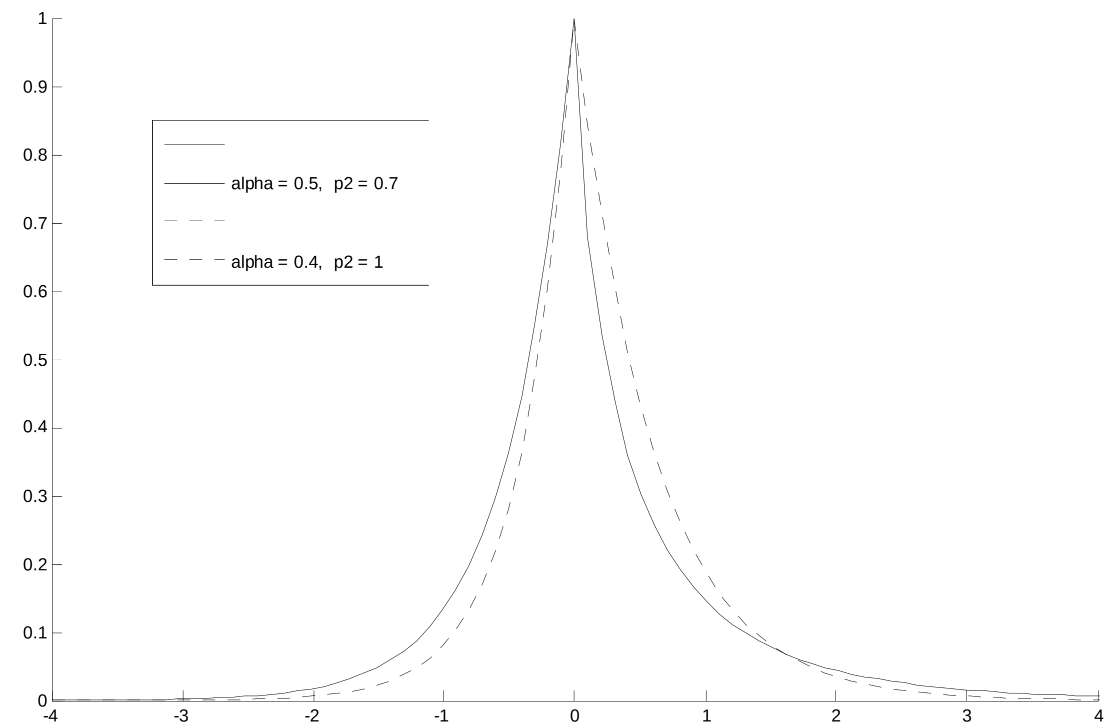

Figure 2: Table. Simulation results for the MLE of the AEPD: averages of estimated parameters (Mean of MLEs) and ratios of simulated standard deviations to theoretical (STD ratio) for different combinations of parameters.

			P2=1						*P2=1.5*			
	alpha=0.3	***p1=0.7***	***p2=1***	**sigma=1**	**mu=0**		**alpha=0.3**	***p1=0.7***	***p2=1.5***	**sigma=1**	**mu=0**	
	0.3042	0.7125	1.0049	1.0008	0.0056		0.3041	0.7100	1.5090	0.9966	0.0053	**T=500**
	0.3021	0.7050	1.0004	0.9986	0.0033		0.3030	0.7068	1.5027	0.9985	0.0038	**T=1000**
Mean of	0.3013	0.7032	1.0008	0.9999	0.0017		0.3013	0.7032	1.5012	0.9992	0.0016	**T=2000**
MLEs	0.3005	0.7013	0.9999	0.9991	0.0007		0.3004	0.7012	1.5006	0.9995	0.0003	**T=4000**
	0.3003	0.7004	0.9993	0.9991	0.0004		0.3002	0.7007	1.4998	0.9995	0.0003	**T=8000**
	alpha=0.3	***p1=0.7***	***p2=1***	**sigma=1**	**mu=0**		**alpha=0.3**	***p1=0.7***	***p2=1.5***	**sigma=1**	**mu=0**	
	1.172	1.137	1.066	1.038	1.437		1.140	1.107	1.086	1.026	1.343	**T=500**
	1.342	1.126	1.115	1.016	1.612		1.331	1.172	1.128	1.032	1.557	**T=1000**
STD Ratio	1.334	1.120	1.077	1.011	1.565		1.318	1.124	1.138	1.031	1.545	**T=2000**
	1.280	1.104	1.069	1.022	1.479		1.230	1.071	1.126	1.039	1.418	**T=4000**
	1.224	1.097	1.048	1.002	1.375		1.196	1.065	1.110	1.041	1.350	**T=8000**
	alpha=0.3	***p1=1***	***p2=1***	**sigma=1**	**mu=0**		**alpha=0.3**	***p1=1***	***p2=1.5***	**sigma=1**	**mu=0**	
	0.3045	1.0272	1.0065	1.0052	0.0062		0.3082	1.0318	1.5104	1.0012	0.0095	**T=500**
	0.3019	1.0141	1.0050	1.0045	0.0024		0.3023	1.0121	1.5104	1.0019	0.0022	**T=1000**
Mean of	0.3005	1.0058	1.0036	1.0024	0.0009		0.3008	1.0048	1.5047	1.0007	0.0007	**T=2000**
MLEs	0.2999	1.0019	1.0017	1.0007	0.0002		0.3005	1.0026	1.5009	1.0000	0.0004	**T=4000**
	0.2999	1.0008	1.0010	1.0004	0.0000		0.3004	1.0021	1.5005	1.0000	0.0004	**T=8000**
	alpha=0.3	***p1=1***	***p2=1***	**sigma=1**	**mu=0**		**alpha=0.3**	***p1=1***	***p2=1.5***	**sigma=1**	**mu=0**	
	1.178	1.185	1.092	1.032	1.264		1.175	1.184	1.094	1.021	1.215	**T=500**
	1.250	1.186	1.076	1.019	1.321		1.201	1.167	1.087	1.023	1.250	**T=1000**
STD Ratio	1.183	1.124	1.060	1.020	1.243		1.165	1.110	1.059	0.987	1.196	**T=2000**
	1.114	1.072	1.042	0.999	1.146		1.077	1.026	1.028	0.982	1.101	**T=4000**
	1.099	1.035	1.034	0.992	1.127		1.043	1.015	1.013	0.982	1.054	**T=8000**

	alpha=0.3	***p1=1.5***	***p2=1***	**sigma=1**	**mu=0**		**alpha=0.3**	***p1=1.5***	***p2=1.5***	**sigma=1**	**mu=0**	
	0.3081	1.5894	1.0014	1.0079	0.0097		0.3126	1.5894	1.5049	1.0064	0.0142	**T=500**
	0.3060	1.5568	0.9989	1.0054	0.0074		0.3067	1.5499	1.5052	1.0047	0.0070	**T=1000**
Mean of	0.3030	1.5293	1.0001	1.0027	0.0039		0.3029	1.5233	1.5035	1.0021	0.0032	**T=2000**
MLEs	0.3015	1.5158	1.0004	1.0020	0.0019		0.3012	1.5108	1.5028	1.0017	0.0013	**T=4000**
	0.3010	1.5092	1.0002	1.0011	0.0012		0.3005	1.5060	1.5025	1.0013	0.0006	**T=8000**
	alpha=0.3	***p1=1.5***	***p2=1***	**sigma=1**	**mu=0**		**alpha=0.3**	***p1=1.5***	***p2=1.5***	**sigma=1**	**mu=0**	
	1.085	1.116	1.111	1.026	1.116		1.048	1.142	1.078	1.027	1.073	**T=500**
	1.187	1.184	1.090	1.032	1.206		1.125	1.152	1.089	1.031	1.135	**T=1000**
STD Ratio	1.137	1.129	1.041	1.023	1.159		1.089	1.102	1.072	1.013	1.092	**T=2000**
	1.104	1.085	1.041	1.011	1.111		1.069	1.069	1.046	0.998	1.064	**T=4000**
	1.068	1.053	1.034	0.996	1.068		1.036	1.046	1.025	1.001	1.036	**T=8000**
	alpha=0.3	***p1=2.5***	***p2=1***	**sigma=1**	**mu=0**		**alpha=0.3**	***p1=2.5***	***p2=1.5***	**sigma=1**	**mu=0**	
	0.3013	2.6811	1.0059	1.0029	0.0037		0.3188	2.8367	1.4925	1.0106	0.0231	**T=500**
	0.3016	2.5976	1.0013	1.0016	0.0034		0.3121	2.6939	1.4951	1.0078	0.0146	**T=1000**
Mean of	0.3018	2.5616	1.0003	1.0025	0.0026		0.3057	2.5887	1.4972	1.0030	0.0068	**T=2000**
MLEs	0.3019	2.5405	0.9998	1.0020	0.0024		0.3027	2.5405	1.4973	1.0006	0.0032	**T=4000**
	0.3007	2.5149	0.9998	1.0004	0.0010		0.3017	2.5265	1.4992	1.0007	0.0019	**T=8000**
	alpha=0.3	***p1=2.5***	***p2=1***	**sigma=1**	**mu=0**		**alpha=0.3**	***p1=2.5***	***p2=1.5***	**sigma=1**	**mu=0**	
	0.994	1.089	1.088	1.007	1.010		0.869	1.007	0.997	1.047	0.893	**T=500**
	1.109	1.142	1.070	1.043	1.135		0.974	1.042	1.025	1.035	0.991	**T=1000**
STD Ratio	1.126	1.125	1.080	1.046	1.129		1.008	1.048	1.030	1.021	1.019	**T=2000**
	1.116	1.114	1.058	1.068	1.122		0.993	1.024	1.004	1.025	0.997	**T=4000**
	1.063	1.079	1.033	1.036	1.072		0.987	1.012	1.000	1.016	0.991	**T=8000**

Figure 3: Table (continued). Simulation results for the MLE of the AEPD: averages of estimated parameters (Mean of MLEs) and ratios of simulated standard deviations to theoretical (STD ratio) for different combinations of parameters.