An Asymmetry-Steepness Parameterization of the Generalized Lambda Distribution

Yohan Chalabi[a,b,*], David J. Scott[c], Diethelm Wuertz[a,b]

[a] *Institute for Theoretical Physics, ETH Zurich, Zurich, Switzerland*
[b] *Computational Science and Engineering, ETH Zurich, Zurich, Switzerland*
[c] *Department of Statistics, University of Auckland, Auckland, New Zealand*

Abstract

The generalized lambda distribution (GLD) is a versatile distribution that can accommodate a wide range of shapes, including fat-tailed and asymmetric distributions. It is defined by its quantile function. We introduce a more intuitive parameterization of the GLD that expresses the location and scale parameters directly as the median and inter-quartile range of the distribution. The remaining two shape parameters characterize the asymmetry and steepness of the distribution respectively. This is in contrasts to the previous parameterizations where the asymmetry and steepness are described by the combination of the two tail indices. The estimation of the GLD parameters is notoriously difficult. With our parameterization, the fitting of the GLD to empirical data can be reduced to a two-parameter estimation problem where the location and scale parameters are estimated by their robust sample estimators. This approach also works when the moments of the GLD do not exist. Moreover, the new parameterization can be used to compare data sets in a convenient asymmetry and steepness shape plot. In this paper, we derive the new formulation, as well as the conditions of the various distribution shape regions and moment conditions. We illustrate the use of the asymmetry and steepness shape plot by comparing equities from the NASDAQ-100 stock index.

Keywords: Quantile distributions, generalized lambda distribution, shape plot representation

*Corresponding author. Email address: chalabi@phys.ethz.ch. Postal address: Institut für Theoretische Physik, HIT G 31.5, Wolfgang-Pauli-Str. 27, 8093 Zürich, Switzerland. Tel.: +41 44 633 70 53.

This version: March 2012

1. Introduction

The generalized lambda distribution (GLD) originated in Tukey's lambda distribution [10, 34, 35]. Tukey's lambda is a symmetric distribution that is defined by its quantile function. It was introduced to efficiently generate random variates that approximate other distributions [34, 36]. Soon after the introduction of Tukey's lambda distribution, Hogben [11], Shapiro et al. [27], Shapiro and Wilk [28], and Filliben [5] used the non-symmetric case in sampling studies. Over the years, it became a versatile distribution with location, scale, and shape parameters that can accommodate a large range of distribution shapes. It was used to analyze data and was no longer restricted to approximating other distributions. Figure 1 illustrates the four different shapes of the GLD: unimodal, U-shape, Monotone and S-shape. In applications, the GLD has been used in various fields, such as in option pricing as a fast generator of financial prices [3] or in fitting income data [32]. It has also been used in meteorology [23], in studying the fatigue lifetime prediction of materials [2], in simulations of queue systems [4], in corrosion studies [21], and in independent component analysis [17]. Finally, it has been used in statistical process control [25, 22, 6].

The parameterization of the lambda distribution family has a long history. The parameterizations used nowadays are those of Ramberg and Schmeiser [26] and Freimer et al. [8]. Ramberg and Schmeiser [26] expressed the Tukey lambda distribution with four parameters (RS) as,

$$Q_{RS}(u) = \lambda_1 + \frac{1}{\lambda_2}\left[u^{\lambda_3} - (1-u)^{\lambda_4}\right], \tag{1}$$

where $Q = F^{-1}$ is the quantile function for probabilities u, λ_1 is the location parameter, λ_2 is the scale parameter, and λ_3 and λ_4 are the shape parameters. However, in order to have a valid distribution function where the probability density function f is positive for all x and integrates to one over the range of possible values,

$$f(x) \geq 0 \quad \text{and} \quad \int_{Q(0)}^{Q(1)} f(x)\,\mathrm{d}x = 1, \tag{2}$$

the RS parameterization has complex constraints on the parameters and support regions as summarized in Table 1 and in Fig. 2. Later, Freimer et al. [8] introduced a new parameterization (FKML) to circumvent the constraints on the parameter values. It is defined as,

$$Q_{FKML}(u) = \lambda_1 + \frac{1}{\lambda_2}\left[\frac{u^{\lambda_3}-1}{\lambda_3} - \frac{(1-u)^{\lambda_4}-1}{\lambda_4}\right]. \tag{3}$$

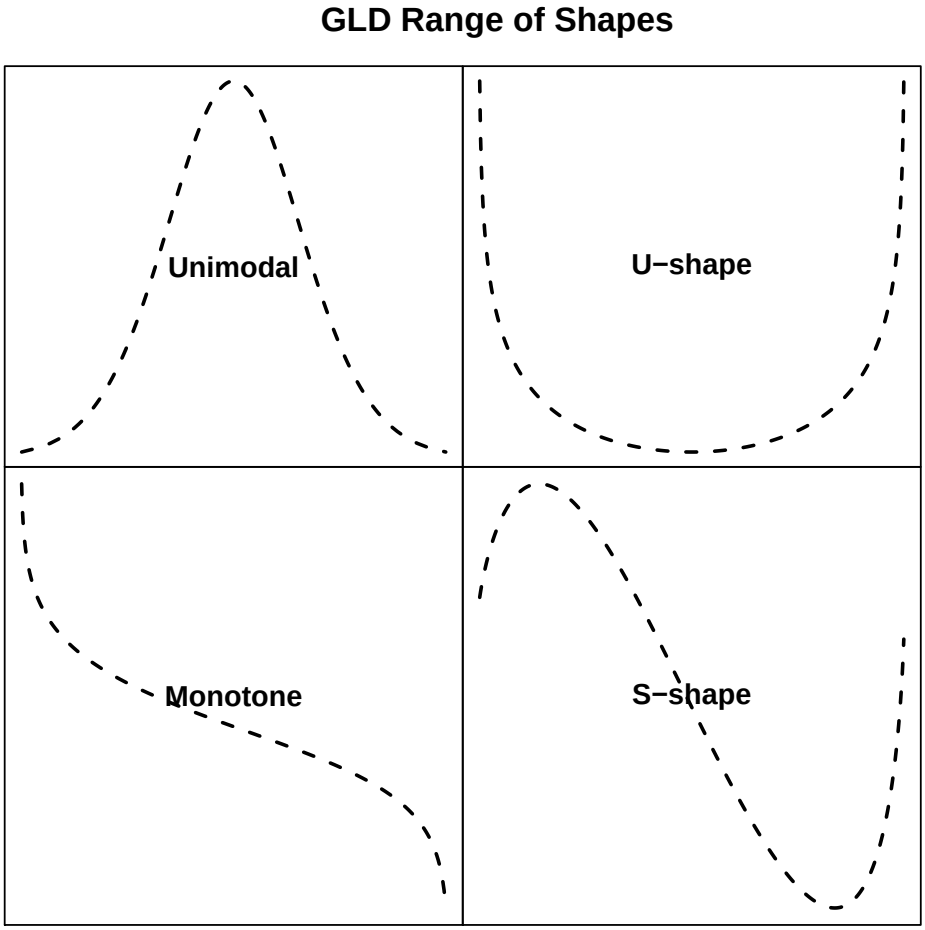

Figure 1: The GLD has four basic shapes: unimodal, U-shape, monotone, and S-shape.

As in the previous parameterization, λ_1 and λ_2 are the location and scale parameters, and λ_3 and λ_4 are the tail index parameters. The advantage over the previous parameterization is that the only constraint on the parameters is that λ_2 must be positive. Figure 3 displays the support regions of the GLD in the FKML parameterization.

The estimation of the GLD parameters for empirical data is notoriously difficult because of the change of the distributional shapes as the parameters are varied in the different regions of the shape parameters. In particular, the support of the distribution can change with the value of the parameters from being a whole real line to an interval, which is infinite in only one direction or finite. In the last decade, several papers have been published to discuss

Region	λ_2	λ_3	λ_4	$Q(0)$	$Q(1)$
1	< 0	≤ -1	≥ 1	$-\infty$	$\lambda_1 + (1/\lambda_2)$
		$-1 < \lambda_3 < 0$ and $\lambda_4 > 1$			
		$\frac{(1-\lambda_3)^{1-\lambda_3}(\lambda_4-1)^{\lambda_4-1}}{(\lambda_4-\lambda_3)^{\lambda_4-\lambda_3}} = \frac{-\lambda_3}{\lambda_4}$			
2	< 0	≥ 1	≤ -1	$\lambda_1 - (1/\lambda_2)$	∞
		$\lambda_3 > 1 \quad \wedge \quad -1 < \lambda_4 < 0$			
		$\frac{(1-\lambda_4)^{1-\lambda_4}(\lambda_3-1)^{\lambda_3-1}}{(\lambda_3-\lambda_4)^{\lambda_3-\lambda_4}} = \frac{-\lambda_4}{\lambda_3}$			
3	> 0	> 0	> 0	$\lambda_1 - (1/\lambda_2)$	$\lambda_1 + (1/\lambda_2)$
		$= 0$	> 0	λ_1	$\lambda_1 + (1/\lambda_2)$
		> 0	$= 0$	$\lambda_1 - (1/\lambda_2)$	λ_1
4	< 0	< 0	< 0	$-\infty$	∞
		$= 0$	< 0	λ_1	∞
		< 0	$= 0$	$-\infty$	λ_1

Table 1: Support regions of the GLD and conditions on the parameters given by the RS parameterization to define a valid distribution function (Karian and Dudewicz [13]). The support regions are displayed in Fig. 2. Note there are no conditions on λ_1.

different parameter estimation approaches. On the one side are the direct estimation methods, such as least-squares estimation with order statistics [24] and with percentiles [15, 13, 7, 18, 14]; the method of moments [23, 9], of L-moments [9, 16], and of trimmed L-moments [1]; and the method of goodness-of-fit with histograms [30] and with maximum likelihood estimation [31]. On the other side, stochastic methods have been introduced with various estimators such as goodness-of-fit [20] or the starship method [19]. Moreover, Shore [29] studied the L2 norm estimator that minimizes the density function distance and the use of nonlinear regression applied to a sample of exact quantile values.

As noted by Gilchrist [9], one of the criticisms of the GLD is that its skewness is expressed in terms of both tail indices $\lambda_{3,4}$ (Eq. 1 and 3). In one approach addressing this concern, a five-parameter GLD was introduced by Joiner and Rosenblatt [12], which, expressed in the FKML parameterization, can be defined as,

$$Q_{FGLD}(u) = \lambda_1 + \frac{1}{2\lambda_2}\left[(1-\lambda_5)\frac{u^{\lambda_3}-1}{\lambda_3} - (1+\lambda_5)\frac{(1-u)^{\lambda_4}-1}{\lambda_4}\right]. \quad (4)$$

It has a location parameter λ_1, a scale parameter λ_2, and an asymmetry

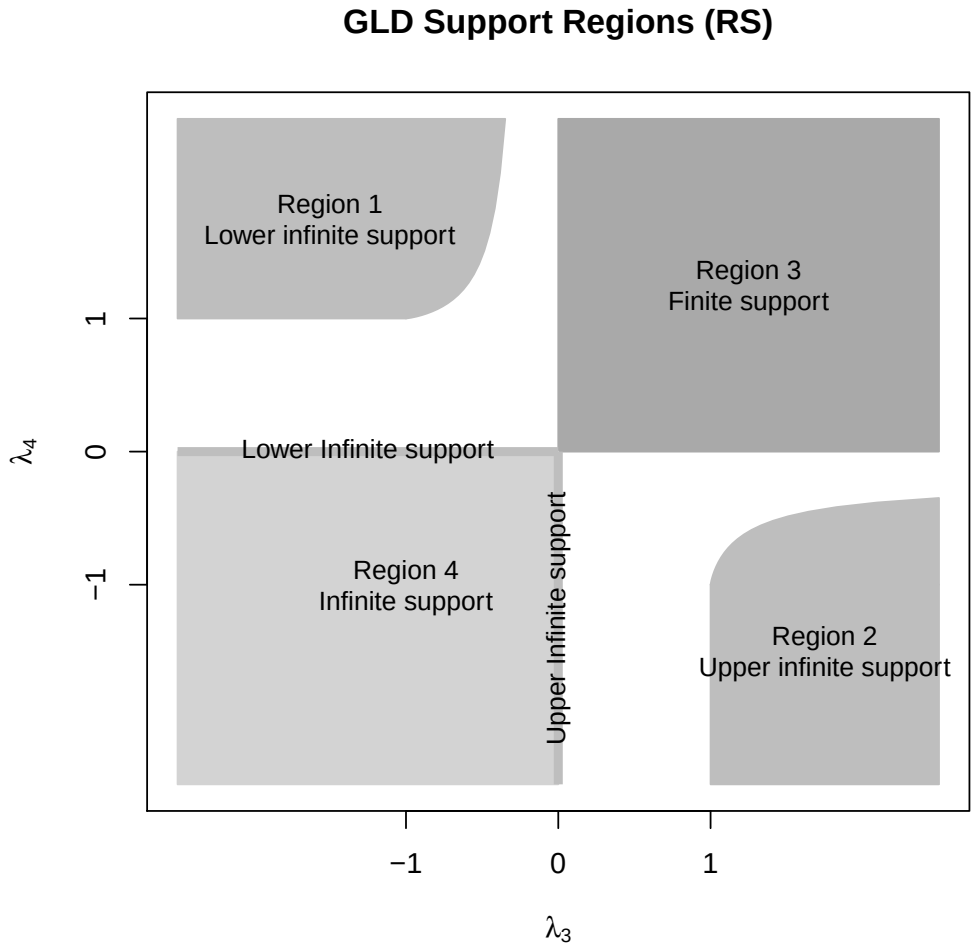

Figure 2: Support regions of the GLD in the RS parameterization that produce valid statistical distributions as described in Table 1.

parameter, λ_5, which weights each side of the distribution and the two tail indices, λ_3 and λ_4. The conditions on the parameters are $\lambda_2 > 0$ and $-1 < \lambda_5 < 1$. However, the addition of a new parameter can make the estimation of the parameters from a data set even more difficult.

In this paper, we show how the four-GLD parameterization (Eq. 3) can be transformed in terms of an asymmetry and steepness parameter without adding a new variable. Moreover, we formulate the location and scale parameters in terms of quantile statistics. The median is used as the location parameter, and the inter-quartile range is used as the scale parameter. The advantage of this parameterization is twofold. First, it brings a clearer inter-

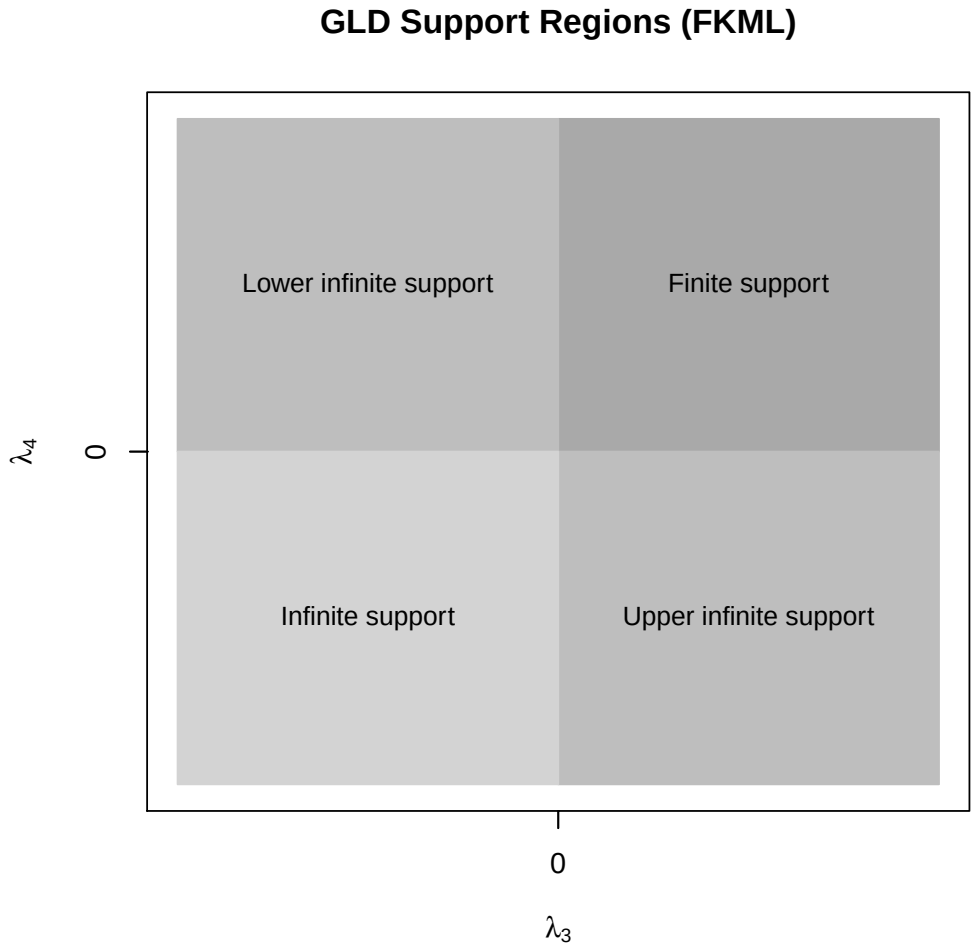

Figure 3: Support regions of the GLD in the FKML parameterization.

pretation of the parameters whereby the asymmetry of the distribution can be distinguished from the steepness of the distribution. This is in contrast to the current parameterizations where the asymmetry is described by the combination of the tail indices. This allows for simpler conditions for the support regions and for the existence of moments for the GLD. Second, the estimation of the parameters for empirical data set can be reduced to a two-parameter estimation problem because the location and scale parameters can be directly estimated by their robust sample estimators. Note that sample quantiles can always be estimated, even when moments of the distribution do not exist. The remaining two shape parameters can be estimated with the usual estimation methods, such as the maximum likelihood estimator.

The remainder of this paper is organized as follows. In §2, we introduce the new asymmetry-steepness parameterization of the GLD. In §3, we derive the support of the distribution in terms of its parameters and illustrate the different support regions in a shape plot. Then, in §4, we give conditions for the existence of its moments, as well as its representation in the shape plot. Based on the support and moment existence conditions, we study in §5 the different shape regions of the distribution and give values of the parameters that make the GLD a good approximation for well-known distributions. We use the new parameterization in §6 to compare the equities of the NASDAQ-100 stock index in the shape plot representation. Conclusions are offered in the last section.

2. Asymmetry-steepness parameterization

In the following section, we express the new parameterization in the FKML form. We consider the GLD quantile function,

$$Q(u) = \lambda_1 + \frac{1}{\lambda_2} S(u|\lambda_3, \lambda_4), \tag{5}$$

where

$$S(u|\lambda_3, \lambda_4) = \begin{cases} \ln(u) - \ln(1-u) & (\lambda_3 = 0,\ \lambda_4 = 0)\,, \\ \ln(u) - \dfrac{1}{\lambda_4}\left[(1-u)^{\lambda_4} - 1\right] & (\lambda_3 = 0,\ \lambda_4 \neq 0)\,, \\ \dfrac{1}{\lambda_3}\left(u^{\lambda_3} - 1\right) - \ln(1-u) & (\lambda_3 \neq 0,\ \lambda_4 = 0)\,, \\ \dfrac{1}{\lambda_3}\left(u^{\lambda_3} - 1\right) - \dfrac{1}{\lambda_4}\left[(1-u)^{\lambda_4} - 1\right] & (\text{otherwise})\,. \end{cases} \tag{6}$$

when $0 < u < 1$. Q is the quantile function for probabilities u; λ_1, λ_2 are the location and scale parameters; and λ_3, λ_4 are the shape parameters jointly related to the strengths of the lower and upper tails. In the limiting cases, when $u = 0$ we have

$$S(0|\lambda_3, \lambda_4) = \begin{cases} -\dfrac{1}{\lambda_3} & (\lambda_3 > 0)\,, \\ -\infty & (\text{otherwise})\,, \end{cases}$$

and when $u = 1$

$$S(1|\lambda_3, \lambda_4) = \begin{cases} \dfrac{1}{\lambda_4} & (\lambda_4 > 0)\,, \\ \infty & (\text{otherwise})\,. \end{cases}$$

We can now use the median, $\tilde{\mu}$, and the inter-quartile range, $\tilde{\sigma}$, to represent the location and scale parameters. Note that we use the tilde to not confuse them with the mean, μ, and the standard deviation, σ. Then, $\tilde{\mu}$ and $\tilde{\sigma}$ are defined by

$$\begin{aligned}\tilde{\mu} &= Q(1/2),\\ \tilde{\sigma} &= Q(3/4) - Q(1/4).\end{aligned}$$

The parameters λ_1 and λ_2 in (6) can therefore be expressed in terms of the median and inter-quartile range as

$$\begin{aligned}\lambda_1 &= \tilde{\mu} - \frac{1}{\lambda_2} S(\frac{1}{2}|\lambda_3, \lambda_4),\\ \lambda_2 &= \frac{1}{\tilde{\sigma}}\left[S(\frac{3}{4}|\lambda_3, \lambda_4) - S(\frac{1}{4}|\lambda_3, \lambda_4)\right].\end{aligned}$$

As mentioned in the introduction, one of the criticisms of the GLD is that the asymmetry and steepness of the distribution depend on both tail indices λ_3 and λ_4. The main idea of this paper is to use distinct shape parameters for the asymmetry and steepness. This is in contrast to the current parameterizations where both tail indices determine the asymmetry and steepness of the distribution. First, it is clear from the definition of the GLD (Eq. 5) that when the tail indices are equal, the distribution is symmetric. Increasing one tail index then produces an asymmetric distribution and a large difference between the tail indices produces a more asymmetric distribution. Second, the steepness of the distribution is related to the size of both tail indices. Indeed, increasing both tail indices would result in a distribution with thinner tails. We can therefore formulate an asymmetry parameter, χ, proportional to the difference between the tail indices and a steepness parameter ξ proportional to the sum of both indices. The remaining step is to map the unbounded interval of $(\lambda_3 - \lambda_4)$ and $(\lambda_3 + \lambda_4)$ to a finite one. To achieve this, we use the transformation

$$y = \frac{x}{\sqrt{1+x^2}} \quad \leftrightarrow \quad x = \frac{y}{\sqrt{1-y^2}} \quad \text{where } y \in (-1, 1) \text{ and } x \in (-\infty, \infty).$$

By taking the proper scaling for the finite intervals, the asymmetry parameter, χ, and the steepness parameter, ξ, can be expressed as

$$\chi = \frac{\lambda_3 - \lambda_4}{\sqrt{1 + (\lambda_3 - \lambda_4)^2}}, \tag{7}$$

$$\xi = \frac{1}{2} - \frac{\lambda_3 + \lambda_4}{2\sqrt{1 + (\lambda_3 + \lambda_4)^2}}. \tag{8}$$

The domain of variation of the shape parameters is given by $\chi \in (-1, 1)$ and $\xi \in (0, 1)$. When χ is equal to 0, the distribution is symmetric. When χ is positive (negative), the distribution is positively (negatively) skewed. Moreover, the GLD becomes steeper when ξ increases. From the parameterization of χ in (7) and ξ in (8), we obtain a system of two equations for the tail indices λ_3 and λ_4,

$$\lambda_3 - \lambda_4 = \frac{\chi}{\sqrt{1 - \chi^2}},$$
$$\lambda_3 + \lambda_4 = \frac{\frac{1}{2} - \xi}{\sqrt{\xi(1 - \xi)}}.$$

This gives, for λ_4 and λ_3,

$$\lambda_3 = \alpha + \beta,$$
$$\lambda_4 = \alpha - \beta,$$

where

$$\alpha = \frac{1}{2} \frac{\frac{1}{2} - \xi}{\sqrt{\xi(1 - \xi)}}, \tag{9}$$

$$\beta = \frac{1}{2} \frac{\chi}{\sqrt{1 - \chi^2}}. \tag{10}$$

We can now formulate the S function in (6) in terms of the shape parameters χ and ξ. We obtain

$$\hat{S}(u|\chi,\xi) = \begin{cases} \ln(u) - \ln(1-u) & \left(\chi = 0,\ \xi = \frac{1}{2}\right), \\ \ln(u) - \frac{1}{2\alpha}\left[(1-u)^{2\alpha} - 1\right] & \left(\chi \neq 0,\ \xi = \frac{1}{2}(1+\chi)\right), \\ \frac{1}{2\beta}(u^{2\beta} - 1) - \ln(1-u) & \left(\chi \neq 0,\ \xi = \frac{1}{2}(1-\chi)\right), \\ \frac{1}{\alpha+\beta}\left(u^{\alpha+\beta} - 1\right) - \frac{1}{\alpha-\beta}\left[(1-u)^{\alpha-\beta} - 1\right] & (\text{otherwise}), \end{cases} \tag{11}$$

where α and β are defined in (9) and (10), and $0 < u < 1$. When $u = 0$, we have

$$\hat{S}(0|\chi,\xi) = \begin{cases} -\frac{1}{\alpha+\beta} & \left(\xi < \frac{1}{2}(1+\chi)\right), \\ -\infty & (\text{otherwise}), \end{cases}$$

and when $u = 1$,

$$\hat{S}(1|\chi,\xi) = \begin{cases} \frac{1}{\alpha-\beta} & \left(\xi < \frac{1}{2}(1-\chi)\right), \\ \infty & (\text{otherwise}). \end{cases}$$

Given the definitions of $\tilde{\mu}$, $\tilde{\sigma}$, χ, ξ, and $\hat{S}$, the quantile function of the GLD becomes

$$\boxed{Q_{CSW}(u|\tilde{\mu},\tilde{\sigma},\chi,\xi) = \tilde{\mu} + \tilde{\sigma}\frac{\hat{S}(u|\chi,\xi) - \hat{S}(\frac{1}{2}|\chi,\xi)}{\hat{S}(\frac{3}{4}|\chi,\xi) - \hat{S}(\frac{1}{4}|\chi,\xi)}}. \tag{12}$$

We will hereinafter use the subscript CSW to denote the new parameterization. Since the cumulative distribution function, F, of the GLD is continuous, we have by definition $F(Q(u)) = u$ for all u. The probability density function $f(x) = F'(x)$ and the quantile density function $q(u) = Q'(u)$ are then related by

$$f\left[Q(u)\right]\ q(u) = 1. \tag{13}$$

$f\left[Q(u)\right]$ is often referred in the literature as the density quantile function, $fQ(u)$. The probability density function of the GLD can then be calculated from the quantile density function. In particular, the quantile density function can be derived from the definition of the quantile function in (12), which gives

$$\boxed{q_{CSW}(u|\tilde{\sigma},\chi,\xi) = \frac{\tilde{\sigma}}{\hat{S}(\frac{3}{4}|\chi,\xi) - \hat{S}(\frac{1}{4}|\chi,\xi)} \frac{\mathrm{d}}{\mathrm{d}\,u}\hat{S}(u|\chi,\xi)}, \tag{14}$$

where

$$\frac{\mathrm{d}}{\mathrm{d}\,u}\hat{S}(u|\chi,\xi) = \begin{cases} \dfrac{1}{u} + \dfrac{1}{1-u} & (\chi = 0,\ \xi = 1/2)\,, \\ \dfrac{1}{u} + (1-u)^{2\alpha-1} & \left(\chi \neq 0,\ \xi = \dfrac{1}{2}(1+\chi)\right), \\ u^{2\beta-1} + \dfrac{1}{1-u} & \left(\chi \neq 0,\ \xi = \dfrac{1}{2}(1-\chi)\right), \\ u^{\alpha+\beta-1} + (1-u)^{\alpha-\beta-1} & (\text{otherwise})\,. \end{cases}$$

It is interesting to note that the limiting sets of shape parameters $\{\chi \to -1, \xi \to 0\}$ and $\{\chi \to 1, \xi \to 0\}$ produce a valid distribution. We obtain the quantile and quantile density functions when $\chi \to -1$ and $\xi \to 0$,

$$\lim_{\substack{\chi\to-1 \\ \xi\to0}} Q_{CSW}(u) = \tilde{\mu} + \tilde{\sigma}\frac{\ln(u) + \ln(2)}{\ln(3)},$$
$$\lim_{\substack{\chi\to-1 \\ \xi\to0}} q_{CSW}(u) = \frac{\tilde{\sigma}}{\ln(3)}\frac{1}{u},$$

and when $\chi \to 1$ and $\xi \to 0$,

$$\lim_{\substack{\chi\to1 \\ \xi\to0}} Q_{CSW}(u) = \tilde{\mu} - \tilde{\sigma}\frac{\ln(1-u) + \ln(2)}{\ln(3)}, \tag{15}$$
$$\lim_{\substack{\chi\to1 \\ \xi\to0}} q_{CSW}(u) = \frac{\tilde{\sigma}}{\ln(3)}\frac{1}{1-u}.$$

However, the other sets of limiting shape parameters do not yield a valid distribution. Details of the calculations for all limiting cases are available in Appendix A.

3. Support

The GLD can accommodate a wide range of distribution shapes and support. In this section, we calculate the conditions on the shape parameters for

the different support regions. The support of the GLD can be derived from the extreme values of $\hat{S}$ in (11)—i.e., when $u = 0, 1$. We obtain

$$\hat{S}(0|\chi,\xi) = \begin{cases} -\dfrac{2\sqrt{\xi(1-\xi)(1-\chi^2)}}{(\frac{1}{2}-\xi)\sqrt{1-\chi^2}+\chi\sqrt{\xi(1-\xi)}} & \left(\xi < \dfrac{1}{2}(1+\chi)\right), \\ -\infty & (\text{otherwise}), \end{cases}$$

and

$$\hat{S}(1|\chi,\xi) = \begin{cases} \dfrac{2\sqrt{\xi(1-\xi)(1-\chi^2)}}{(\frac{1}{2}-\xi)\sqrt{1-\chi^2}-\chi\sqrt{\xi(1-\xi)}} & \left(\xi < \dfrac{1}{2}(1-\chi)\right), \\ \infty & (\text{otherwise}). \end{cases}$$

The GLD thus has infinite support when $\frac{1}{2}(1+|\chi|) \leq \xi$, lower infinite support when $\chi < 0$ and $\frac{1}{2}(1+\chi) \leq \xi < \frac{1}{2}(1-\chi)$, and upper infinite support when $0 < \chi$ and $\frac{1}{2}(1-\chi) \leq \xi < \frac{1}{2}(1+\chi)$. The distribution has finite support in the remaining region. The support regions can be nicely drawn using triangle regions in a shape plot with the asymmetry parameter χ versus steepness parameter ξ as shown in Fig. 4. This contrasts with the complex region supports in the RS parameterization displayed in Fig. 2. Of course, the region supports share the same intuitiveness as the FKML region supports (Fig. 3) since the shape-asymmetry parameterization is based on the FKML parameterization. However, the advantage of the CSW is that its shape parameters have a finite domain of variation and can therefore be represented in a unique plot.

4. Moments

In this section we derive conditions for the existence of moments depending on the shape parameters χ and ξ. The conditions for existence of moments can be determined by expanding the definition of moments to a binomial series and using the existence conditions of the Beta function as Ramberg and Schmeiser [26] and Freimer et al. [8] do. In the λ's representation of FKML (Eq. 3), the condition of existence of the k^{th} moment is $\min(\lambda_3, \lambda_4) = -1/k$. For the parameters α and β defined in (9) and (10), this gives the condition of existence

$$\min(\alpha+\beta, \alpha-\beta) > -1/k.$$

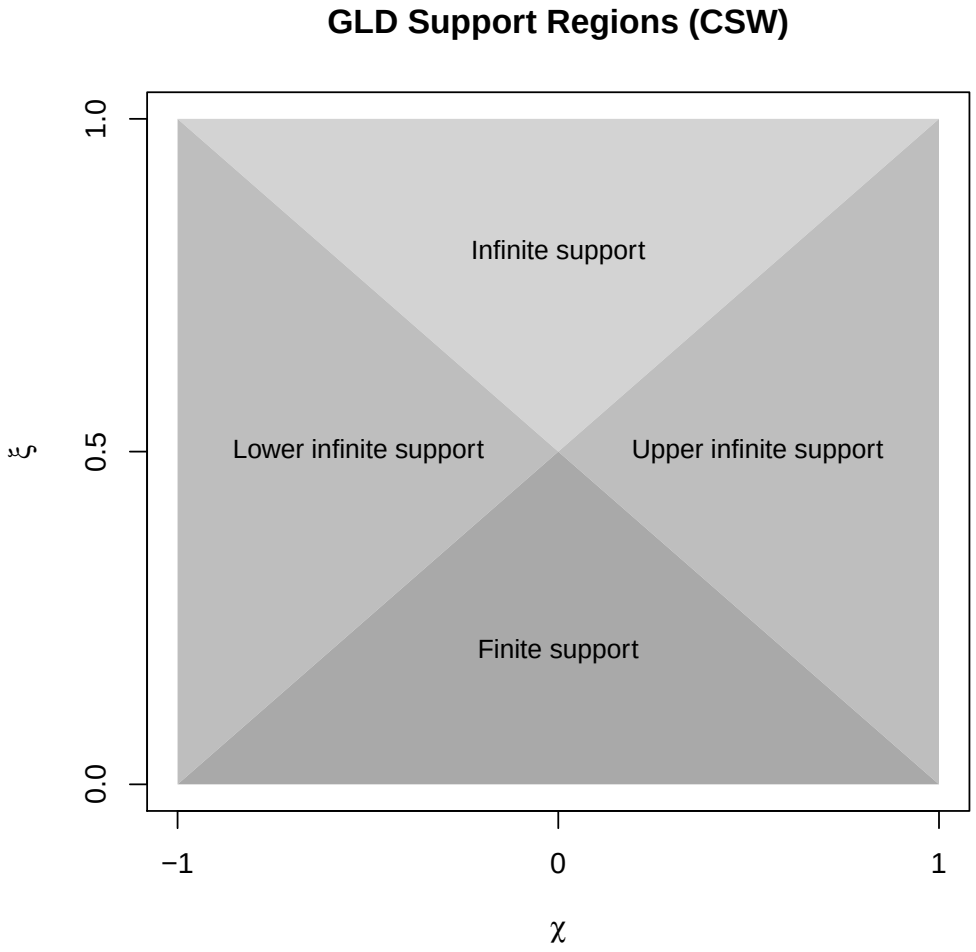

Figure 4: The four support regions of the GLD displayed in the shape plot with asymmetry parameter χ versus the steepness parameter ξ.

After some basic algebraic manipulation, we obtain the condition of existence of the k^{th} moment in terms of the shape parameters χ and ξ,

$$\xi < \frac{1}{2} - H\left(|\chi| - \sqrt{\frac{4}{4+k^2}}\right)\sqrt{\frac{1-2k|\beta|+k^2\beta^2}{4-8k|\beta|+k^2+4k^2\beta^2}},$$

where H is the step function

$$H(x) = \begin{cases} 1 & (x \geq 0), \\ -1 & (\text{otherwise}), \end{cases}$$

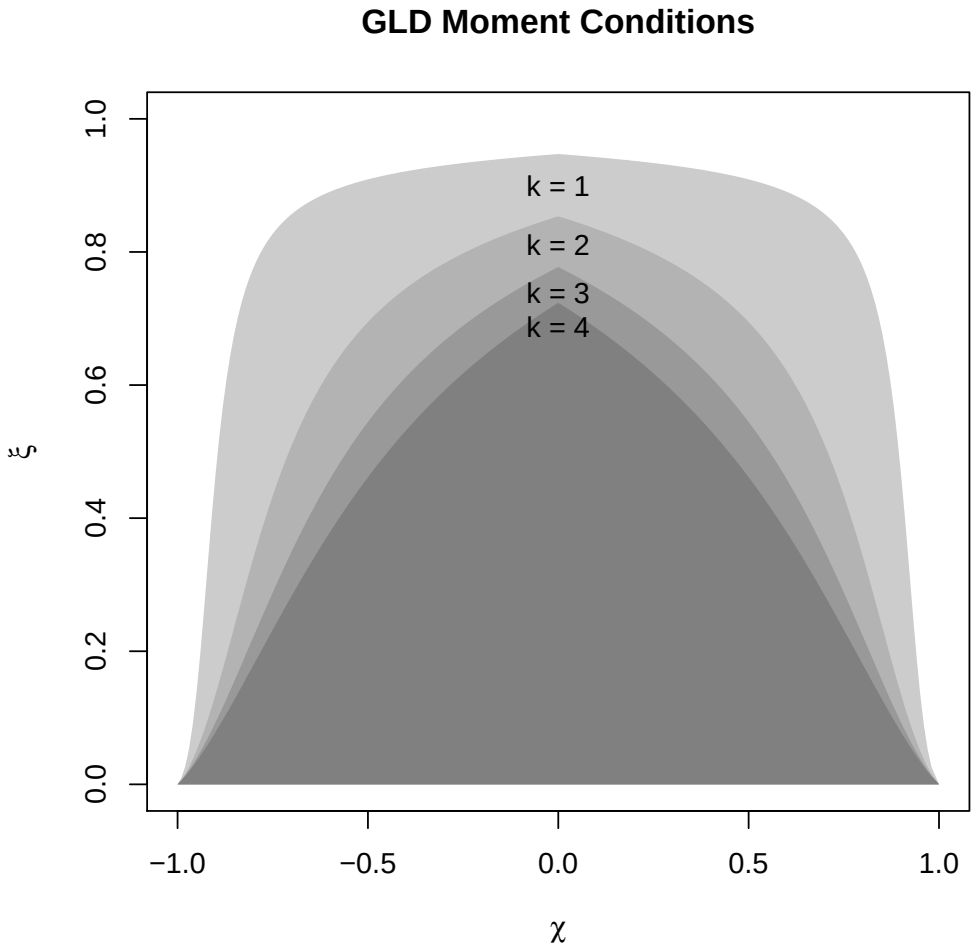

Figure 5: Shape conditions for the existence of moments k. Any set of parameters χ and ξ that is under the k^{th} line defines a distribution with finite k^{th} moment.

and β as defined in (10). Note that in the limiting case when $\chi \to \pm 1$, the k^{th} moment exists when $\xi \to 0$. Figure 5 shows the condition line of existence for the first four moments in the shape diagram. Any set of shape parameters χ and ξ that is under the k^{th} condition line defines a distribution with a finite k^{th} moment.

5. Distribution shape

The GLD has four distribution shapes: unimodal, U-shape, monotone, and S-shape. In this section, we derive the conditions for each distribution

shape based on the derivatives of $\hat{S}$ in (11). Indeed, $\hat{S}$ is proportional to the quantile function, and its derivatives thus provide information on the distribution shapes. First, $\hat{S}$ is an increasing function since it is proportional to the quantile function. Second, the density quantile function is defined as the multiplicative inverse of the quantile density function as noted in (13). The first derivative of $\hat{S}$ is thus always larger than or equal to zero. The second and third derivatives of $\hat{S}$ are

$$\frac{\mathrm{d}^2}{\mathrm{d}\,u^2}\hat{S}(u|\chi,\xi) = \begin{cases} \dfrac{1}{(1-u)^2} - \dfrac{1}{u^2} & (\chi = 0,\ \xi = 1/2)\,, \\ -\dfrac{1}{u^2} - (2\alpha-1)\,(1-u)^{2\alpha-2} & \left(\xi < \dfrac{1}{2}(1+\chi)\right), \\ (2\beta-1)u^{2\beta-2} + \dfrac{1}{(1-u)^2} & \left(\xi < \dfrac{1}{2}(1-\chi)\right), \\ Au^{A-1} - B\,(1-u)^{B-1} & (\text{otherwise})\,, \end{cases}$$

and

$$\frac{\mathrm{d}^3}{\mathrm{d}\,u^3}\hat{S}(u|\chi,\xi) = \begin{cases} \dfrac{2}{u^3} + \dfrac{2}{(1-u)^3} & (\chi = 0,\ \xi = 1/2)\,, \\ \dfrac{2}{u^3} + (2\alpha-2)(2\alpha-1)\,(1-u)^{2\alpha-3} & \left(\xi < \dfrac{1}{2}(1+\chi)\right), \\ (2\beta-2)(2\beta-1)u^{2\beta-3} + \dfrac{2}{(1-u)^3} & \left(\xi < \dfrac{1}{2}(1-\chi)\right), \\ A(A-1)u^{A-2} + B(B-1)\,(1-u)^{B-2} & (\text{otherwise})\,, \end{cases}$$

where $A = \alpha+\beta-1$, $B = \alpha-\beta-1$, and α and β are defined in (9)–(10). We can now deduce the parameter conditions for the four shape regions of the GLD. Figure 6 summarizes the shape regions as described in the remaining part of this section.

5.1. One turning point

The density has one turning point when it is either convex or concave, and its derivative has one zero. Therefore, the density function has one mode when the second derivative of $\hat{S}$ has one zero, and its third derivative is either positive or negative. Note that the second derivative of $\hat{S}$ has at most one zero because $\hat{S}$ is an increasing function.

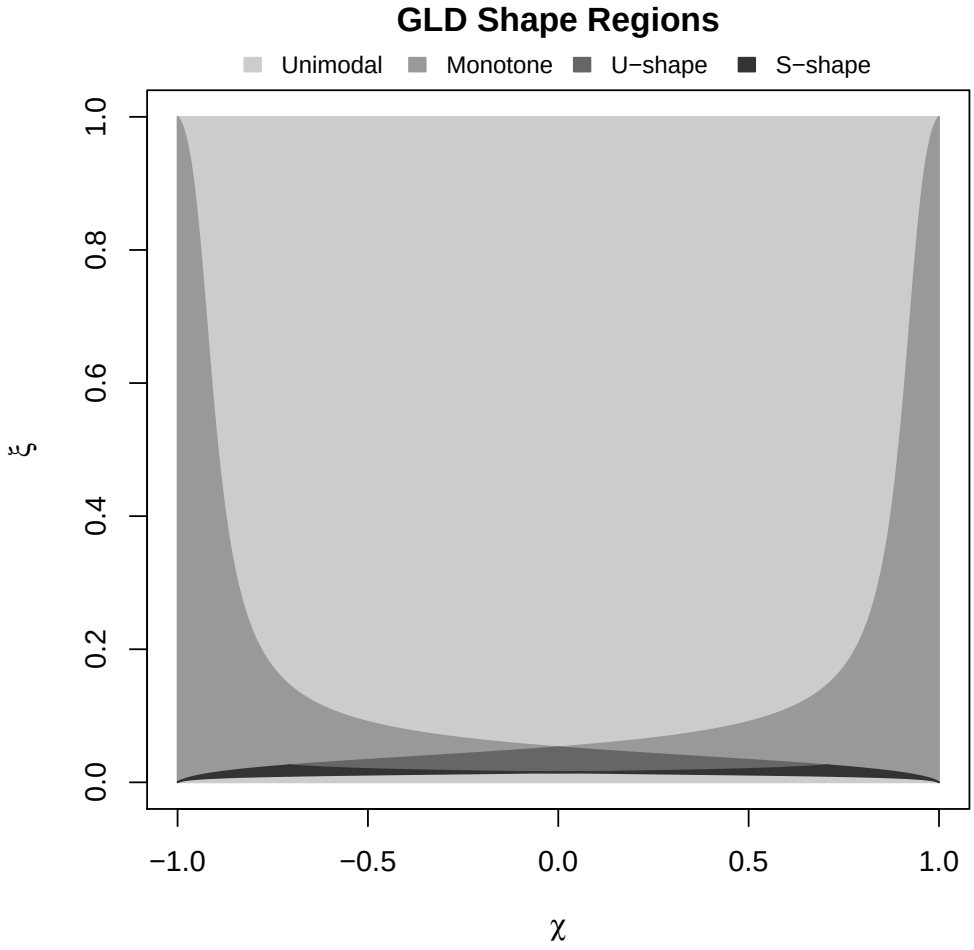

Figure 6: Parameter regions of the GLD shapes in the (χ, ξ) space.

5.1.1. Unimodal

We first consider the case where the third derivative of $\hat{S}$ is positive—that is, when the distribution is unimodal. Note throughout this paper that a unimodal distribution refers to a distribution with a single local maximum. We thus have the conditions on the third derivative and on the end points of the second derivative of S,

$$\frac{\mathrm{d}^3}{\mathrm{d}\,u^3}\hat{S}(u|\chi,\xi) > 0,$$
$$\frac{\mathrm{d}^2}{\mathrm{d}\,u^2}\hat{S}(0|\chi,\xi) < 0,$$
$$\frac{\mathrm{d}^2}{\mathrm{d}\,u^2}\hat{S}(1|\chi,\xi) > 0.$$

These conditions are fulfilled when $\alpha+\beta > 2$ and $\alpha-\beta > 2$ or when $\alpha+\beta < 1$ and $\alpha - \beta < 1$. The calculation is tedious but straightforward. In terms of the shape parameters χ and ξ, the conditions are

$$0 < \xi < \frac{1}{34}\left(17 - 4\sqrt{17}\right),$$
$$-2\sqrt{\frac{4-4\alpha+\alpha^2}{17-16\alpha+4\alpha^2}} < \chi < 2\sqrt{\frac{4-4\alpha+\alpha^2}{17-16\alpha+4\alpha^2}},$$

or

$$\frac{1}{10}\left(5 - 2\sqrt{5}\right) < \xi < 1,$$
$$-2\sqrt{\frac{1-2\alpha+\alpha^2}{5-8\alpha+4\alpha^2}} < \chi < 2\sqrt{\frac{1-2\alpha+\alpha^2}{5-8\alpha+4\alpha^2}}.$$

5.1.2. U-shape

Likewise, the probability density function is U-shaped when the third derivative of $\hat{S}$ is negative and the end points of the second derivative are of different signs. This is the case when $1 < \alpha + \beta < 2$ and $1 < \alpha - \beta < 2$, giving the conditions,

$$\frac{1}{20}\left(10 - 3\sqrt{10}\right) \leq \xi < \frac{1}{10}\left(5 - 2\sqrt{5}\right),$$
$$-2\sqrt{\frac{1-2\alpha+\alpha^2}{5-8\alpha+4\alpha^2}} < \chi < 2\sqrt{\frac{1-2\alpha+\alpha^2}{5-8\alpha+4\alpha^2}},$$

or

$$\frac{1}{34}\left(17 - 4\sqrt{17}\right) < \xi < \frac{1}{20}\left(10 - 3\sqrt{10}\right),$$
$$-2\sqrt{\frac{4-4\alpha+\alpha^2}{17-16\alpha+4\alpha^2}} < \chi < 2\sqrt{\frac{4-4\alpha+\alpha^2}{17-16\alpha+4\alpha^2}}.$$

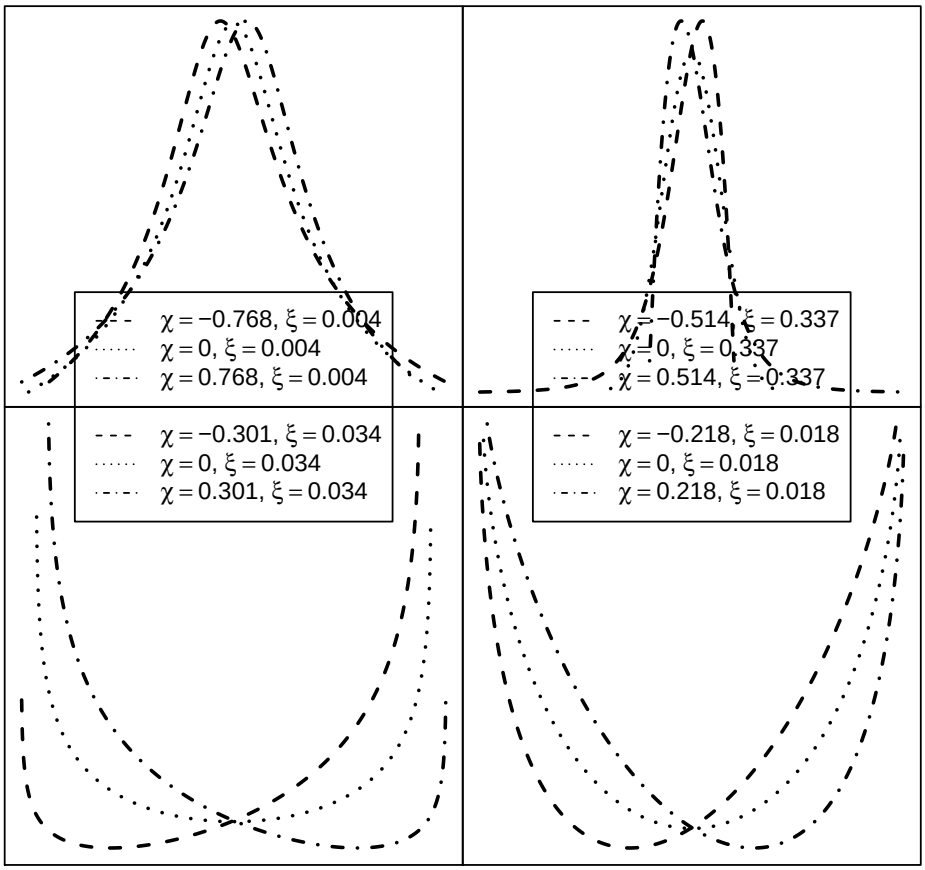

Figure 7: One turning point probability density function of the GLD with different sets of shape parameters.

Note that the parameter conditions of the U-shaped density yield a distribution with finite support, whereas the unimodal shape can have either finite or infinite support. Figure 7 illustrates both the unimodal and U-shaped distribution shapes of the GLD.

5.2. Monotone

The probability density function is monotone when its derivative has no zeros—that is, when the second derivative of $\hat{S}$ does not have any zeros. In

other words, it means that $\hat{S}$ has no inflection point. This is the case when

$$\alpha + \beta > 1 \qquad \text{and} \qquad \alpha - \beta < 1,$$

or

$$\alpha + \beta < 1 \qquad \text{and} \qquad \alpha - \beta > 1.$$

In terms of the shape parameters, the shape conditions are

$$0 < \chi \leq \frac{2}{\sqrt{5}},$$
$$\frac{1}{2} - \sqrt{\frac{1 + 2\beta + \beta^2}{5 + 8\beta + 4\beta^2}} < \xi < \frac{1}{2} - \sqrt{\frac{1 - 2\beta + \beta^2}{5 - 8\beta + 4\beta^2}},$$

or

$$\frac{2}{\sqrt{5}} < \chi < 1,$$
$$\frac{1}{2} - \sqrt{\frac{1 + 2\beta + \beta^2}{5 + 8\beta + 4\beta^2}} < \xi < \frac{1}{2} + \sqrt{\frac{1 - 2\beta + \beta^2}{5 - 8\beta + 4\beta^2}},$$

or

$$-1 < \chi \leq -\frac{2}{\sqrt{5}},$$
$$\frac{1}{2} - \sqrt{\frac{1 - 2\beta + \beta^2}{5 - 8\beta + 4\beta^2}} < \xi < \frac{1}{2} + \sqrt{\frac{1 + 2\beta + \beta^2}{5 + 8\beta + 4\beta^2}},$$

or

$$-\frac{2}{\sqrt{5}} < \chi < 0,$$
$$\frac{1}{2} - \sqrt{\frac{1 - 2\beta + \beta^2}{5 - 8\beta + 4\beta^2}} < \xi < \frac{1}{2} - \sqrt{\frac{1 + 2\beta + \beta^2}{5 + 8\beta + 4\beta^2}}.$$

Figure 8 illustrates the monotone shape of the GLD with a different set of shape parameters.

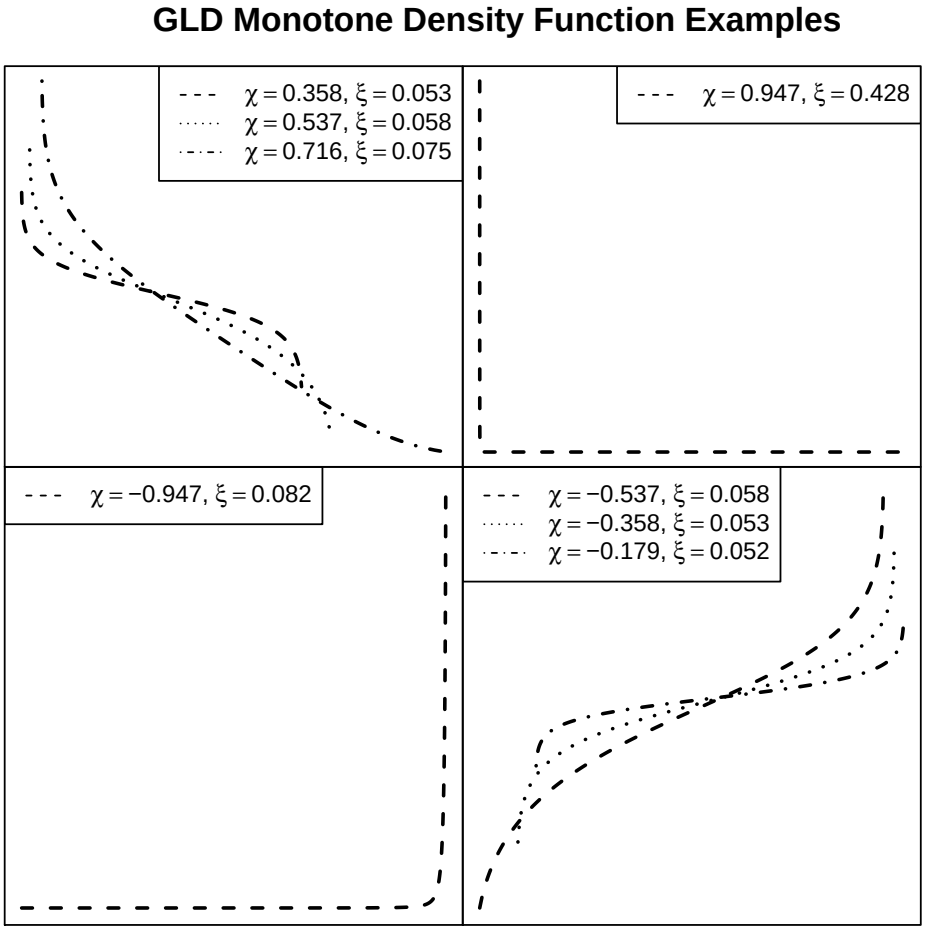

Figure 8: Monotone probability density function of the GLD with different sets of shape parameters.

5.3. S-shape

Given the previous shape regions, the last region is defined by

$$\begin{aligned} \alpha + \beta > 2 \qquad &\text{and} \qquad 1 < \alpha - \beta < 2, \\ 1 < \alpha + \beta < 2 \qquad &\text{and} \qquad \alpha - \beta > 2. \end{aligned}$$

We observe that the second derivative of $\hat{S}$ for $u \in \{0, 1\}$ has the same sign. This indicates that the slope direction at the end points of the probability density function tend toward the same direction. Moreover, the second

derivative of $\hat{S}$ is bounded by a negative and positive number, implying that it has at least two zeros. The probability density function, therefore, has at least two modes. Moreover, the second derivative can be decomposed in two parts, which are both monotonically increasing or decreasing, implying that both functions can be equal, at most, twice. The second derivative, therefore, has exactly two zeros in this region and has an S-shape density when

$$0 < \chi \leq \frac{1}{\sqrt{2}},$$
$$\frac{1}{2} - \sqrt{\frac{4 + 4\beta + \beta^2}{17 + 16\beta + 4\beta^2}} < \xi < \frac{1}{2} - \sqrt{\frac{4 - 4\beta + \beta^2}{17 - 16\beta + 4\beta^2}},$$

or

$$\frac{1}{\sqrt{2}} < \chi < 1,$$
$$\frac{1}{2} - \sqrt{\frac{4 + 4\beta + \beta^2}{17 + 16\beta + 4\beta^2}} < \xi < \frac{1}{2} - \sqrt{\frac{1 + 2\beta + \beta^2}{5 + 8\beta + 4\beta^2}},$$

and

$$-1 < \chi \leq -\frac{1}{\sqrt{2}},$$
$$\frac{1}{2} - \sqrt{\frac{4 - 4\beta + \beta^2}{17 - 16\beta + 4\beta^2}} < \xi < \frac{1}{2} - \sqrt{\frac{1 - 2\beta + \beta^2}{5 - 8\beta + 4\beta^2}},$$

or

$$-\frac{1}{\sqrt{2}} < \chi < 0,$$
$$\frac{1}{2} - \sqrt{\frac{4 - 4\beta + \beta^2}{17 - 16\beta + 4\beta^2}} < \xi < \frac{1}{2} - \sqrt{\frac{4 + 4\beta + \beta^2}{17 + 16\beta + 4\beta^2}}.$$

Figure 9 illustrates the S-shape probability density of the GLD.

5.4. Special Cases

As seen in the previous section, the GLD can accommodate a wide range of distribution shapes. In this section, we estimate the parameters of Q_{CSW}

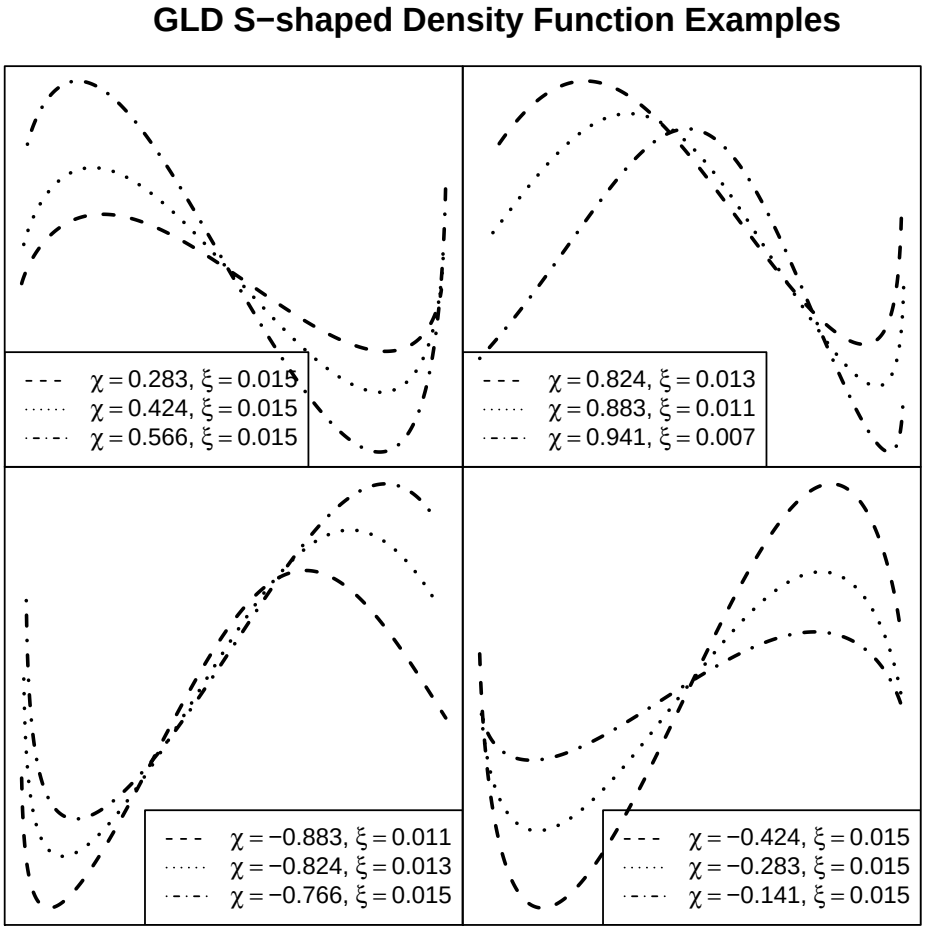

Figure 9: S-shape probability density function of the GLD with different sets of shape parameters.

that replicate common distributions. We start from a vector of equidistant probabilities of length $N = 500$ with probabilities $p_i = \dfrac{i}{N+1}$ with $i = 1, \ldots, N$. We calculate their respective quantiles, x_i, and densities, $d_i = f(x_i)$, for the different distributions. We then fit the shape parameters, $\hat{\chi}$ and $\hat{\xi}$, which minimize the maximum absolute quantile error (MQE),

$$\sup_{\forall i} \left| Q_{_{CSW}}(p_i|\tilde{\mu}, \tilde{\sigma}, \hat{\chi}, \hat{\xi}) - x_i \right| .$$

We use the median and inter-quartile range of the target distributions for the

Distribution	Parameters	$\hat{\chi}$	$\hat{\xi}$	$sup\|Q\|$	sup$\|F\|$	sup$\|f\|$
Normal	$\mu = 0,\ \sigma = 1$	0.0000	0.3661	0.012	0.001	0.001
Student's t	$\nu = 1$	0.0000	0.9434	1.587	0.005	0.012
	$\nu = 5$	0.0000	0.5778	0.069	0.003	0.004
	$\nu = 10$	0.0000	0.4678	0.033	0.002	0.003
Laplace	$\mu = 0,\ b = 1$	0.0000	0.6476	0.257	0.015	0.093
Stable	$\alpha = 1.9,\ \beta = 0$	0.0000	0.5107	0.399	0.010	0.010
	$\alpha = 1.9,\ \beta = 0.5$	0.0730	0.5307	0.584	0.014	0.013
Gamma	$k = 4,\ \theta = 1$	0.4120	0.3000	0.120	0.008	0.012
χ^2	$k = 3$	0.6671	0.1991	0.295	0.015	0.076
	$k = 5$	0.5193	0.2644	0.269	0.011	0.017
	$k = 10$	0.3641	0.3150	0.233	0.007	0.004
Weibull	$k = 3,\ \lambda = 1$	0.0908	0.3035	0.007	0.003	0.016
Log Normal	$\mu = 0,\ \sigma = 0.25$ (log scale)	0.2844	0.3583	0.011	0.007	0.052
Gumbel	$\alpha = 0.5,\ \beta = 2$	-0.3813	0.3624	0.222	0.010	0.010
Inv. Gaussian	$\mu = 1,\ \lambda = 3$	0.5687	0.2957	0.096	0.022	0.175
	$\mu = 0.5,\ \lambda = 6$	0.3267	0.3425	0.008	0.008	0.125
NIG	$\mu = 0,\ \delta = 1,\ \alpha = 2,\ \beta = 1$	0.2610	0.4975	0.124	0.014	0.029
Hyperbolic	$\mu = 0,\ \delta = 1,\ \alpha = 2,\ \beta = 1$	0.2993	0.4398	0.198	0.021	0.030

Table 2: Shape parameters of the GLD to approximate common distributions.

location and scale parameters, $\tilde{\mu}$ and $\tilde{\sigma}$. Note that we explicitly use $N = 500$ in order to compare our estimates to the previous studies of Gilchrist [9], King and MacGillivray [18] and Tarsitano [33].

Table 2 lists the fitted shape parameters for common distributions with the minimum quantile absolute error function in the column denoted by, $\sup|\hat{Q}|$. We also report the maximum probability error, $\sup|\hat{F}|$, as well as the maximum density error, $\sup|\hat{f}|$. We defined the maximum probability error (MPE) as

$$\sup_{\forall i} \left| F_{CSW}(x_i|\tilde{\mu}, \tilde{\sigma}, \hat{\chi}, \hat{\xi}) - p_i \right|,$$

and the maximum density error (MDE) as

$$\sup_{\forall i} \left| f_{CSW}(x_i|\tilde{\mu}, \tilde{\sigma}, \hat{\chi}, \hat{\xi}) - d_i \right|.$$

An important distinction of our method to previous studies is that we used the known values for the location and scale parameters since they are the median and inter-quartile range of the target distributions. This is in contrast with previous work where the location and scale parameters were included in the fitting of the parameters. However, by including the location and scale parameters, the fitted parameters might yield a GLD that does not well approximate the target distribution around its center. This behavior

arises from the change in the estimates of the location and scale, which can alleviate poor fit in the tails of the distribution. However, this is at the expense of having a worse fit around the location. Figure 10 reflects this issue with the Student t distribution with 2 degrees of freedom. The left hand side plot displays the fitted GLD using the exact values for $\tilde{\mu}$ and $\tilde{\sigma}$. The right hand side plot is the fitted GLD obtained when including the location and scale parameters in the estimation. One clearly sees that the center of the distribution is not well described when the location and scale parameters are included in the optimization. However, using the known values for $\tilde{\mu}$ and $\tilde{\sigma}$ yields a higher MQE, 0.307, compared to the other approach, 0.097.

It is interesting to note that the MPE could have been used rather than the MQP to fit the shape parameters. It would have resulted in estimates with smaller MPE. However, we have noticed in practice that the fitted GLD with the MPE estimator does not well approximate the tails of the distribution. This is especially the case with fat-tailed distributions. Figure 11 illustrates this behavior. The left hand side plot displays the fitted log-CDF obtained by minimizing the MQE for the Student t distribution with 2 degrees of freedom. The right hand side plot is the fitted log-CDF obtained with the MPE estimator. Note we used the true median and inter-quartile range in both cases. One clearly sees that the parameters set obtained with the MQE has a better visual fit.

Another important aspect in our approach was to ensure that the fitted parameters yield a GLD with support that include all x_i. This is especially the case for the gamma, χ^2 and Wald distributions. Indeed, one can find a parameter set for which the MQE estimator is smaller, although the fitted distribution does not includes all points.

These considerations motivated our choice to use the known values for the location and scale parameters, to ensure that all points are included in the support of the fitted distribution and to use the MQE estimator to fit the shape parameters. To recall, the fitted shape parameters for common distribution are listed in Table 2.

Besides accommodating a wide range of distributions, the GLD includes the uniform, logistic and exponential distribution. From Eq. (11), (12) and (15) we obtain the parameters of the GLD that replicate the uniform, logistic and exponential distributions as summarized in Table 3. Note that the exponential distribution corresponds to the limiting case $\{\chi \to 1, \xi \to 0\}$ as mentioned at the end of §2.

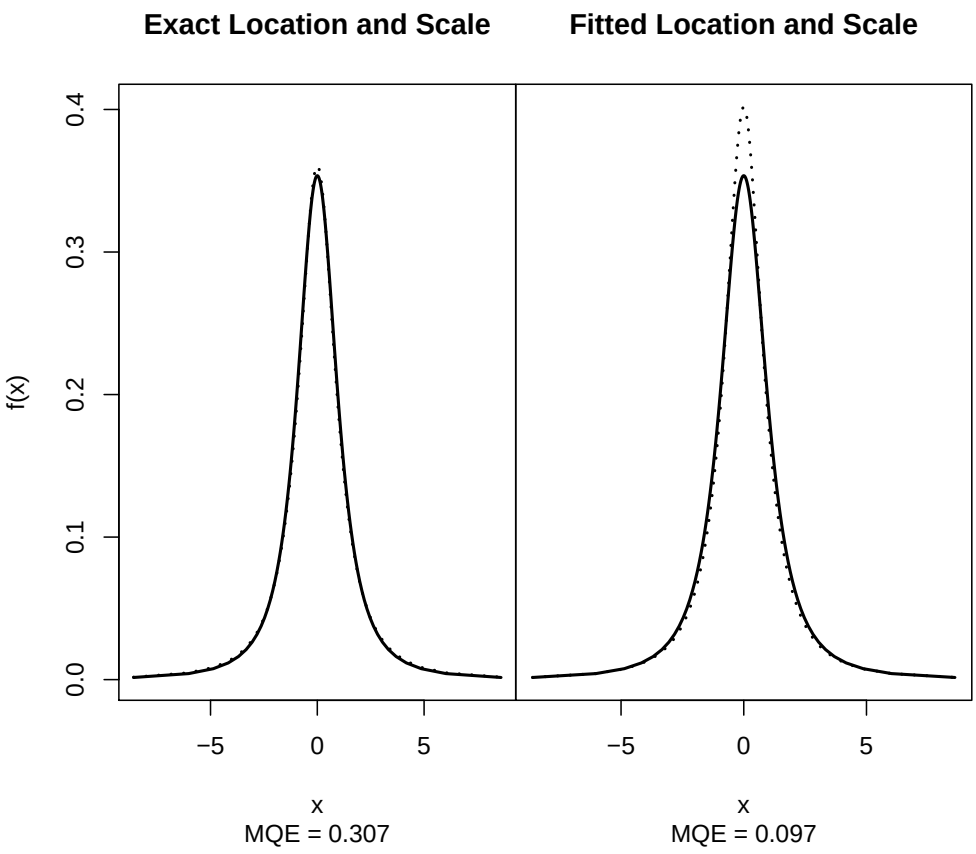

Figure 10: Approximation of the Student t probability density function (dotted lines) with 2 degrees of freedom by either fitting or using the true values for the location and scale parameters, $\tilde{\mu}$ and $\tilde{\sigma}$. In the left hand side plot, the fitted GLD where the true values for the median and inter-quartile range were used. In the right hand side plot, the location and scale parameters were included in the estimation. Note the maximum quantile error (MQE) for the left hand side figure, 0.307, is larger than the one for the right hand side figure, 0.097, although it has a better visual fit.

	$\tilde{\mu}$	$\tilde{\sigma}$	χ	ξ
Uniform (a, b)	$\frac{1}{2}(a+b)$	$\frac{1}{2}(b-a)$	0	$\frac{1}{2} - \frac{1}{\sqrt{5}}$
	$\frac{1}{2}(a+b)$	$\frac{1}{2}(b-a)$	0	$\frac{1}{2} - \frac{2}{\sqrt{17}}$
Logistic (μ, β)	μ	$\beta \ln(9)$	0	$\frac{1}{2}$
Exponential (λ)	$\frac{1}{\lambda}\ln(2)$	$\frac{1}{\lambda}\ln(3)$	1	0

Table 3: Special cases of the GLD in the CSW parameterization.

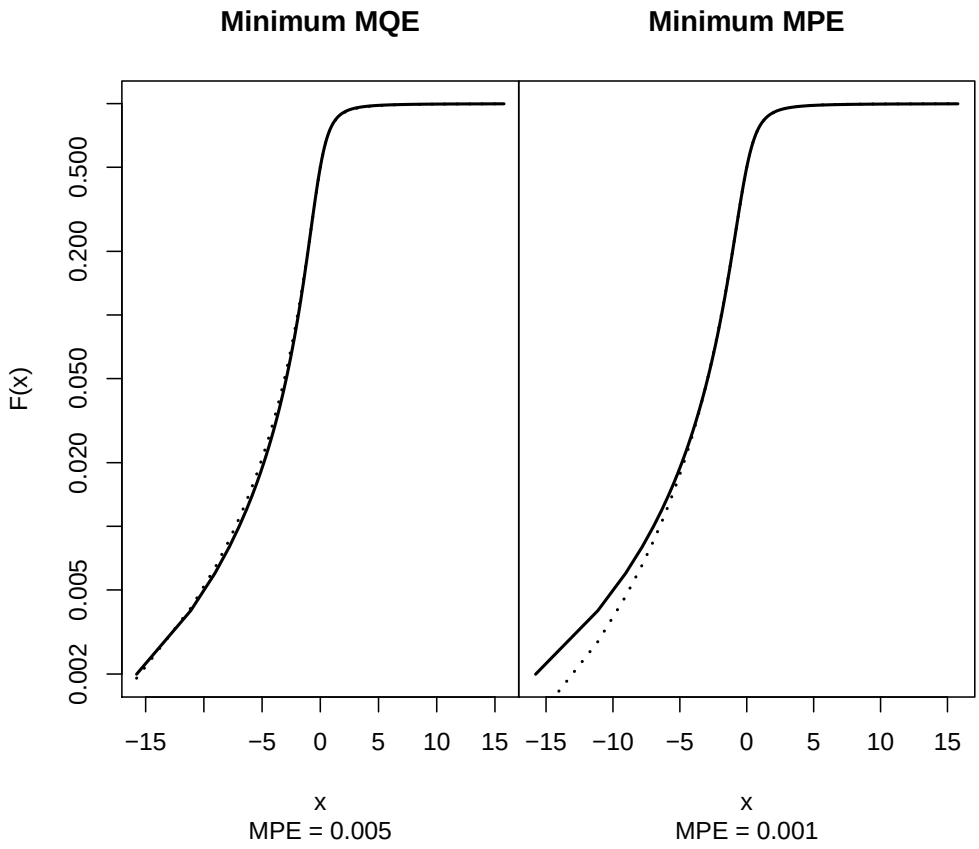

Figure 11: Comparison of the fitted GLD (dotted lines) to the Student t probability density function with 2 degrees of freedom by either using the MQE or the MPE estimator. The left hand side plot was obtained by minimizing the MQE and the right hand side plot with the MPE. The MPE for the left hand side plot, 0.005, is larger than the one for the right hand side figure, 0.001, although it has a better visual fit.

6. Shape Plot Representation

A nice property of having a closed domain of variation for the shape parameters χ and ξ is that we can represent them in a shape plot. In Fig. 12 we illustrate the different shapes, depending on the asymmetry and steepness parameters, with location $\tilde{\mu} = 0$ and scale $\tilde{\sigma} = 1$. The shape plot affords a simple interpretation. The x-axis measures the asymmetry, and the y-axis expresses the heaviness of the tails. The shape plot is thus ideal to compare the fitted shape parameters of a data set. In this section, we illustrate the use of the shape plot with the fitted shape parameters of the equities from the NASDAQ-100 index.

The NASDAQ-100 Index includes 100 of the largest US domestic and

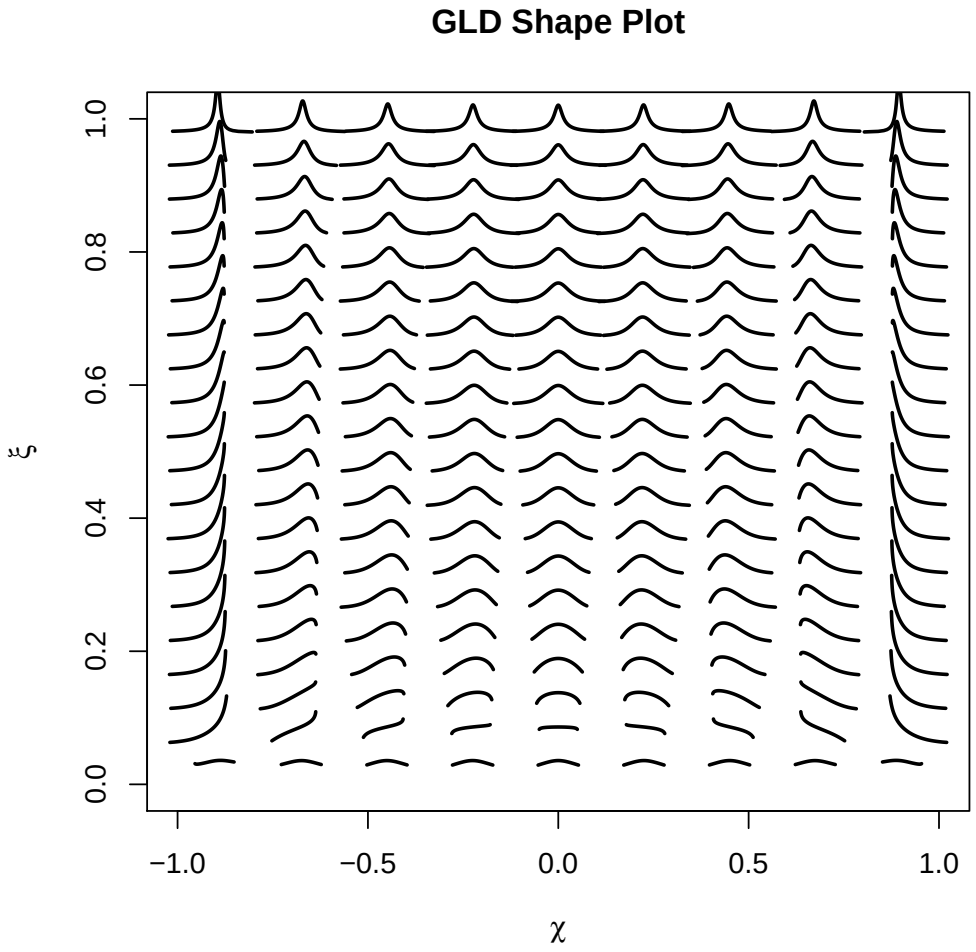

Figure 12: This figure illustrates the different probability density shapes for various steepness and asymmetry parameters. The location and scale of the distribution are, respectively, $\tilde{\mu} = 0$ and $\tilde{\sigma} = 1$.

international non-financial companies listed on the NASDAQ Stock Market. The index reflects the share prices of companies across major industry groups, including computer hardware and software, telecommunications, retail/wholesale trade, and biotechnology. We expect that the listed equities have a wide range of distribution shapes. The GLD is therefore a good candidate for modeling their distribution. We downloaded the financial index series for the NASDAQ-100 from Yahoo's finance web portal and selected the time series with records from 2000–01–03 to 2011–12–31 as reported in Table 4. We used the log-returns of the adjusted closing prices.

AKAM	CHRW	DLTR	FISV	KLAC	MSFT	ORCL	RIMM	SYMC	XLNX
AMAT	CMCSA	EBAY	FLEX	LIFE	MU	ORLY	ROST	TEVA	XRAY
CA	COST	ERTS	FLIR	LLTC	MXIM	PAYX	SBUX	URBN	YHOO
CELG	CSCO	ESRX	HSIC	LRCX	MYL	PCAR	SIAL	VOD	
CEPH	CTSH	EXPD	INFY	MAT	NTAP	PCLN	SNDK	VRSN	
CERN	CTXS	FAST	INTC	MCHP	NVDA	QCOM	SPLS	VRTX	
CHKP	DELL	FFIV	INTU	MICC	NWSA	QGEN	SRCL	WFMI	

Table 4: NASDAQ Symbols. The 66 components of the NASDAQ-100 index that have records from 2001–01–03 to 2011–12–31. Data downloaded from finance.yahoo.com.

We first estimated the location and scale parameters with their sample estimators. We then used the maximum likelihood estimator to fit the shape parameters χ and ξ. Figure 13 shows the fitted shape parameters. It is interesting to note that the fitted parameters are close to the symmetric vertical line at $\chi = 0$. However, the fitted shape parameters are well above the shape parameters that best describe the standard normal distribution, represented by a triangle in the shape plot. The fitted GLD for the components of the NASDAQ has "fatter" tails that the normal distribution. This is one of the typical so-called stylized facts of financial returns.

7. Conclusions

In this paper, we have introduced a new parameterization of the GLD that provides an intuitive interpretation of its parameters. The median and inter-quartile range are the location and scale of the distribution, respectively. The shape parameters describe the asymmetry and the steepness of the distribution. This approach contrasts to previous parameterizations where the skewness of GLD is expressed in terms of both tail indices $\lambda_{3,4}$ (Eq. 1 and 3). Another advantage of this parameterization is that the location and scale parameterization can be estimated from their sample estimators. This reduces the complexity of the estimator used to fit the remaining shape parameters. Moreover, this parameterization enables the use of shape plots that can be used to represent the fitted parameters. However, this new parameterization comes with the cost of more intricate expressions for the condition of the existence of moments and the condition of the shape parameters for the different distribution shapes. Nevertheless, calculating these expressions remains straightforward. The R package **gldist** implements the new parameterization and is available from CRAN.

Acknowledgments

This work is part of the PhD thesis of Yohan Chalabi. He acknowledges financial support from Finance Online GmbH and from the Swiss Federal

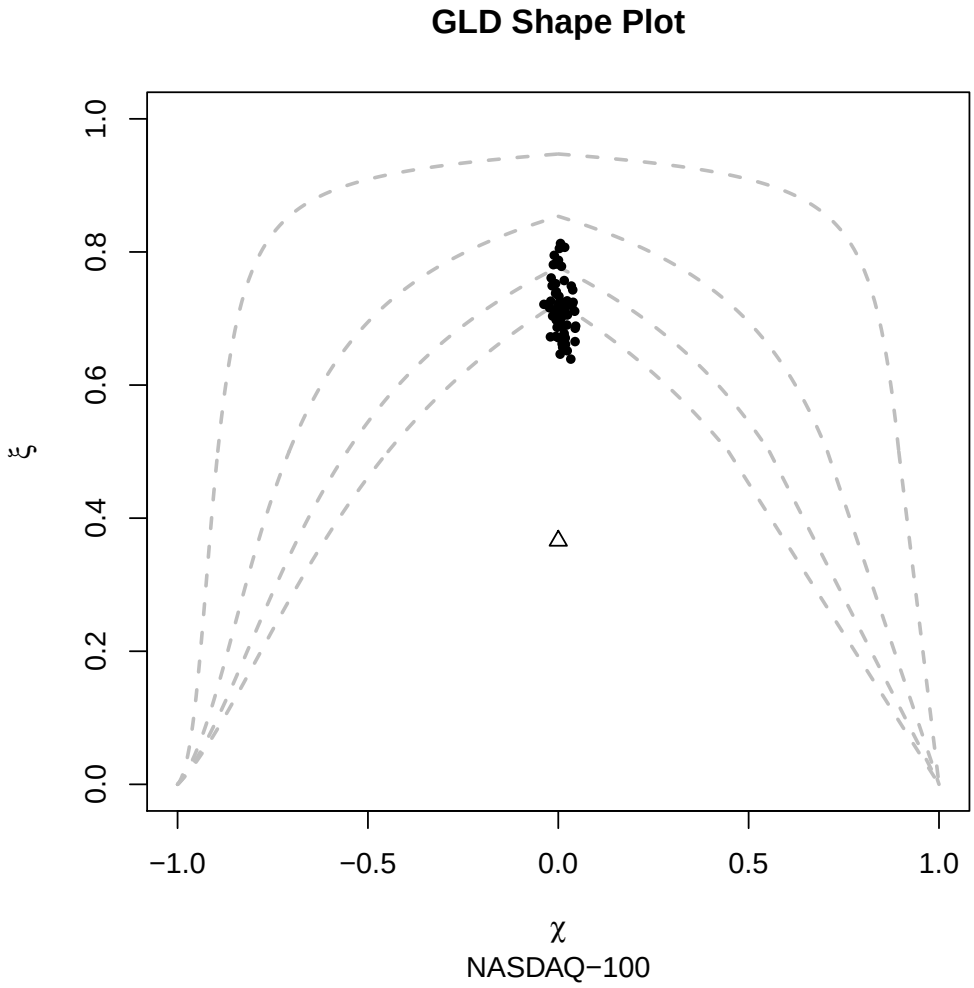

Figure 13: Shape plot of the fitted shape parameters of the NASDAQ Index. The dashed lines represent the existence conditions of the moments of the GLD starting from the existence of the first moment from the top to the existence of higher moments. The triangle symbol corresponds to the GLD shape parameters that best approximate the standard normal distribution.

Institute of Technology in Zurich. David Scott is grateful for the hospitality of the Institute of Theoretical Physics at ETH Zurich.

References

[1] W. H. Asquith. L-moments and TL-moments of the generalized lambda distribution. *Computational Statistics & Data Analysis*, 51(9):4484–4496, May 2007.

[2] M. Bigerelle, D. Najjar, B. Fournier, N. Rupin, and A. Iost. Application of lambda distributions and bootstrap analysis to the prediction of fatigue lifetime and confidence intervals. *International Journal of Fatigue*, 28(3):223–236, Mar. 2006.

[3] C. J. Corrado. Option pricing based on the generalized lambda distribution. *Journal of Futures Markets*, 21(3):213–236, 2001.

[4] B. Dengiz. The generalized lambda distribution in simulation of M/M/1 queue systems. *Journal of the Faculty of Engineering and Architecture of Gazi University*, 3:161–171, 1988.

[5] J. J. Filliben. The probability plot correlation coefficient test for normality. *Technometrics*, 17(1):111–117, Feb. 1975.

[6] B. Fournier, N. Rupin, M. Bigerelle, D. Najjar, and A. Iost. Application of the generalized lambda distributions in a statistical process control methodology. *Journal of Process Control*, 16(10):1087 – 1098, 2006.

[7] B. Fournier, N. Rupin, M. Bigerelle, D. Najjar, A. Iost, and R. Wilcox. Estimating the parameters of a generalized lambda distribution. *Computational Statistics & Data Analysis*, 51(6):2813 – 2835, 2007.

[8] M. Freimer, G. Kollia, G. Mudholkar, and C. Lin. A study of the generalized Tukey lambda family. *Communications in Statistics-Theory and Methods*, 17(10):3547–3567, 1988.

[9] W. Gilchrist. *Statistical modelling with quantile functions*. CRC Press, 2000.

[10] J. Hastings, Cecil, F. Mosteller, J. W. Tukey, and C. P. Winsor. Low moments for small samples: A comparative study of order statistics. *The Annals of Mathematical Statistics*, 18(3):pp. 413–426, 1947.

[11] D. Hogben. Some properties of Tukey's test for non-additivity. Unpublished Ph.D. thesis, Rutgers-The State University, 1963.

[12] B. L. Joiner and J. R. Rosenblatt. Some properties of the range in samples from Tukey's symmetric lambda distributions. *Journal of the American Statistical Association*, 66(334):pp. 394–399, 1971.

[13] Z. Karian and E. Dudewicz. *Fitting statistical distributions: the generalized lambda distribution and generalized bootstrap methods*. Chapman & Hall/CRC, 2000.

[14] Z. Karian and E. Dudewicz. Comparison of GLD fitting methods: superiority of percentile fits to moments in L2 norm. *Journal of Iranian Statistical Society*, 2(2):171–187, 2003.

[15] Z. A. Karian and E. J. Dudewicz. Fitting the generalized lambda distribution to data: a method based on percentiles. *Communications in Statistics - Simulation and Computation*, 28(3):793–819, 1999.

[16] J. Karvanen and A. Nuutinen. Characterizing the generalized lambda distribution by L-moments. *Computational Statistics & Data Analysis*, 52(4):1971–1983, Jan. 2008.

[17] J. Karvanen, J. Eriksson, and V. Koivunen. Maximum likelihood estimation of ICA model for wide class of source distributions. In *Neural Networks for Signal Processing X, 2000. Proceedings of the 2000 IEEE Signal Processing Society Workshop*, volume 1, pages 445 –454 vol.1, 2000.

[18] R. King and H. MacGillivray. Fitting the generalized lambda distribution with location and scale-free shape functionals. *American Journal of Mathematical and Management Sciences*, 27(3-4):441–460, 2007.

[19] R. A. R. King and H. L. MacGillivray. A starship estimation method for the generalized λ distributions. *Australian & New Zealand Journal of Statistics*, 41(3):353–374, 1999.

[20] A. Lakhany and H. Mausser. Estimating the parameters of the generalized lambda distribution. *Algo Research Quarterly*, 3(3):47–58, 2000.

[21] D. Najjar, M. Bigerelle, C. Lefebvre, and A. Iost. A new approach to predict the pit depth extreme value of a localized corrosion process. *ISIJ international*, 43(5):720–725, 2003.

[22] A. Negiz and A. Çinar. Statistical monitoring of multivariable dynamic processes with state-space models. *AIChE Journal*, 43(8):2002–2020, 1997.

[23] A. Ozturk and R. Dale. A study of fitting the generalized lambda distribution to solar-radiation data. *Journal of Applied Meteorology*, 21(7): 995–1004, 1982.

[24] A. Ozturk and R. Dale. Least-squares estimation of the parameters of the generalized lambda-distribution. *Technometrics*, 27(1):81–84, 1985.

[25] S. Pal. Evaluation of nonnormal process capability indices using generalized lambda distribution. *Quality Engineering*, 17(1):77–85, 2004.

[26] J. S. Ramberg and B. W. Schmeiser. An approximate method for generating asymmetric random variables. *Commun. ACM*, 17(2):78–82, 1974.

[27] S. Shapiro, M. Wilk, and H. Chen. A comparative study of various tests for normality. *Journal of the American Statistical Association*, 63(324): 1343–&, 1968.

[28] S. S. Shapiro and M. B. Wilk. An analysis of variance test for normality (complete samples). *Biometrika*, 52(Part 3–4):591–611, Dec. 1965.

[29] H. Shore. Comparison of generalized lambda distribution (GLD) and response modeling methodology (RMM) as general platforms for distribution fitting. *Communications In Statistics-Theory and Methods*, 36 (13-16):2805–2819, 2007.

[30] S. Su. A discretized approach to flexibly fit generalized lambda distributions to data. *Journal of Modern Applied Statistical Methods*, 4(2): 408–424, 2005.

[31] S. Su. Numerical maximum log likelihood estimation for generalized lambda distributions. *Computational Statistics & Data Analysis*, 51(8): 3983–3998, May 2007.

[32] A. Tarsitano. Fitting the generalized lambda distribution to income data. In *COMPSTAT'2004 Symposium*, pages 1861–1867. Physica-Verlag/Springer, 2004.

[33] A. Tarsitano. Comparing estimation methods for the FPLD. *Journal of Probability and Statistics*, 2010.

[34] J. W. Tukey. The practical relationship between the common transformations of percentages or fractions and of amounts. Technical Report Technical Report 36, Statistical Research Group, Princeton, 1960.

[35] J. W. Tukey. The future of data analysis. *The Annals of Mathematical Statistics*, 33(1):1–67, Mar. 1962.

[36] J. Van Dyke. Numerical investigation of the random variable y = $c(u^\lambda - (1-u)^\lambda)$. Unpublished working paper, National Bureau of Standards, 1961.

Appendix A. Limiting cases

In this appendix, we derive the quantile and quantile density functions given in (12) and (14) when the shape parameters tend to their limiting values. We go through all combinations of limiting sets: (i) $\chi \to -1$ and $\xi \to 0$, (ii) $\chi \to 1$ and $\xi \to 0$, (iii) $\chi \to 1$ and $\xi \to 1$, (iv) $\chi \to -1$ and $\xi \to 1$, (v) $\chi \to -1$, (vi) $\chi \to 1$, (vii) $\xi \to 1$, (viii) $\xi \to 0$.

Appendix A.1. $\chi \to -1$ *and* $\xi \to 0$

When $\chi \to -1$ and $\xi \to 0$ we are in the particular case where $\xi = \frac{1}{2}(1+\chi)$ and $\hat{S}(u) = \ln(u) - \frac{1}{a}\left[(1-u)^a - 1\right]$ with $a = \frac{\frac{1}{2} - \xi}{\sqrt{\xi(1-\xi)}}$. We therefore have the limit

$$\begin{aligned}\lim_{\substack{\chi\to-1\\ \xi\to0}} \hat{S}(u) &= \ln(u) - \lim_{a\to\infty} \frac{1}{a}\left[(1-u)^a - 1\right] \\ &= \ln(u).\end{aligned}$$

Using Eq. (12) we obtain the quantile function

$$\lim_{\substack{\chi\to-1\\ \xi\to0}} Q_{CSW}(u) = \tilde{\mu} + \tilde{\sigma}\frac{\ln(u) + \ln(2)}{\ln(3)},$$

and the quantile density function, $q = Q'$, becomes

$$\lim_{\substack{\chi\to-1\\ \xi\to0}} q_{CSW}(u) = \frac{\tilde{\sigma}}{\ln(3)}\frac{1}{u}.$$

Moreover, it is straightforward to show that the obtained quantile function describes a valid probability distribution function.

Appendix A.2. $\chi \to 1$ *and* $\xi \to 0$

When $\chi \to 1$ and $\xi \to 0$, we have $\hat{S}(u) = \frac{1}{b}(u^b - 1) - \ln(1-u)$ with $b = \frac{\chi}{\sqrt{1-\chi^2}}$. We then obtain the limit

$$\begin{aligned}\lim_{\substack{\chi\to1\\ \xi\to0}} \hat{S}(u) &= \lim_{b\to\infty} \frac{1}{b}(u^b - 1) - \ln(1-u) \\ &= -\ln(1-u).\end{aligned}$$

The quantile function becomes

$$\lim_{\substack{\chi\to 1\\ \xi\to 0}} Q_{CSW}(u) = \tilde{\mu} - \tilde{\sigma}\frac{\ln(1-u)+\ln(2)}{\ln(3)},$$

and the quantile density function is

$$\lim_{\substack{\chi\to 1\\ \xi\to 0}} q_{CSW}(u) = \frac{\tilde{\sigma}}{\ln(3)}\frac{1}{1-u}.$$

Again, it is straightforward to show that the obtained quantile function describes a valid probability distribution function.

Appendix A.3. $\chi \to 1$ *and* $\xi \to 1$

When $\chi \to 1$ and $\xi \to 1$ we have $\hat{S}(u) = \ln(u) - \frac{1}{a}\left[(1-u)^a - 1\right]$ where $a = \frac{\frac{1}{2}-\xi}{\sqrt{\xi(1-\xi)}}$. We obtain the limit

$$\begin{aligned}
\lim_{\substack{\chi\to 1\\ \xi\to 0}} Q_{CSW}(u) &= \tilde{\mu} + \tilde{\sigma}\lim_{a\to-\infty}\frac{\ln(u) - \frac{1}{a}\left[(1-u)^a - 1\right] + \ln(2) + \frac{1}{a}\left[\left(\frac{1}{2}\right)^a - 1\right]}{\ln(3) - \frac{1}{a}\left[\left(\frac{1}{4}\right)^a - 1\right] + \frac{1}{a}\left[\left(\frac{3}{4}\right)^a - 1\right]} \\
&= \tilde{\mu} + \tilde{\sigma}\lim_{a\to-\infty}\frac{\left(\frac{1}{2}\right)^a - (1-u)^a}{\left(\frac{3}{4}\right)^a - \left(\frac{1}{4}\right)^a} \\
&= \tilde{\mu} + \tilde{\sigma}\lim_{a\to-\infty}\frac{2^a - (4-4u)^a}{3^a - 1} \\
&= \tilde{\mu} + \tilde{\sigma}\lim_{a\to-\infty}(4-4u)^a \\
&= \begin{cases} \tilde{\mu} & (0 < u < 3/4)\,, \\ \tilde{\mu} + \tilde{\sigma} & (u = 3/4)\,, \\ \infty & (3/4 < u < 1)\,. \end{cases}
\end{aligned}$$

The resulting function diverges for $3/4 < u < 1$, and does therefore not produce a valid probability distribution function.

Appendix A.4. $\chi \to -1$ *and* $\xi \to 1$

When $\chi \to -1$ and $\xi \to 1$ we have $\hat{S}(u) = \frac{1}{b}(u^b - 1) - \ln(1-u)$ with $b = \frac{\chi}{\sqrt{1-\chi^2}}$. We obtain the limit

$$\begin{aligned}
\lim_{\substack{\chi\to 1\\ \xi\to 0}} Q_{CSW}(u) &= \tilde{\mu} + \tilde{\sigma} \lim_{b\to-\infty} \frac{\frac{1}{b}\left(u^b - 1\right) - \ln(1-u) - \frac{1}{b}\left[\left(\frac{1}{2}\right)^b - 1\right]}{\frac{1}{b}\left[\left(\frac{3}{4}\right)^b - 1\right] - \frac{1}{b}\left[\left(\frac{1}{4}\right)^b - 1\right]} \\
&= \tilde{\mu} + \tilde{\sigma} \lim_{b\to-\infty} \frac{u^b - \left(\frac{1}{2}\right)^b}{\left(\frac{3}{4}\right)^b - \left(\frac{1}{4}\right)^b} \\
&= \tilde{\mu} - \tilde{\sigma} \lim_{b\to-\infty} (4u)^b \\
&= \begin{cases} -\infty & (0 < u < 1/4)\,, \\ \tilde{\mu} - \tilde{\sigma} & (u = 1/4)\,, \\ \tilde{\mu} & (1/4 < u < 1)\,. \end{cases}
\end{aligned}$$

The resulting function diverges for $0 < u < 1/4$, and does therefore not produce a valid distribution function.

Appendix A.5. $\chi \to -1$

For the remaining limiting cases, the quantile function is defined as

$$\tilde{Q}_{CSW}(u) = \tilde{\mu} + \tilde{\sigma} \frac{\frac{u^{\alpha+\beta}-1}{\alpha+\beta} - \frac{(1-u)^{\alpha-\beta}-1}{\alpha-\beta} - \frac{\left(\frac{1}{2}\right)^{\alpha+\beta}-1}{\alpha+\beta} + \frac{\left(\frac{1}{2}\right)^{\alpha-\beta}-1}{\alpha-\beta}}{\frac{\left(\frac{3}{4}\right)^{\alpha+\beta}-1}{\alpha+\beta} - \frac{\left(\frac{1}{4}\right)^{\alpha-\beta}-1}{\alpha-\beta} - \frac{\left(\frac{1}{4}\right)^{\alpha+\beta}-1}{\alpha+\beta} + \frac{\left(\frac{3}{4}\right)^{\alpha-\beta}-1}{\alpha-\beta}}, \tag{A.1}$$

where α and β are defined in (9)–(10). When $\chi \to -1$ and $0 < \xi < 1$, we have to consider the limit of $\tilde{Q}_{CSW}$ for $\beta \to -\infty$. From Eq. (A.1) we obtain

$$\begin{aligned}
\lim_{\chi\to-1} Q_{CSW}(u) &= \lim_{\beta\to-\infty} \tilde{Q}_{CSW}(u) \\
&= \tilde{\mu} + \tilde{\sigma} \lim_{\beta\to-\infty} \frac{u^\beta - \left(\frac{1}{2}\right)^\beta}{\left(\frac{3}{4}\right)^\beta - \left(\frac{1}{4}\right)^\beta} \\
&= \tilde{\mu} - \tilde{\sigma} \lim_{\beta\to-\infty} (4u)^\beta\,.
\end{aligned}$$

We thus obtain when $\chi \to -1$,

$$\lim_{\chi\to-1} Q_{CSW}(u) = \begin{cases} -\infty & \left(0 < u < \dfrac{1}{4}\right), \\ \mu - \sigma & \left(u = \dfrac{1}{4}\right), \\ \mu & \left(\dfrac{1}{4} < u < 1\right). \end{cases}$$

The resulting function diverges for $0 < u < 1/4$, and does therefore not produce a valid probability distribution function.

Appendix A.6. $\chi \to 1$

When $\chi \to 1$ and $0 < \xi < 1$, we have $\beta \to \infty$. From Eq. (A.1), we obtain

$$\begin{aligned}\lim_{\chi\to 1} Q_{CSW}(u) &= \lim_{\beta\to\infty} \tilde{Q}_{CSW}(u) \\ &= \tilde{\mu} + \tilde{\sigma} \lim_{\beta\to\infty} \frac{-(1-u)^{-\beta} + \left(\frac{1}{2}\right)^{-\beta}}{\left(\frac{3}{4}\right)^{-\beta} - \left(\frac{1}{4}\right)^{-\beta}} \\ &= \tilde{\mu} + \tilde{\sigma} \lim_{\beta\to\infty} \frac{1}{(4-4u)^{\beta}}.\end{aligned}$$

We then obtain,

$$\lim_{\chi\to 1} Q_{CSW}(u) = \begin{cases} \tilde{\mu} & (0 < u < 3/4)\,, \\ \tilde{\mu} + \tilde{\sigma} & (u = 3/4)\,, \\ \infty & (3/4 < u < 1)\,. \end{cases}$$

The resulting function diverges for $3/4 < u < 1$, and does therefore not produce a valid probability distribution function.

Appendix A.7. $\xi \to 1$

When $\xi \to 1$ and $-1 < \chi < 1$ we have $\alpha \to -\infty$. From Eq. (A.1) we obtain

$$\begin{aligned}\lim_{\xi\to 1} Q_{CSW}(u) &= \lim_{\alpha\to-\infty} \tilde{Q}_{CSW}(u) \\ &= \tilde{\mu} + \tilde{\sigma} \lim_{\alpha\to-\infty} \frac{u^{\alpha+\beta} - (1-u)^{\alpha-\beta} - 2^{-\alpha-\beta} + 2^{\beta-\alpha}}{\left(\frac{3}{4}\right)^{\alpha+\beta} - 4^{\beta-\alpha} - 4^{-\alpha-\beta} + \left(\frac{3}{4}\right)^{\alpha-\beta}} \\ &= \tilde{\mu} + \tilde{\sigma} \lim_{\alpha\to-\infty} \frac{4^{\alpha+\beta}u^{\alpha+\beta} - 2^{\alpha+\beta} + 2^{\alpha+3\beta}}{3^{\alpha+\beta} - 16^{\beta} - 1 + 16^{b}3^{a-b}} \\ &\quad - \tilde{\sigma} \lim_{\alpha\to-\infty} \frac{(4-4u)^{\alpha-\beta} - 2^{\alpha-3\beta} + 2^{\alpha-\beta}}{16^{-\beta}3^{\alpha+\beta} - 1 - 16^{\beta} + 3^{\alpha-\beta}} \\ &= \tilde{\mu} - \frac{\tilde{\sigma}}{1+16^{\beta}} \left[\lim_{\alpha\to-\infty} (4u)^{\alpha+\beta} - 16^{\beta} \lim_{\alpha\to-\infty} (4-4u)^{\alpha-\beta} \right].\end{aligned}$$

We thus obtain

$$\lim_{\xi\to 1} Q_{CSW}(u) = \begin{cases} -\infty & (0 < u < 1/4)\,, \\ \tilde{\mu} - \tilde{\sigma}\left(1+16^{\beta}\right)^{-1} & (u = 1/4)\,, \\ \tilde{\mu} & (1/4 < u < 3/4)\,, \\ \tilde{\mu} + 16^{\beta}\ \tilde{\sigma}\left(1+16^{\beta}\right)^{-1} & (u = 3/4)\,, \\ \infty & (3/4 < u < 1)\,. \end{cases}$$

The resulting function diverges for $1 < u < 1/4$ and $3/4 < u < 1$, and does therefore not produce a valid probability distribution function.

Appendix A.8. $\xi \to 0$

When $\xi \to 0$ and $-1 < \chi < 1$ we have $\alpha \to \infty$. Equation (A.1) becomes

$$\begin{aligned} \lim_{\xi\to 0} Q_{CSW}(u) &= \lim_{\alpha\to\infty} \tilde{Q}_{CSW}(u) \\ &= \tilde{\mu} + \tilde{\sigma} \lim_{\alpha\to\infty} \frac{u^{\alpha+\beta} - (1-u)^{\alpha-\beta} - 2^{-\alpha-\beta} + 2^{\beta-\alpha}}{\left(\frac{3}{4}\right)^{\alpha+\beta} - 4^{\beta-\alpha} - 4^{-\alpha-\beta} + \left(\frac{3}{4}\right)^{\alpha-\beta}} \\ &= \tilde{\mu} + \tilde{\sigma} \lim_{\alpha\to\infty} \frac{\left(\frac{4}{3}\right)^{\alpha+\beta} u^{\alpha+\beta}}{1 - 16^{\beta}3^{-\alpha-\beta} - 3^{-\alpha-\beta} + 9^{-\beta}16^{\beta}} \\ &\quad - \tilde{\sigma} \lim_{\alpha\to\infty} \frac{\left(\frac{4}{3}\right)^{\alpha-\beta} (1-u)^{\alpha-\beta}}{9^{\beta}16^{-\beta} - 3^{\beta-\alpha} - 16^{-\beta}3^{\beta-\alpha} + 1} \\ &= \tilde{\mu} + \frac{\tilde{\sigma}}{9^{\beta}+16^{\beta}} \left[9^{\beta} \lim_{\alpha\to\infty} \left(\frac{4u}{3}\right)^{\alpha+\beta} - 16^{\beta} \lim_{\alpha\to\infty} \left(\frac{4}{3} - \frac{4u}{3}\right)^{\alpha-\beta} \right]. \end{aligned}$$

$$\lim_{\xi\to 0} Q_{CSW}(u) = \begin{cases} -\infty & (0 < u < 1/4)\,, \\ \tilde{\mu} - 16^{\beta}\ \tilde{\sigma}\left(9^{\beta}+16^{\beta}\right)^{-1} & (u = 1/4)\,, \\ \tilde{\mu} & (1/4 < u < 3/4)\,, \\ \tilde{\mu} + 9^{\beta}\ \tilde{\sigma}\left(9^{\beta}+16^{\beta}\right)^{-1} & (u = 3/4)\,, \\ \infty & (3/4 < u < 1)\,. \end{cases}$$

The resulting function diverges for $1 < u < 1/4$ and $3/4 < u < 1$, and does therefore not produce a valid probability distribution function.