



Eidgenössische Technische Hochschule Zürich  
Swiss Federal Institute of Technology Zurich

# Flexible Distribution Modeling with the Generalized Lambda Distribution

Yohan Chalabi and Diethelm Würtz

No. 2012-05



ETH Econophysics Working and White Papers Series  
Online at <https://www.rmetrics.org/WhitePapers>

# Flexible Distribution Modeling with the Generalized Lambda Distribution

Yohan Chalabi\*      Diethelm Wuertz

Institute for Theoretical Physics, ETH Zurich, Switzerland  
Computational Science and Engineering, ETH Zurich, Switzerland

David J. Scott

Department of Statistics, University of Auckland, New Zealand

November 2012

We consider the use of the generalized lambda distribution (GLD) family as a flexible distribution with which to model financial data sets. The GLD can assume distributions with a large range of shapes. Analysts can therefore work with a single distribution to model almost any class of financial assets, rather than needing several. This becomes especially useful in the estimation of risk measures, where the choice of the distribution is crucial for accuracy. We introduce a new parameterization of the GLD, wherein the location and scale parameters are directly expressed as the median and interquartile range of the distribution. The two remaining parameters characterize the asymmetry and steepness of the distribution. Conditions are given for the existence of its moments, and for it to have the appropriate support. The tail behavior is also studied. The new parameterization brings a clearer interpretation of the parameters, whereby the distribution's asymmetry can be more readily distinguished from its steepness. This is in contrast to current parameterizations of the GLD, where the asymmetry and steepness are simultaneously described by a combination of the tail indices. Moreover, the new parameterization can be used to compare data sets in a convenient asymmetry and steepness shape plot.

---

\*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.

**Keywords** Quantile distributions · Generalized lambda distribution · Shape plot representation

**JEL Classification** C16 · C13

## 1 Introduction

It is well known that distributions of financial returns can exhibit heavy-tails, skewness, or both. This trait is often taken as stylized fact, and can be modeled by  $\alpha$ -stable distributions. The drawback to using these distributions is that they do not all have closed form for their densities and distribution functions. In this regard, the generalized lambda distribution (GLD) offers an alternative. The four-parameter GLD family is known for its high flexibility. It can create distributions with a large range of different shapes. It shares the heavy-tail and skewness properties of the  $\alpha$ -stable distribution. Excepting Corrado [2001] and Tarsitano [2004], there has been little application of the GLD to financial matters. However, as will be covered in this chapter, the shape range of the GLD family is so large that it can accommodate almost any financial time series. Analysts can therefore work with a single distribution, rather than needing several. This becomes especially useful in the estimation of risk measures, where the choice of the distribution is crucial for accuracy. Furthermore, the GLD is defined by a quantile function. Its parameters can be estimated even when its moments do not exist.

This paper introduces a more intuitive parameterization of the GLD that expresses the location and scale parameters directly as the median and interquartile range of the distribution. The two remaining shape parameters respectively characterize the asymmetry and steepness of the distribution. This is in contrast to the standard parameterization, where the asymmetry and steepness are described by a combination of the two tail indices. In Section 2 we review the history of the generalized lambda distribution. Then, in Section 3, the new formulation is derived, along with the conditions of the various distribution shape regions and the moment conditions.

The estimation of standard GLD parameters is notoriously difficult. For the new parameterization, fitting the GLD to empirical data can be reduced to a two-step estimation problem, wherein 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. Section 4 describes an extensive empirical study that compares different GLD parameter estimation methods. The author's contribution is to introduce the maximum product of spacing estimator (MPS) for the GLD. Also presented, is a new robust moment matching method where the estimation of the two shape parameters is independent of the location and scale parameters. The methods developed in this work allow easy fitting of the GLD to empirical data, such as financial returns. An extensive simulation study to compare different GLD parameter estimators is also performed.

Applications of the new GLD parameterization are presented in Section 5. First, the new parameterization is used to compare data sets in a convenient asymmetry and steepness shape plot. The use of the asymmetry and steepness shape plot is illustrated by comparing equities returns from the NASDAQ-100 stock index. Second, widely used financial risk measures, such as the value-at-risk and expected shortfall, are calculated for the GLD. Third, the shape plot is used to illustrate the apparent scaling law and self-similarity of high frequency returns.

## 2 The generalized lambda distribution

The GLD is an extension of Tukey's lambda distribution [Hastings et al., 1947; Tukey, 1960, 1962]. Tukey's lambda distribution is symmetric and is defined by its quantile function. It was introduced to efficiently redistribute uniform random variates into approximations of other distributions [Tukey, 1960; Van Dyke, 1961]. Soon after the introduction of Tukey's lambda distribution, Hogben [1963], Shapiro et al. [1968], Shapiro and Wilk [1965], and Filliben [1975] introduced the non-symmetric case in sampling studies. Over the years, Tukey's lambda distribution became a versatile distribution, with location, scale, and shape parameters, that could accommodate a large range of distribution forms. It gained use as a data analysis tool and was no longer restricted to approximating other distributions. Figure 1 illustrates the four qualitatively different shapes of the GLD: unimodal, U-shaped, monotone and S-shaped. The GLD has been applied to numerous fields, including option pricing as a fast generator of financial prices [Corrado, 2001], and in fitting income data [Tarsitano, 2004]. It has also been used in meteorology [Ozturk and Dale, 1982], in studying the fatigue lifetime prediction of materials [Bigerelle et al., 2006], in simulations of queue systems [Dengiz, 1988], in corrosion studies [Najjar et al., 2003], in independent component analysis [Karvanen et al., 2000], and in statistical process control [Pal, 2004; Negiz and Çinar, 1997; Fournier et al., 2006].

The parameterization of the extended Tukey's lambda distributions has a long history of development. Contemporary parameterizations are those of Ramberg and Schmeiser [1974] and Freimer et al. [1988]. Ramberg and Schmeiser [1974] generalized the Tukey lambda distribution to four parameters called the RS parameterization:

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

where  $Q_{\text{RS}} = F_{\text{RS}}^{-1}$  is the quantile function for probabilities  $u$ ,  $\lambda_1$  is the location parameter,  $\lambda_2$  is the scale parameter, and  $\lambda_3$  and  $\lambda_4$  are the shape parameters. 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 allowed domain, i.e.,

$$f(x) \geq 0 \quad \text{and} \quad \int_{Q(0)}^{Q(1)} f(x) dx = 1,$$

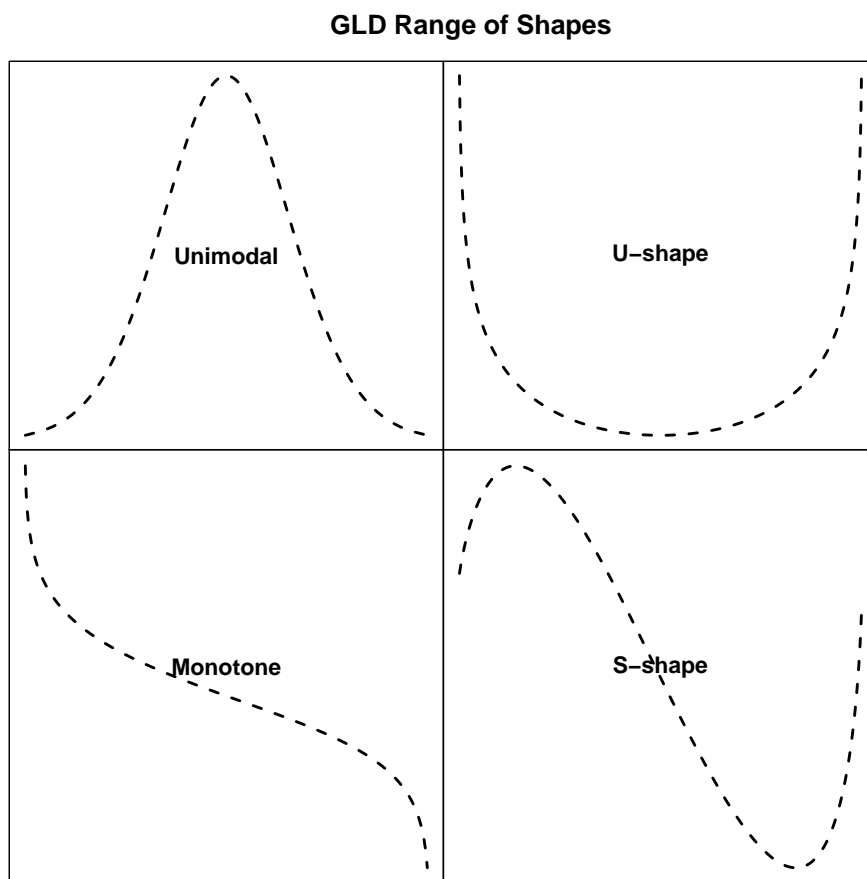


Figure 1: The GLD has four basic shapes: unimodal, U-shape, monotone, and S-shape.

Region	$\lambda_2$	$\lambda_3$	$\lambda_4$	$Q(0)$	$Q(1)$
1	$< 0$	$\leq -1$	$\geq 1$	$-\infty$	$\lambda_1 + (1/\lambda_2)$
		$\begin{cases} -1 < \lambda_3 < 0 \text{ and } \lambda_4 > 1 \\ \frac{(1-\lambda_3)^{1-\lambda_3}(\lambda_4-1)^{\lambda_4-1}}{(\lambda_4-\lambda_3)^{\lambda_4-\lambda_3}} = \frac{-\lambda_3}{\lambda_4} \end{cases}$			
2	$< 0$	$\geq 1$	$\leq -1$	$\lambda_1 - (1/\lambda_2)$	$\infty$
		$\begin{cases} \lambda_3 > 1 \quad \wedge \quad -1 < \lambda_4 < 0 \\ \frac{(1-\lambda_4)^{1-\lambda_4}(\lambda_3-1)^{\lambda_3-1}}{(\lambda_3-\lambda_4)^{\lambda_3-\lambda_4}} = \frac{-\lambda_4}{\lambda_3} \end{cases}$			
3	$> 0$	$> 0$	$> 0$	$\lambda_1 - (1/\lambda_2)$	$\lambda_1 + (1/\lambda_2)$
		$= 0$	$> 0$	$\lambda_1$	$\lambda_1 + (1/\lambda_2)$
		$> 0$	$= 0$	$\lambda_1 - (1/\lambda_2)$	$\lambda_1$
4	$< 0$	$< 0$	$< 0$	$-\infty$	$\infty$
		$= 0$	$< 0$	$\lambda_1$	$\infty$
		$< 0$	$= 0$	$-\infty$	$\lambda_1$

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 [2000]). The support regions are displayed in Fig. 2. Note that there are no conditions on  $\lambda_1$  to obtain a valid distribution.

the RS parameterization has complex constraints on the parameters and support regions as summarized in Table 1 and Fig. 2.

Later, Freimer et al. [1988] introduced a new parameterization called FKML to circumvent the constraints on the RS parameter values. It is expressed as

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

As in the previous parameterization,  $\lambda_1$  and  $\lambda_2$  are the location and scale parameters, and  $\lambda_3$  and  $\lambda_4$  are the tail index parameters. The advantage over the previous parameterization is that the only constraint on the parameters is that  $\lambda_2$  must be positive. Figure 3 displays the support regions of the GLD in the FKML parameterization.

Estimation of the GLD parameters for empirical data is notoriously difficult because of interdependencies between the parameters, rapid changes in the distribution forms and supports. In particular, the support of the distribution is dependent upon the parameter values, and can vary from the entire real line to a semi-infinite, or finite, interval. In the last decade, several papers have been published to discuss two different parameter estimation philosophies. On one side of the discourse are the direct estimation methods, such as least-squares estimation with

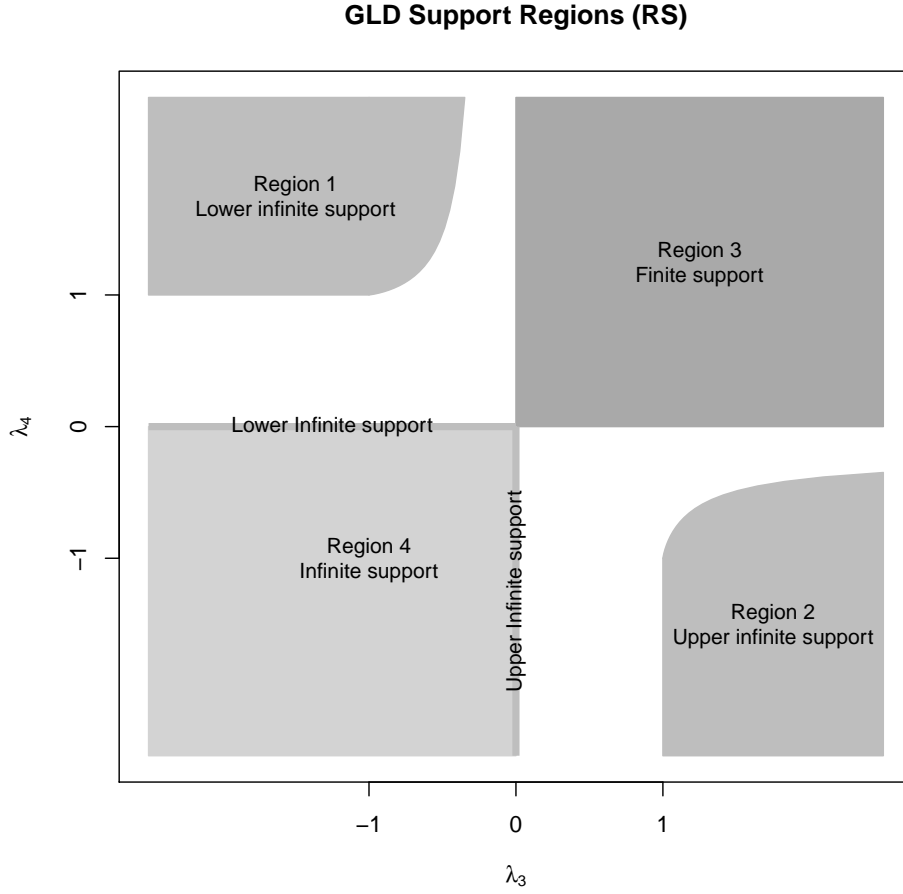


Figure 2: Support regions of the GLD in the RS parameterization that produce valid statistical distributions as described in Table 1.

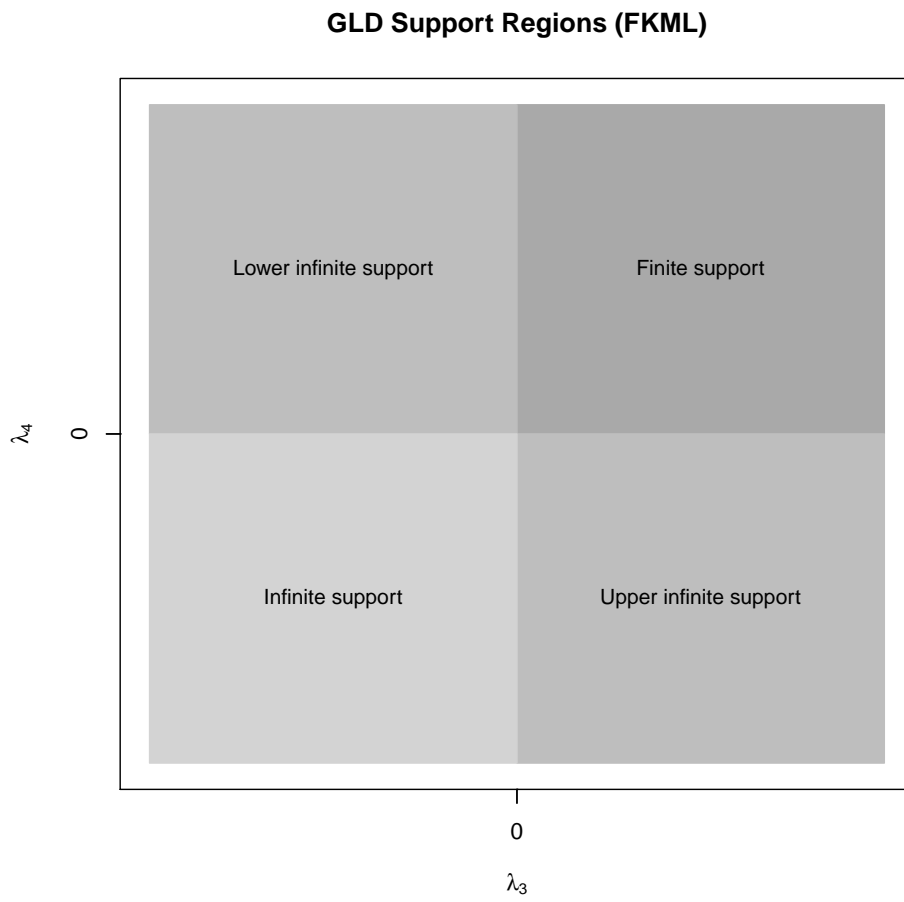


Figure 3: Support regions of the GLD in the FKML parameterization.



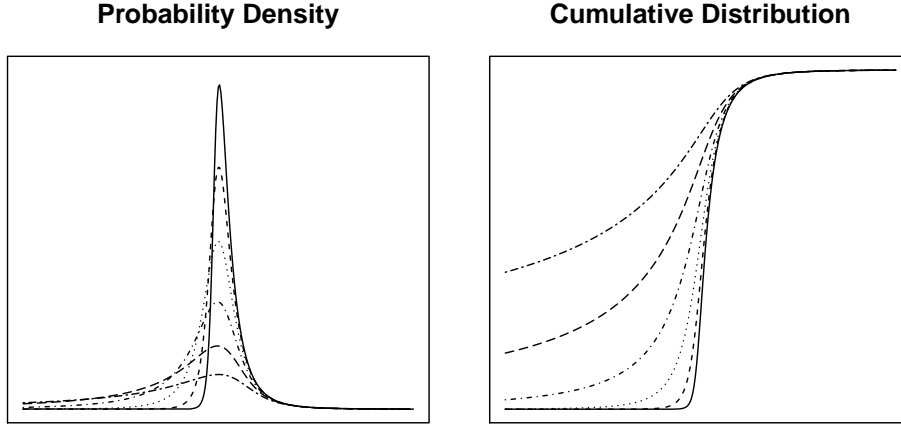


Figure 4: Probability density and cumulative distribution functions for the GLD in the RS parameterization. The location and scale parameters are,  $\lambda_1 = 0$  and  $\lambda_2 = -1$ , respectively. The right tail index is fixed at  $\lambda_4 = -1/4$  and the left tail index,  $\lambda_3$ , varies in the range  $\{-1/16, -1/8, -1/4, -1/2, -1, -2\}$ .

order statistics [Ozturk and Dale, 1985] and with percentiles [Karian and Dudewicz, 1999, 2000; Fournier et al., 2007; King and MacGillivray, 2007; Karian and Dudewicz, 2003]; the methods of moments [Ozturk and Dale, 1982; Gilchrist, 2000], L-moments [Gilchrist, 2000; Karvanen and Nuutinen, 2008], and trimmed L-moments [Asquith, 2007]; and the goodness-of-fit method with histograms [Su, 2005] and with maximum likelihood estimation [Su, 2007]. On the other side, stochastic methods have been introduced with various estimators such as goodness-of-fit [Lakhany and Mausser, 2000] or the starship method [King and MacGillivray, 1999]. Moreover, Shore [2007] studied the  $L^2$ -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 [2000], one of the criticisms of the GLD is that its skewness is expressed in terms of both tail indices  $\lambda_3$  and  $\lambda_4$  in Eqs. (1) and (2). In one approach addressing this concern, a five-parameter GLD was introduced by Joiner and Rosenblatt [1971], which, expressed in the FKML parameterization, can be written as,

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

It has a location parameter  $\lambda_1$ , a scale parameter  $\lambda_2$ , and an asymmetry parameter,  $\lambda_5$ , which weights each side of the distribution and the two tail indices,  $\lambda_3$  and  $\lambda_4$ . The conditions on the parameters are  $\lambda_2 > 0$  and  $-1 < \lambda_5 < 1$ . However, the additional parameter can make the estimation of the parameter values even more difficult.

### 3 A shape parameterization of the GLD

In this chapter, it is shown how the four-parameter GLD of Eq. (2) can be transformed in terms of an asymmetry and steepness parameter without adding a new variable. Moreover, the location and scale parameters are reformulated in terms of quantile statistics. The median is used as the location parameter, and the interquartile range is used as the scale parameter.

The new parameterization brings a clearer interpretation of the parameters, whereby the asymmetry of the distribution can be more readily distinguished from the steepness of the distribution. This is in contrast to the RS and FKML parameterizations, where the asymmetry is described by a combination of the tail indices. The new formulation allows for relaxed constraints upon the support regions, and for the existence of moments for the GLD.

Another advantage of the new parameterization is that the estimation of the parameters for an empirical data set can be reduced to a two-step 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 two shape parameters can then be estimated by one of the usual estimation methods, such as the maximum likelihood estimator.

The remainder of this chapter is organized as follows. In Section 3.1, the new asymmetry-steepness parameterization of the GLD is introduced. In Section 3.2, the support of the distribution is derived in terms of its parameters, and the different support regions are illustrated in a shape plot. Then, in Section 3.3, conditions are given for the existence of its moments, and a shape plot representation is made. In Section 3.4, the tail behavior of the GLD in the new parameterization is studied. Based on the support and moment existence conditions, Section 3.5 studies the different shape regions of the distribution. Section 3.6 lists sets of parameter values that make the GLD a good approximation for several well-known distributions.

#### 3.1 An asymmetry-steepness parameterization

In the following section, the new parameterization is expressed in the FKML form. Consider the GLD quantile function:

$$Q(u) = \lambda_1 + \frac{1}{\lambda_2} S(u|\lambda_3, \lambda_4), \quad (3)$$

where

$$S(u|\lambda_3, \lambda_4) = \begin{cases} \ln(u) - \ln(1-u) & \text{if } \lambda_3 = 0, \lambda_4 = 0, \\ \ln(u) - \frac{1}{\lambda_4} \left[ (1-u)^{\lambda_4} - 1 \right] & \text{if } \lambda_3 = 0, \lambda_4 \neq 0, \\ \frac{1}{\lambda_3} (u^{\lambda_3} - 1) - \ln(1-u) & \text{if } \lambda_3 \neq 0, \lambda_4 = 0, \\ \frac{1}{\lambda_3} (u^{\lambda_3} - 1) - \frac{1}{\lambda_4} \left[ (1-u)^{\lambda_4} - 1 \right] & \text{otherwise,} \end{cases} \quad (4)$$

for  $0 < u < 1$ .  $Q$  is the quantile function for probabilities  $u$ ;  $\lambda_1$  and  $\lambda_2$  are the location and scale parameters; and  $\lambda_3$  and  $\lambda_4$  are the shape parameters jointly related to the strengths of the lower and upper tails. In the limiting case  $u = 0$ :

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

In the limiting case  $u = 1$ :

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

The median,  $\tilde{\mu}$ , and the interquartile range,  $\tilde{\sigma}$ , can now be used to represent the location and scale parameters. These are defined by

$$\tilde{\mu} = Q(1/2), \tag{5a}$$

$$\tilde{\sigma} = Q(3/4) - Q(1/4). \tag{5b}$$

The parameters  $\lambda_1$  and  $\lambda_2$  in Eq. (4) can therefore be expressed in terms of the median and interquartile range as

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

As mentioned in the introduction, one of the criticisms of the GLD is that the asymmetry and steepness of the distribution are both dependent upon both of the tail indices,  $\lambda_3$  and  $\lambda_4$ . The main idea in this chapter is to use distinct shape parameters for the asymmetry and steepness. First, it is clear, from the definition of the GLD in Eq. (3), that when the tail indices are equal, the distribution is symmetric. Increasing one tail index then produces an asymmetric distribution. Second, the steepness of the distribution is related to the size of both tail indices: increasing both tail indices results in a distribution with thinner tails. Now, formulate an asymmetry parameter,  $\chi$ , proportional to the difference between the two tail indices, and a steepness parameter,  $\xi$ , proportional to the sum of both tail indices. The remaining step is to map the unbounded interval of  $(\lambda_3 - \lambda_4)$  to the interval  $(-1, 1)$ , and  $(\lambda_3 + \lambda_4)$  to the interval  $(0, 1)$ . To achieve this, we use the transformation

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

The asymmetry parameter,  $\chi$ , and the steepness parameter,  $\xi$ , can then be expressed as

$$\chi = \frac{\lambda_3 - \lambda_4}{\sqrt{1 + (\lambda_3 - \lambda_4)^2}}, \tag{6a}$$

$$\xi = \frac{1}{2} - \frac{\lambda_3 + \lambda_4}{2\sqrt{1 + (\lambda_3 + \lambda_4)^2}}, \tag{6b}$$

where the respective domains of the shape parameters are  $\chi \in (-1, 1)$  and  $\xi \in (0, 1)$ . When  $\chi$  is equal to 0, the distribution is symmetric. When  $\chi$  is positive (negative), the distribution is positively (negatively) skewed. Moreover, the GLD becomes steeper when  $\xi$  increases. The parameterization of  $\chi$  and  $\xi$  in Eq. (6), yields a system of two equations for the tail indices  $\lambda_3$  and  $\lambda_4$ :

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

This gives

$$\begin{aligned}\lambda_3 &= \alpha + \beta, \\ \lambda_4 &= \alpha - \beta,\end{aligned}$$

where

$$\alpha = \frac{1}{2} \frac{\frac{1}{2} - \xi}{\sqrt{\xi(1 - \xi)}}, \quad (7a)$$

$$\beta = \frac{1}{2} \frac{\chi}{\sqrt{1 - \chi^2}}. \quad (7b)$$

The  $S$  function of Eq. (4) can now be formulated in terms of the shape parameters  $\chi$  and  $\xi$ ;

$$S(u|\chi, \xi) = \begin{cases} \ln(u) - \ln(1 - u) & \text{if } \chi = 0, \xi = \frac{1}{2}, \\ \ln(u) - \frac{1}{2\alpha} \left[ (1 - u)^{2\alpha} - 1 \right] & \text{if } \chi \neq 0, \xi = \frac{1}{2}(1 + \chi), \\ \frac{1}{2\beta} (u^{2\beta} - 1) - \ln(1 - u) & \text{if } \chi \neq 0, \xi = \frac{1}{2}(1 - \chi), \\ \frac{1}{\alpha + \beta} (u^{\alpha + \beta} - 1) - \frac{1}{\alpha - \beta} \left[ (1 - u)^{\alpha - \beta} - 1 \right] & \text{otherwise,} \end{cases} \quad (8)$$

where  $\alpha$  and  $\beta$  are defined in Eq. (7), and  $0 < u < 1$ . When  $u = 0$ ;

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

and when  $u = 1$ ;

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

Given the definitions of  $\tilde{\mu}$ ,  $\tilde{\sigma}$ ,  $\chi$ ,  $\xi$ , and  $S$  in Eqs. (5), (6) and (8), the quantile function of the GLD becomes

$$Q_{\text{CSW}}(u|\tilde{\mu}, \tilde{\sigma}, \chi, \xi) = \tilde{\mu} + \tilde{\sigma} \frac{S(u|\chi, \xi) - S(\frac{1}{2}|\chi, \xi)}{S(\frac{3}{4}|\chi, \xi) - S(\frac{1}{4}|\chi, \xi)}. \quad (9)$$

Hereinafter the subscript CSW shall be used to denote the new parameterization.

Since the quantile function,  $Q_{\text{CSW}}$ , is continuous, it immediately follows by definition that the cumulative distribution function is  $F_{\text{CSW}}[Q_{\text{CSW}}(u)] = u$  for all probabilities  $u \in [0, 1]$ . The probability density function,  $f(x) = F'(x)$ , and the quantile density function,  $q(u) = Q'(u)$ , are then related by

$$f[Q(u)] q(u) = 1. \quad (10)$$

The literature often refers to  $f[Q(u)]$  as the density quantile function,  $fQ(u)$ .

The probability density function of the GLD can be calculated from the quantile density function. In particular, the quantile density function can be derived from the definition of the quantile function in Eq. (9). This gives

$$q_{\text{CSW}}(u|\tilde{\sigma}, \chi, \xi) = \frac{\tilde{\sigma}}{S(\frac{3}{4}|\chi, \xi) - S(\frac{1}{4}|\chi, \xi)} \frac{d}{du} S(u|\chi, \xi), \quad (11)$$

where

$$\frac{d}{du} S(u|\chi, \xi) = u^{\alpha+\beta-1} + (1-u)^{\alpha-\beta-1},$$

with  $\alpha$  and  $\beta$  defined in Eq. (7).

It is interesting to note that the limiting sets of shape parameters;  $\{\chi \rightarrow -1, \xi \rightarrow 0\}$  and  $\{\chi \rightarrow 1, \xi \rightarrow 0\}$ , produce valid distributions. When  $\chi \rightarrow -1$  and  $\xi \rightarrow 0$ , the quantile function is

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

and the quantile density function is

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

When, instead,  $\chi \rightarrow 1$  and  $\xi \rightarrow 0$ ,

$$\lim_{\substack{\chi \rightarrow 1 \\ \xi \rightarrow 0}} Q_{\text{CSW}}(u) = \tilde{\mu} - \tilde{\sigma} \frac{\ln(1-u) + \ln(2)}{\ln(3)}, \quad (12)$$

and

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

Other sets of limiting shape parameters do not, however, yield valid distributions.

### 3.2 Support

The GLD can accommodate a wide range of distribution shapes and supports. In this section, the conditions imposed upon the shape parameters for the different support regions are calculated.

The support of the GLD can be derived from the extreme values of  $S$  in Eq. (8). When  $u = 0$ ;

$$S(0|\chi, \xi) = \begin{cases} -\frac{2\sqrt{\xi(1-\xi)(1-\chi^2)}}{(\frac{1}{2}-\xi)\sqrt{1-\chi^2}+\chi\sqrt{\xi(1-\xi)}} & \text{if } \xi < \frac{1}{2}(1+\chi), \\ -\infty & \text{otherwise,} \end{cases} \quad (13)$$

and when  $u = 1$ ;

$$S(1|\chi, \xi) = \begin{cases} \frac{2\sqrt{\xi(1-\xi)(1-\chi^2)}}{(\frac{1}{2}-\xi)\sqrt{1-\chi^2}-\chi\sqrt{\xi(1-\xi)}} & \text{if } \xi < \frac{1}{2}(1-\chi), \\ \infty & \text{otherwise.} \end{cases} \quad (14)$$

The GLD thus has unbounded infinite support,  $(-\infty, \infty)$ , when  $\frac{1}{2}(1-|\chi|) \leq \xi$ ; semi-infinite support bounded above,  $(-\infty, Q_{\text{CSW}}(1)]$ , when  $\chi < 0$  and  $\frac{1}{2}(1+\chi) \leq \xi < \frac{1}{2}(1-\chi)$ ; and semi-infinite support bounded below,  $[Q_{\text{CSW}}(0), \infty)$ , 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,  $\xi \leq \frac{1}{2}(1-|\chi|)$ .

As shown in Fig. 5, the support regions can be depicted using triangular regions in a shape plot with the asymmetry parameter  $\chi$  versus steepness parameter  $\xi$ . This contrasts with the complex region supports of the RS parameterization displayed in Fig. 2. Of course, the region supports share the same intuitiveness as the FKML region supports in Fig. 3, since the shape-asymmetry parameterization is based on the FKML parameterization. The advantage of the CSW parameterization is that its shape parameters have finite domains of variation and the distribution family can therefore be represented by a single plot.

### 3.3 Moments

In this section, shape parameter ( $\chi$  and  $\xi$ ) dependent conditions for the existence of GLD moments are derived. As in Ramberg and Schmeiser [1974] and Freimer et al. [1988], conditions for the existence of moments can be determined by expanding the definition of the  $k$ th moment of the GLD into a binomial series where the binomial series can be represented by the Beta function. The existence condition of the moment then follow the existence condition of the obtained Beta function. In the  $\lambda$ s representation of FKML in Eq. (2), the existence condition of the  $k$ th moment is  $\min(\lambda_3, \lambda_4) > -1/k$ . For the parameters  $\alpha$  and  $\beta$  defined in Eq. (7), this gives the existence condition

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

Some simple algebraic manipulation produces the condition of existence of the  $k$ th moment in terms of the shape parameters  $\chi$  and  $\xi$ ;

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

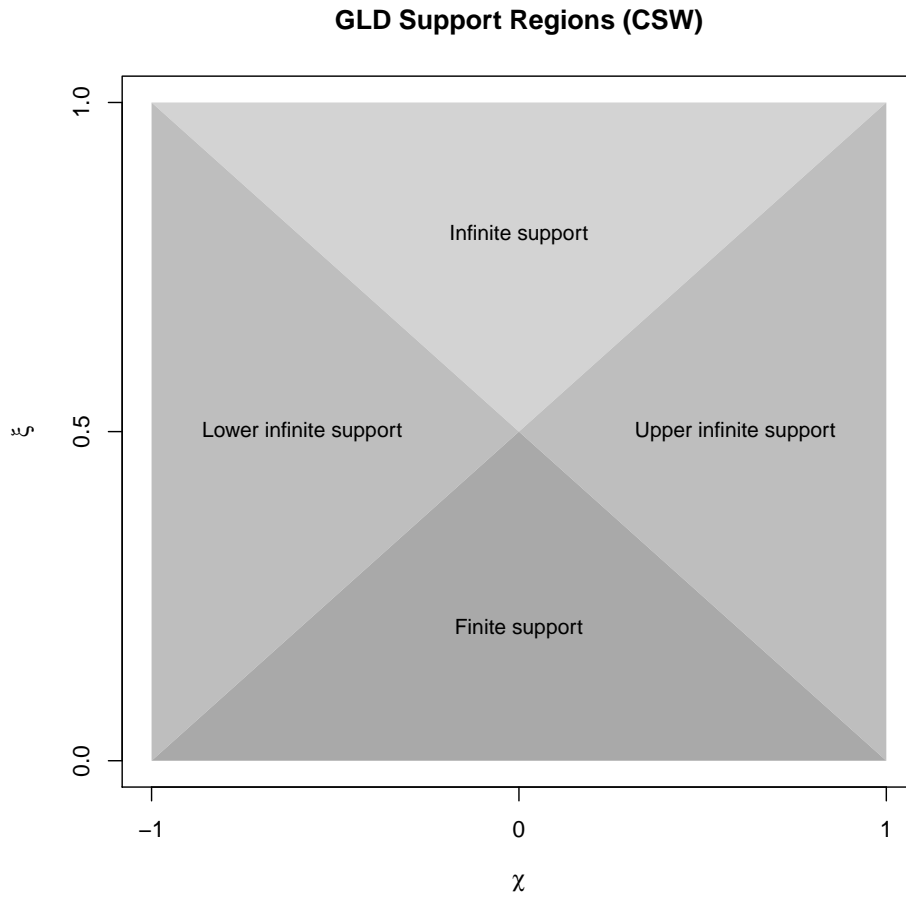


Figure 5: The four support regions of the GLD displayed in the shape plot with asymmetry parameter  $\chi$  versus the steepness parameter  $\xi$ .

where  $H$  is a discontinuous function such that

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

and  $\beta$  is defined in Eq. (7b). Note that in the limiting cases of  $\chi \rightarrow \pm 1$ , the  $k$ th moment exists when and only when  $\xi \rightarrow 0$ . Figure 6 shows the condition lines of existence for the first four moments in the shape diagram. Any set of shape parameters  $\chi$  and  $\xi$  that is under the  $k$ th condition line defines a distribution with a finite  $k$ th moment.

### 3.4 Tail behavior

The tail behavior plays an important role in the modeling of stochastic processes. This is especially the case for financial returns that have been shown to have fat tail distributions. An important class of distributions are the regularly varying tail distribution such as the Student  $t$ , Fréchet,  $\alpha$ -stable and Pareto distributions that exhibit power law decays. They are suitable for modeling the large and small values typically observed in financial data set. Moreover, power law distributions become very handy in the derivation of the asymptotic distribution of estimators such as the regression estimators encountered in econometric models [Jessen and Mikosch, 2006].

We now restrict our attention to the shape parameter range for which the GLD has infinite right tail,  $x \rightarrow \infty$ ; i.e., when  $\xi > \frac{1}{2}(1 - \chi)$  as shown in Eqs. (13) and (14). From Eq. (9),

$$x = \tilde{\mu} + \tilde{\sigma} \frac{S(F(x)|\chi, \xi) - S(\frac{1}{2}|\chi, \xi)}{S(\frac{3}{4}|\chi, \xi) - S(\frac{1}{4}|\chi, \xi)},$$

where  $S$  is defined in Eq. (8),  $\alpha$  and  $\beta$  are defined in Eq. (7). In region  $\xi > \frac{1}{2}(1 - \chi)$ ,

$$\begin{aligned} \lim_{x \rightarrow \infty} Q_{\text{CSW}}(x) &\propto (1 - F_{\text{CSW}}(x))^{\alpha - \beta} \\ &\propto \frac{\beta - \alpha}{\tilde{\sigma}} x, \end{aligned}$$

and hence

$$\frac{1 - F_{\text{CSW}}(tx)}{1 - F_{\text{CSW}}(x)} \sim t^{1/(\alpha - \beta)}.$$

The right tail of the GLD is thus regularly varying at  $+\infty$  with index  $-\frac{1}{\alpha - \beta}$  [see Embrechts et al., 1997, p. 129]. A similar argument shows that the left tail of the GLD is regularly varying at  $-\infty$  with index  $-\frac{1}{\alpha + \beta}$ .

In the shape parameter region,  $\xi > \frac{1}{2}(1 + |\chi|)$ , the left and right tail probabilities of the GLD are power laws in the asymptotic limit with index  $-\frac{1}{\alpha + \beta}$  and  $-\frac{1}{\alpha - \beta}$  respectively. Given the existence conditions of regularly varying distributions [Embrechts et al., 1997, Proposition A3.8],



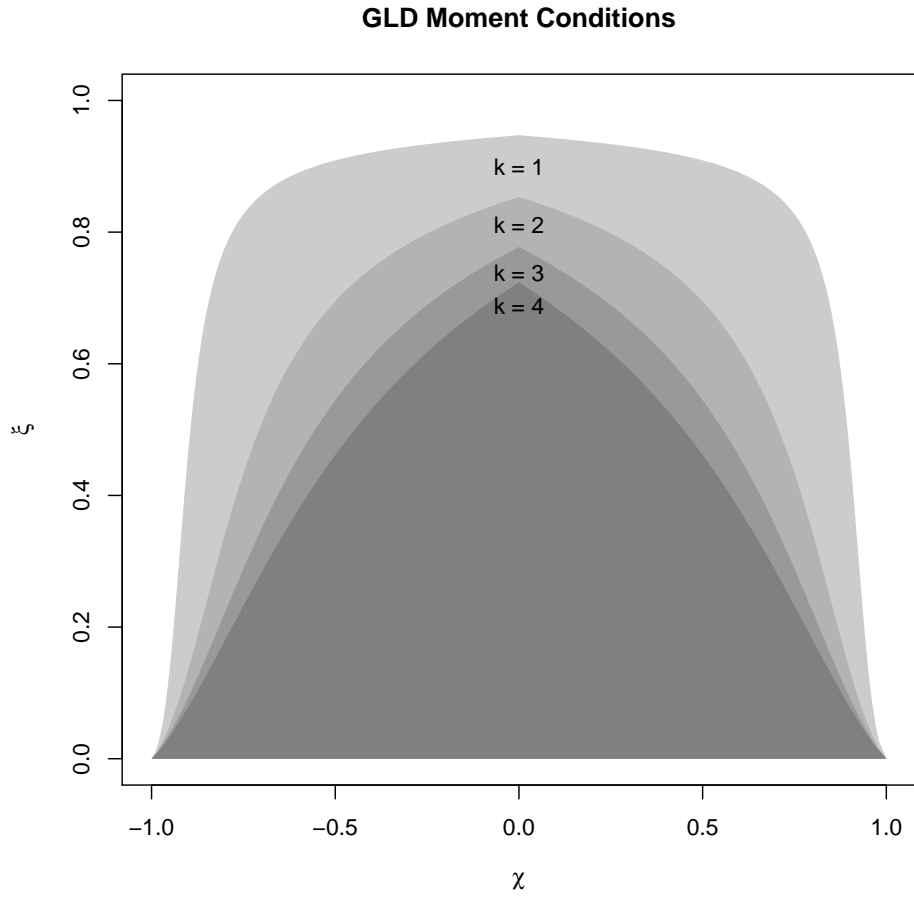


Figure 6: Shape conditions for the existence of moments  $k$ . Any set of parameters  $\chi$  and  $\xi$  that is under the  $k$ th line defines a distribution with finite  $k$ th moment.

the  $k$ th moments of the GLD is therefore infinite when  $k > \max(-\frac{1}{\alpha+\beta}, -\frac{1}{\alpha-\beta})$ . This is in agreement with the existence conditions of moments presented in Section 3.3.

Moreover, from the monotone density theorem [Embrechts et al., 1997, Theorem A3.7], the density function is

$$f(x) \propto \frac{1}{\alpha - \beta} \left[ \frac{\beta - \alpha}{\tilde{\sigma}} x \right]^{\frac{1}{\alpha - \beta} - 1} \quad \text{as } x \rightarrow \infty,$$

and with the analogous result as  $x \rightarrow -\infty$ . Hence, The decay exponent for the left and right tail of the probability density function are:

$$\begin{aligned} \tau_{\text{left}} &= \frac{1}{\alpha + \beta} - 1, \\ \tau_{\text{right}} &= \frac{1}{\alpha - \beta} - 1, \end{aligned}$$

respectively. Since the left and right tail indices are related to the shape parameters  $\chi$  and  $\xi$ , the parameter estimates can be used to form an estimate of the tail indices.

### 3.5 Distribution shape

The GLD exhibits four distribution shapes: unimodal, U-shape, monotone, and S-shape. In this section, conditions based on the derivatives of the quantile density  $q$  in Eq. (11) are derived for each distribution shape. Indeed, the density quantile function is defined as the multiplicative inverse of the quantile density function, as noted in Eq. (10). The first and second derivatives of  $q$  are

$$\frac{d}{du} q(u|\tilde{\sigma}, \chi, \xi) \propto (\alpha + \beta - 1)u^{\alpha+\beta-2} - (\alpha - \beta - 1)(1 - u)^{\alpha-\beta-2}, \quad (15a)$$

$$\begin{aligned} \frac{d^2}{du^2} q(u|\tilde{\sigma}, \chi, \xi) &\propto (\alpha + \beta - 1)(\alpha + \beta - 2)u^{\alpha+\beta-3} + \\ &(\alpha - \beta - 1)(\alpha - \beta - 2)(1 - u)^{\alpha-\beta-3}, \end{aligned} \quad (15b)$$

where  $\alpha$  and  $\beta$  are defined in Eq. (7). It is now possible to deduce the parameter conditions for each of the four shape regions of the GLD. Figure 7 summarizes the shape regions as described in the remaining part of this section.

#### 3.5.1 Unimodal

The GLD distribution is unimodal when its quantile density function is strictly convex; that is when  $\frac{d^2}{du^2} q(u) > 0$  for all  $u \in [0, 1]$ . Note that throughout this paper, a unimodal distribution refers to a distribution with a single local maximum. From Eq. (15b), the GLD density function is then unimodal when

$$\alpha + \beta - 2 > 0 \quad \text{and} \quad \alpha - \beta - 2 > 0,$$

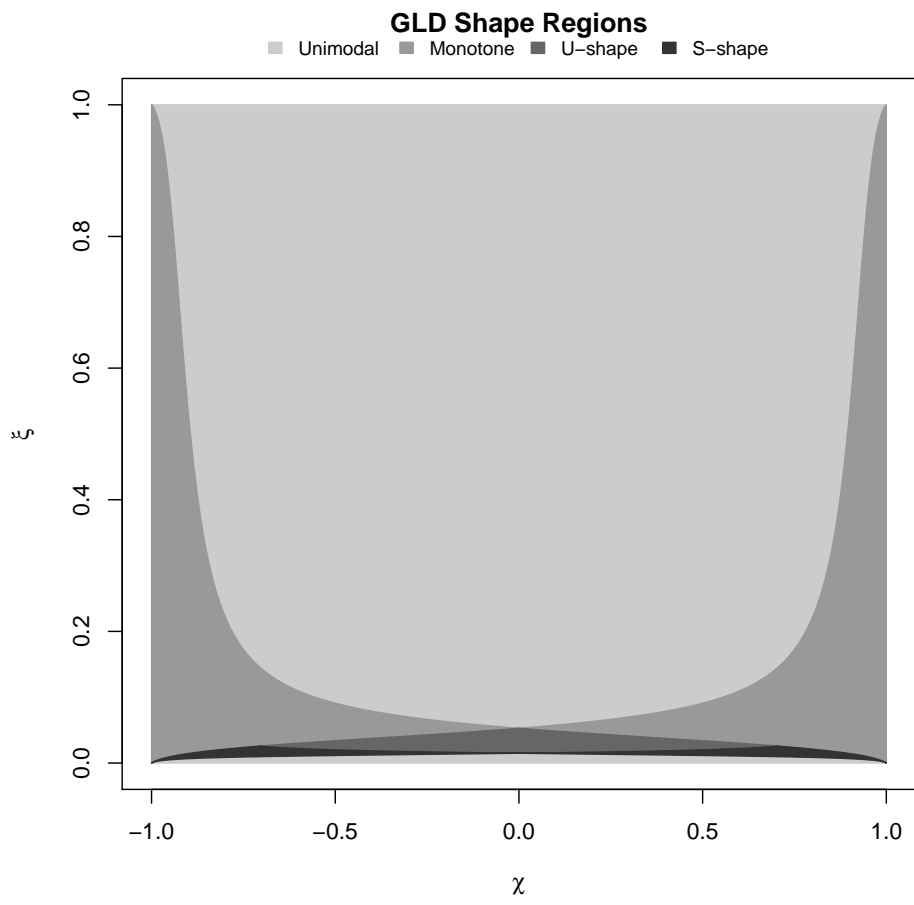


Figure 7: Parameter regions of the GLD shapes in the  $(\chi, \xi)$ -space.

or when

$$\alpha + \beta - 1 < 0 \quad \text{and} \quad \alpha - \beta - 1 < 0.$$

After some tedious but straightforward calculations, the conditions are, in terms of the shape parameters  $\chi$  and  $\xi$ ,

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

or

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

Figure 8 illustrates the unimodal distribution shapes of the GLD.

### 3.5.2 U-shape

Likewise, the GLD is U-shaped when the quantile density function is strictly concave,  $\frac{d^2}{du^2}q(u) < 0$  for all  $u \in [0, 1]$ . This is the case when

$$0 < \alpha + \beta - 1 < 1 \quad \text{and} \quad 0 < \alpha - \beta - 1 < 1,$$

giving the equivalent 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 8 illustrates U-shaped distribution shapes of the GLD.

### 3.5.3 Monotone

The probability density function is monotone when its derivative is either positive or negative for all possible values in its support range. From Eq. (15a), this is the case when

$$\alpha + \beta - 1 > 0 \quad \text{and} \quad \alpha - \beta - 1 < 0,$$

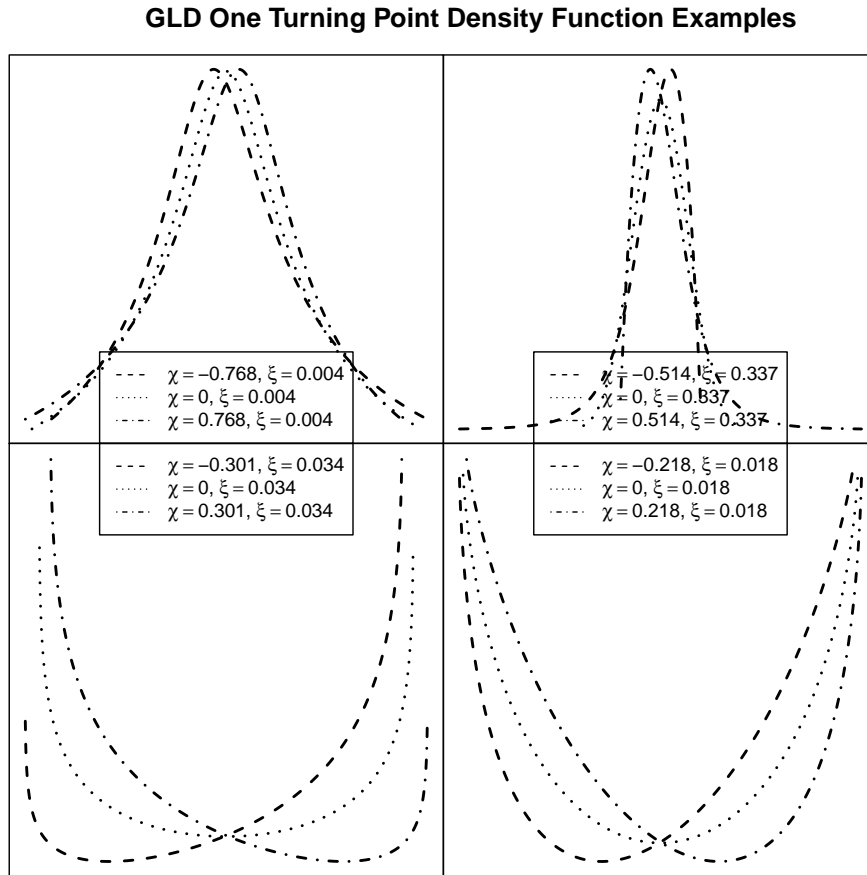


Figure 8: One turning point probability density function of the GLD with different sets of shape parameters.

or

$$\alpha + \beta - 1 < 0 \quad \text{and} \quad \alpha - \beta - 1 > 0.$$

In terms of the shape parameters, the monotone 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 9 illustrates monotone shape of the GLD with different sets of shape parameters.

#### 3.5.4 S-shape

Since  $\alpha \in \mathbb{R}$  and  $\beta \in \mathbb{R}$ , the remaining shape parameter regions are:

$$\alpha + \beta > 2 \quad \text{and} \quad 1 < \alpha - \beta < 2,$$

and

$$1 < \alpha + \beta < 2 \quad \text{and} \quad \alpha - \beta > 2.$$

Observe that the first derivative of the quantile density function at  $u = 0$  has the same sign as the first derivative of the quantile density function at  $u = 1$ . This indicates that the gradients at the two end points of the probability density function tend toward the same direction. However,

### GLD Monotone Density Function Examples

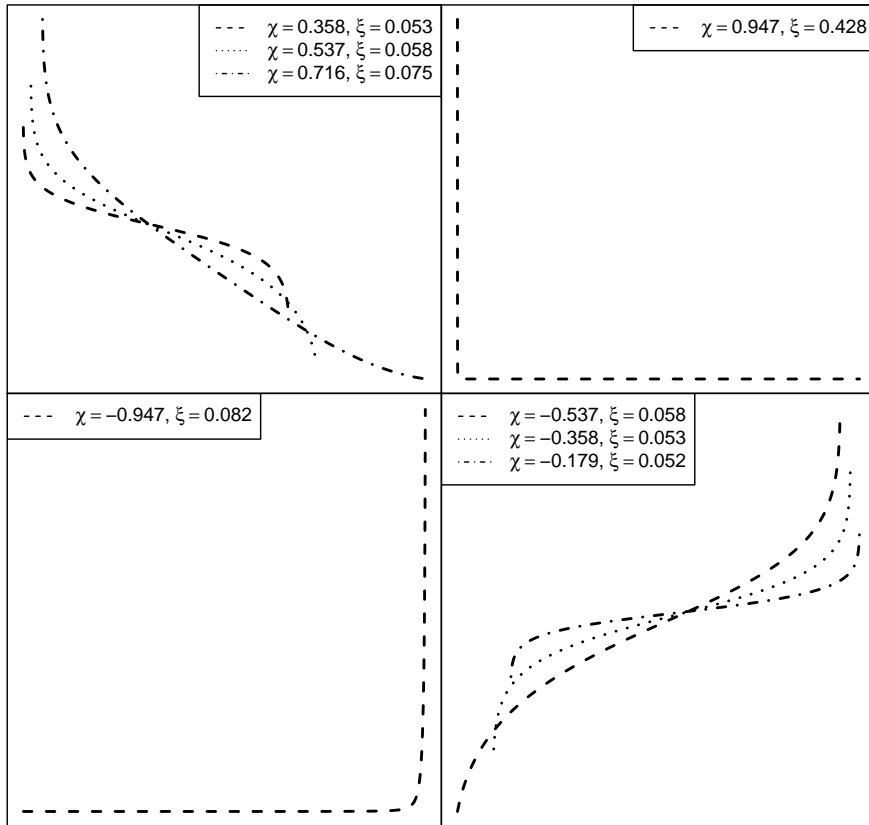


Figure 9: Monotone probability density function of the GLD with different sets of shape parameters.

the gradient of the quantile density function has not the same sign for all possible values of the distribution. The GLD therefore 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 10 illustrates S-shaped probability densities of the GLD.

**Sturges breaks** The approach of Sturges [1926] is currently the default method for calculating the number of histogram bins in the R Environment for Statistical Computing [R Core Team]. Since Sturges' formula computes bin sizes from the range of the data, it can perform quite poorly when the sample size is small in particular when,  $n$  is less than 30. For this method

$$b_n = \lceil \log_2(n + 1) \rceil.$$

The brackets represent the ceiling function, by which  $\log_2(n + 1)$  will be rounded up to the next integer value.

**Scott breaks** Scott's choice [see Scott, 1979] is derived from the normal distribution and relies upon an estimate of the standard error. It offers a more flexible approach than the fixed number of bins used by Sturges. In Scott's method, the number of bins is

$$b_n = \left\lceil \frac{\max(x) - \min(x)}{h} \right\rceil,$$



### GLD S-Shaped Density Function Examples

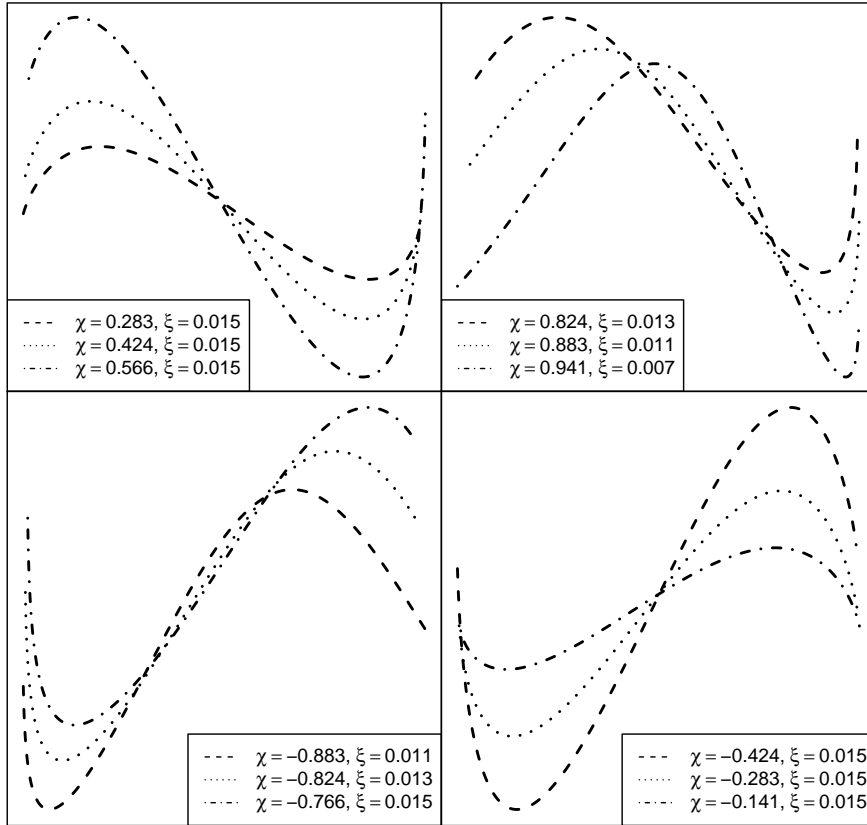


Figure 10: S-shape probability density function of the GLD with different sets of shape parameters.

where

$$h = 3.49\hat{\sigma}n^{1/3}.$$

When all data are equal,  $h = 0$ , and  $b_n = 1$ .

**Freedman–Diaconis breaks** The Freedman–Diaconis choice [Freedman and Diaconis, 1981] is a robust selection method based upon the full and interquartile ranges of the data. The number of bins is

$$b_n = \left\lceil \frac{\max(x) - \min(x)}{h} \right\rceil,$$

where

$$h = \frac{\pi(0.75) - \pi(0.25)}{n^{1/3}}.$$

If the interquartile range is zero, then  $h$  is instead set to the median absolute deviation.

The Sturges breaks do not use any information concerning the form of the underlying distribution, and so the Sturges histogram bins are often far from being optimal. The Scott approach is reliant upon the distribution being approximately normal, and so, for GLD applications, the histogram bins are not expected to be optimally chosen over the tails. Since it is a robust method based on the interquartile range, the Freedman–Diaconis approach seems to be the most promising option for creating histograms from random variates drawn from the heavy-tailed GLD.

### 3.6 Special cases

As seen in the previous section, the GLD can assume a wide range of distribution shapes. Here, values of the parameters in Eq. (9) are found such that  $Q_{CSW}$  replicates several common distributions. The method used to accomplish this is as follows. Taking a set of equidistant probabilities,  $p_i = \frac{i}{N+1}$ , with  $i = 1, \dots, N$  and  $N = 500$ , the respective quantiles,  $x_i$ , and densities,  $d_i = f(x_i)$ , are calculated for each of the target distributions. The shape parameters,  $\hat{\chi}$  and  $\hat{\xi}$ , are then fitted, by minimizing the maximum absolute quantile error (MQE):

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

The median and interquartile range of the target distributions are used for the location and scale parameters,  $\tilde{\mu}$  and  $\tilde{\sigma}$ . Note that  $N = 500$  was explicitly chosen in order to compare the results of this work with those of the previous studies of Gilchrist [2000], King and MacGillivray [2007] and Tarsitano [2010].

Table 2 lists the fitted shape parameters for some common distributions along with the maximum MQE appearing in the  $\sup |\hat{Q}|$  column. Also reported are the maximum probability

Distribution	Parameters	$\hat{\chi}$	$\hat{\xi}$	$\sup \hat{Q} $	$\sup \hat{F} $	$\sup \hat{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
	$\nu = 10$	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
$\chi^2$	$k = 3$	0.6671	0.1991	0.295	0.015	0.072
	$k = 5$	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.

error,  $\sup|\hat{F}|$ , and the maximum density error,  $\sup|\hat{f}|$ . The maximum probability error (MPE) is defined as

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

The maximum density error (MDE) is defined as

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

An important distinction between this method and those of previous studies, is that here the appropriate values of the location and scale parameters, being identified with the median and interquartile range of the target distribution, are known. In previous works, the location and scale parameters had to be estimated alongside the tail parameters. In so doing, the fitted parameters can sometimes produce a GLD that does not well approximate the target distribution around its center. This problem arises from adjustment of the location and scale parameter values to improve a poor fit over the tails of the distribution. This is at the expense of having a worse fit around the center. Figure 11 demonstrates this issue with the Student  $t$ -distribution of two degrees of freedom. The left-hand-side plot displays the fitted GLD obtained using the exact values for  $\tilde{\mu}$  and  $\tilde{\sigma}$ . The right-hand-side plot shows the fitted GLD obtained when including the location and scale parameters in the inexact estimation. Clearly, the center of the distribution

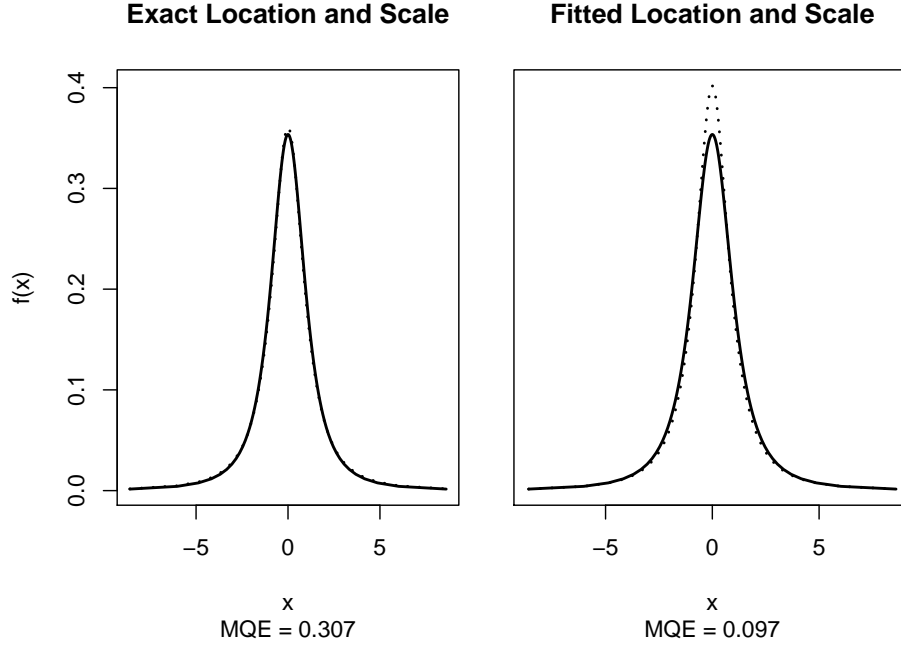


Figure 11: Approximation of the Student  $t$ -distribution (dotted lines) with two degrees of freedom by either fitting or using the true values for the location and scale parameters,  $\tilde{\mu}$  and  $\tilde{\sigma}$ . In the left-hand-side figure, the fitted GLD where the true values for the median and interquartile range were used. In the right-hand-side figure, 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.

is not well described by the latter solution. Using the known values for  $\tilde{\mu}$  and  $\tilde{\sigma}$  yields a greater MQE value (of 0.307, in this case) than that of the alternate approach (for which, the MQE value was 0.097).

It is interesting to note that the MPE could have been used in place of the MQE while fitting the shape parameters. This would have resulted in a smaller MPE value. However, in practice, GLDs fitted according to the MPE estimator do not well approximate the tails of the target distribution. This is especially the case with fat-tailed distributions. Figure 12 illustrates this tendency. The left-hand-side plot displays the fitted log-CDF obtained by minimizing the MQE for the Student  $t$ -distribution of two degrees of freedom. The right-hand-side plot is the fitted log-CDF obtained with the MPE estimator. Note that the true median and interquartile range were used in both cases. The parameter-value set obtained with the MQE has produced the better visual fit.

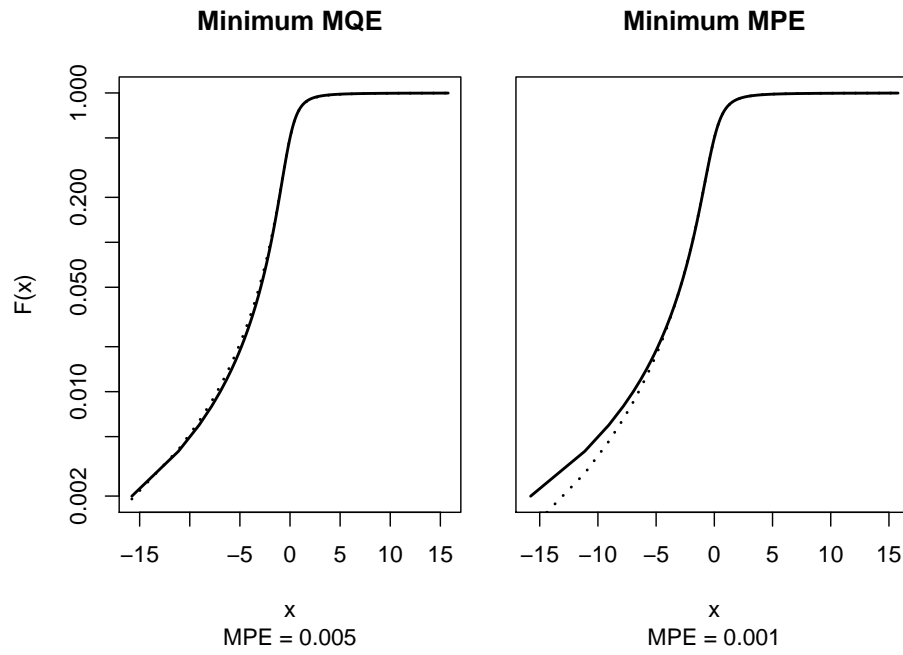


Figure 12: Comparison of the fitted GLD (dotted lines) to the Student  $t$ -distribution with two degrees of freedom by either using the MQE or the MPE estimator. The left-hand-side figure was obtained by minimizing the MQE and the right-hand-side plot with the MPE. The MPE for the left hand side plot, 0.005, is larger than the one for the right hand side figure, 0.001, although it has a better visual fit.

	$\tilde{\mu}$	$\tilde{\sigma}$	$\chi$	$\xi$
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 $(\mu, \beta)$	$\mu$	$\beta \ln(9)$	0	$\frac{1}{2}$
Exponential $(\lambda)$	$\frac{1}{\lambda} \ln(2)$	$\frac{1}{\lambda} \ln(3)$	1	0

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

Another important consideration in our approach is to ensure that the fitted parameters yield a GLD with a support that includes all  $x_i$ . This is especially the case for the gamma,  $\chi^2$  and Wald distributions. Indeed, it is possible to find parameter-value sets for which the MQE is small, but where the fitted distribution does not include all  $x_i$ . We therefore add constraints in the optimization routine to ensure that all  $x_i$  are included in the fitted GLD support range.

These considerations motivated the decisions 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.

The GLD includes the uniform, logistic, and exponential distributions as special cases. The corresponding GLD parameter values, obtained from Eqs. (8), (9) and (12), are summarized in Table 3. Note that the exponential distribution corresponds to the limiting case  $\{\chi \rightarrow 1, \xi \rightarrow 0\}$ , as mentioned at the end of Section 3.1.

## 4 Parameter estimation

Estimating parameter values for the generalized lambda distribution (GLD) is notoriously difficult. This is due to abrupt distribution shape changes that accompany variations of the parameter values in the different shape regions (Fig. 2). In this section, an extensive Monte-Carlo study of different GLD estimators is presented. This includes maximum log-likelihood, maximum product of spacing, goodness-of-fit, histogram binning, quantile sequence, and linear moments. Besides these estimators, we introduce the maximum product of spacing and a robust moment matching to estimate the shape parameters of the GLD.

### 4.1 Robust moments matching

A robust moment approach, for the estimation of the distribution parameters, is presented. The method is based on the median, interquartile range, Bowley's skewness and Moors' kurtosis. It yields estimators for the two shape parameters, which are independent of the location and scale parameters. A similar approach has been introduced by King and MacGillivray [2007].

The method suggested by Ramberg et al. [1979] for fitting GLD distributions, uses a method of moments wherein the first four moments are matched. Clearly, this approach can only be applied to regions within the first four moments exist. In addition, parameter estimates based upon sample moments are highly sensitive to outliers. This is especially true of third and fourth moment. To circumvent the estimation of sample moments, Karian and Dudewicz [1999] considered a quantile approach that estimates the parameters from four sample quantile statistics. In that approach, the four statistics depend upon the four parameters of the GLD, thereby leading to a system of four nonlinear equations to be solved.

In the CSW parameterization, because the location and scale parameters are, respectively, the median and interquartile range, they can be directly estimated by their sample estimators. What remains to be estimated, are two shape parameters;  $\chi$  and  $\xi$ . This is achieved by using the robust skewness ratio  $\tilde{s}$  of Bowley [1920], and the robust kurtosis ratio  $\tilde{\kappa}$  of Moors [1988]. These robust moments are defined as

$$\tilde{s} = \frac{\pi_{3/4} + \pi_{1/4} - 2\pi_{2/4}}{\pi_{3/4} - \pi_{1/4}},$$

$$\tilde{\kappa} = \frac{\pi_{7/8} - \pi_{5/8} + \pi_{3/8} - \pi_{1/8}}{\pi_{6/8} - \pi_{2/8}},$$

where  $\pi_q$  indicates the  $q$ th quantile, and the tilde indicates the robust versions of the moments. For a detailed discussion of Bowley's skewness and Moors' kurtosis statistics, please refer to [Kim and White, 2004].

Recall that the quantile function of the GLD in the CSW parameterization is

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

as seen in Eq. (9) where  $S$  is defined in Eq. (8). The population robust skewness and robust kurtosis thus depend only upon the shape parameters. Explicitly,

$$\tilde{s} = \frac{S(3/4|\chi, \xi) + S(1/4|\chi, \xi) - 2S(2/4|\chi, \xi)}{S(3/4|\chi, \xi) - S(1/4|\chi, \xi)}, \quad (16a)$$

$$\tilde{\kappa} = \frac{S(7/8|\chi, \xi) - S(5/8|\chi, \xi) + S(3/8|\chi, \xi) - S(1/8|\chi, \xi)}{S(6/8|\chi, \xi) - S(2/8|\chi, \xi)}. \quad (16b)$$

Estimates of the shape parameters,  $\chi$  and  $\xi$ , can then be calculated by solving the nonlinear system of equations in Eq. (16). The advantage of this method, wherein the GLD shape parameters are readily obtained from quantile-based estimators, is that it reduces the nonlinear system from four to two unknowns. As a secondary application, this simplified two-equation system can be efficiently solved by a lookup method in order to obtain good initial estimates of the shape parameters for use by other fitting methods that are sensitive to initial values and might otherwise get trapped in local minima.

## 4.2 Histogram approach

The histogram approach is appealing and simple; the empirical data are binned in a histogram and the resulting probabilities, taken to be at the midpoints of the histogram bins, are fitted to the true GLD density. This approach was considered by Su [2005] for the GLD. To obtain the best estimates, it is vital to choose an appropriate number of bins for the histogram. Three different methods of estimating the number of bins were investigated. In the following discussions of these,  $n$  denotes the sample size, and  $b_n$  denotes the number of bins for that sample size.

## 4.3 Goodness-of-fit approaches

To test the goodness-of-fit of a continuous cumulative distribution function, to a set of data, statistics based on the empirical distribution function can be used. Examples of these are the Kolmogorov–Smirnov, Cramér–von Mises, and Anderson–Darling statistics. These statistics are measures of the difference between the hypothetical GLD distribution and the empirical distribution.

To use these statistics as parameter estimators, they are minimized with respect to the unknown parameters of the distribution. It is known, from the generalized Pareto distribution, that these approaches are able to successfully estimate the parameters even when the maximum likelihood method fails. This is shown by Luceño [2006], who calls this approach the method of maximum goodness-of-fit. Definitions and computational forms of the three goodness-of-fit statistics named above are also given in the paper of Luceño [2006].

The following discussions consider a random sample of  $n$  *iid* observations,  $\{x_1, x_2, \dots, x_n\}$ , with the order statistics of the sample denoted by  $\{x_{(1)}, x_{(2)}, \dots, x_{(n)}\}$ , and write  $F_i = F(x_{(i)})$  the corresponding probability of the event  $x_{(1)}$ .

### 4.3.1 The Kolmogorov–Smirnov statistic

The Kolmogorov–Smirnov statistic measures the maximum distance between the empirical cumulative distribution function and the theoretical probabilities,  $S_n$ ;

$$D_n = \sup_x |F_x - S_n(x)|,$$

where the sample estimator is

$$\widehat{D}_n = \frac{1}{2n} + \max_{1 \leq i \leq n} \left| F_i - \frac{1 - i/2}{n} \right|.$$

Minimizing this leads to the optimization of a discontinuous objective function.



#### 4.3.2 The Cramér–von Mises statistic

The Cramér–von Mises statistic uses mean-squared differences, thereby leading to a continuous objective function for parameter optimization:

$$W_n^2 = n \int_{-\infty}^{\infty} [F_x - S_n(x)]^2 dF(x),$$

which reduces to

$$\widehat{W_n^2} = \frac{1}{12n} + \sum_{i=1}^n \left( F_i - \frac{1 - i/2}{n} \right)^2.$$

#### 4.3.3 The Anderson–Darling statistic

The Anderson–Darling statistic is a tail-weighted statistic. It gives more weight to the tails and less weight to the center of the distribution.

This makes the Anderson–Darling statistic an interesting candidate for estimating parameters of the heavy-tailed GLD. It is defined as

$$A_n^2 = n \int_{-\infty}^{\infty} \frac{[F_x - S_n(x)]^2}{F(x)(1 - F(x))} dF(x),$$

where its sample estimator is

$$\widehat{A_n^2} = -n - \frac{1}{n} \sum_{i=1}^n (2i - 1)(\ln(F_i) + \ln(1 - F_{n+1-i})).$$

### 4.4 Quantile matching

As described by Su [2010], the quantile matching method consists of finding the parameter values that minimize the difference between the theoretical and sample quantiles. This approach is especially interesting for the GLD, since the distribution is defined by a quantile function.

Consider an indexed set of probabilities,  $p$ , as a sequence of values in the range  $0 < p_i < 1$ . The quantile matching estimator is defined For a cumulative probability distribution  $F$  with a set of parameters  $\theta$ , the quantile matching estimator yield the parameter values

$$\hat{\theta} = \arg \min_{\theta} \frac{1}{n} \sum_i^n (F^{-1}(p_i|\theta) - Q(p_i))^2,$$

where  $n$  is the cardinality of  $p$ ,  $F^{-1}(p|\theta)$  is the quantile function, and  $Q$  the sample quantile function.

### 4.5 Trimmed L-moments

A drawback of using moment matching to fit a distribution, is that sample moment estimators are sensitive to outliers. To circumvent this issue, the L-moments have been defined as linear

combinations of the ordered data values. The method of L-moments is to find the distribution parameter values that minimize the difference between the sample L-moments and the theoretical ones. Elamir and Seheult [2003] introduced the trimmed L-moments, a robust extension of the L-moments.

The trimmed L-moments,  $l_r(t_1, t_2)$ , on the order statistics  $X_{1:n}, \dots, X_{n:n}$  of a random sample  $\mathbf{X}$  of size  $n$  with trimming parameters  $t_1$  and  $t_2$ , are defined as

$$TL(r, t_1, t_2) = \frac{1}{r} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} E[r + t_1 - k, r + t_1 + t_2], \quad (17)$$

where

$$E[i, r] = \frac{r!}{(i-1)!(r-i)!} \int_0^1 Q(u) u^{i-1} (1-u)^{r-i} du.$$

$Q$  is the quantile function of the distribution. Note that when  $t_1 = t_2 = 0$ , this reduces to the definition of the standard L-moment. The concept behind the trimmed L-moments is to extend the sample size by  $t_1 + t_2$  and trim the  $t_1$  smallest and  $t_2$  largest order statistics.

Using the unbiased estimator of the expectation of the order statistics,  $E$ , of Downton [1966], Elamir and Seheult [2003] obtains the sample estimator  $\hat{E}$ :

$$\hat{E}(i, r) = \frac{1}{\binom{n}{r}} \sum_{t=1}^n \binom{t-1}{i-1} \binom{n-t}{r-i} X_{t:n},$$

which can be used to estimate the sample's  $r$ th trimmed L-moment in Eq. (17).

The derivation of the trimmed L-moments for the GLD is cumbersome. Asquith [2007] calculated them for the RS parameterization.

## 4.6 MLE and MPS approaches

The Kullback–Leibler (KL) divergence [Kullback, 1959], also known by the names information divergence, information gain, and relative entropy, measures the difference between two probability distributions:  $P$  and  $Q$ . Typically,  $P$  represents the empirical distribution, and  $Q$  comes from a theoretical model. The KL divergence allows for the interpretation of many other measures in information theory. Two of them are the maximum log-likelihood estimator and the maximum product of spacing estimator.

### 4.6.1 Maximum log-likelihood estimation

The maximum log-likelihood method was introduced by Fisher [1922]. For a more recent review, please refer to Aldrich [1997]. Consider a random sample of  $n$  *iid* observations,  $\{x_1, x_2, \dots, x_n\}$  drawn from the probability distribution with parameter set  $\theta$ . Then the maximum value of the log-likelihood function,  $\mathcal{L}$ , defined by

$$\mathcal{L}(\theta) = \sum_{i=1}^n \ln f(x_i|\theta),$$

returns optimal estimates for the parameters  $\theta$  of the probability density function  $f$ . The maximum of this expression can be found numerically using a non-linear optimizer that allows for linear constraints.

#### 4.6.2 Maximum product of spacing estimation

The maximum likelihood method may break down in certain circumstances. An example of this is when the support depends upon the parameters to be estimated [see Cheng and Amin, 1983]. Under such circumstances, the maximum product of spacing estimation (MPS), introduced separately by Cheng and Amin [1983]; Ranneby [1984], may be more successful. This method is derived from the probability integral transform, wherein the random variates of any continuous distribution can be transformed to variables of a uniform distribution. The parameter estimates are the values that make the spacings between the cumulative distribution functions of the random variates equal. This condition is equivalent to maximizing the geometric means of the probability spacings.

Consider the order statistics,  $\{x_{(1)}, x_{(2)}, \dots, x_{(n)}\}$ , of the observations  $\{x_1, x_2, \dots, x_n\}$ , and compute their spacings or gaps at adjacent points,

$$D_i = F(x_{(i)}) - F(x_{(i-1)}),$$

where  $F(x_{(0)}) = 0$  and  $F(x_{(n+1)}) = 1$ . The estimate of the maximum product of spacing becomes

$$\hat{\theta} = \arg \max_{\theta \in \Theta} \sum_{i=1}^{n+1} \ln D(x_{(i)}).$$

The maximizing solution can be found numerically with a non-linear optimizer.

Note that the concept underlying the MPS method is similar to that of the starship estimation GLD estimation method introduced by King and MacGillivray [1999]. However, the starship method uses the Anderson–Darling test to find the parameters that make the spacings closest to a uniform distribution.

### 4.7 Empirical study

In this section, we compare the aforementioned estimators. In this study we specifically take: the Anderson–Darling statistic for the goodness-of-fit approach, the Friedman–Draconis breaks for the histogram approach, the L-moments estimator, the MLE, the MPS estimator, the quantile matching estimator, the robust moment matching and the trimmed L-moments estimator with  $t_1 = 1$  and  $t_2 = 1$ . All methods have been implemented for the GLD in the statistical software environment **R** [R Core Team] and are available within the package **gldist**.

Figure 13 shows the fitted log-density of the GLD for the Google equity log-returns with the different methods. Overall, The estimates differ by small amounts in the lower tail. However, the upper tail fit of the histogram approach is substantially different from the other methods.

To compare the estimation methods, we generated one thousand time series of one thousand points each, for each of the set GLD parameter sets:  $\{\tilde{\mu} = 0, \tilde{\sigma} = 1, \chi = 0, \xi = 0.35\}$ . We perform both approaches, when the location and scale parameters,  $(\chi, \xi)$ , are estimated by their sample estimators, and when they are included in the estimator. Table 4 shows the median and interquartile range of the minimum discrimination information criterion (MDE). The MDE is a divergence between two distributions measure used in information statistics. The MDE between two continuous distributions,  $f$  and  $g$ , is defined as

$$D_{\phi}(f, g) = \int_{-\infty}^{\infty} \phi\left(\frac{f(x)}{g(x)}\right) g(x) dx,$$

where

$$\phi(x) = -\log(x) - x + 1.$$

A divergence measure is used to evaluate how closely the estimator can fit the distribution. This is in contrast to the traditional mean absolute deviation and mean absolute error of the parameter estimates, which focus on how closely the estimator can fit the parameters. The distinction is important because, although the fitted parameters might substantially differ from those used to generate the time series, their combination can produce a fitted distribution that closely resembles the generating distribution. Table 4 also shows how many of the estimation routines have successfully converged to a minimum, and the median computation times. From the results, the performance of the estimators in terms of the MDE is not affected by estimating the parameter values of the GLD using the two-step procedure: the location and scale parameter estimated by their sample estimates, and the two shape parameter values fitted by the estimator into consideration. However, the computation times of convergence are much shorter in the former case. This is expected, since the optimization routine only seeks solutions for the shape parameters, leading to a simpler optimization path. This result can be useful when dealing with a large data set, in which case the computation time is commonly critical.

## 5 Applications

### 5.1 Shape plot representation

A nice consequence of having a closed domain of variation for the shape parameters  $\chi$  and  $\xi$ , is that they can be represented by a shape plot, such as. Figure 14 illustrates the different shapes associated with location parameter  $\tilde{\mu} = 0$  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

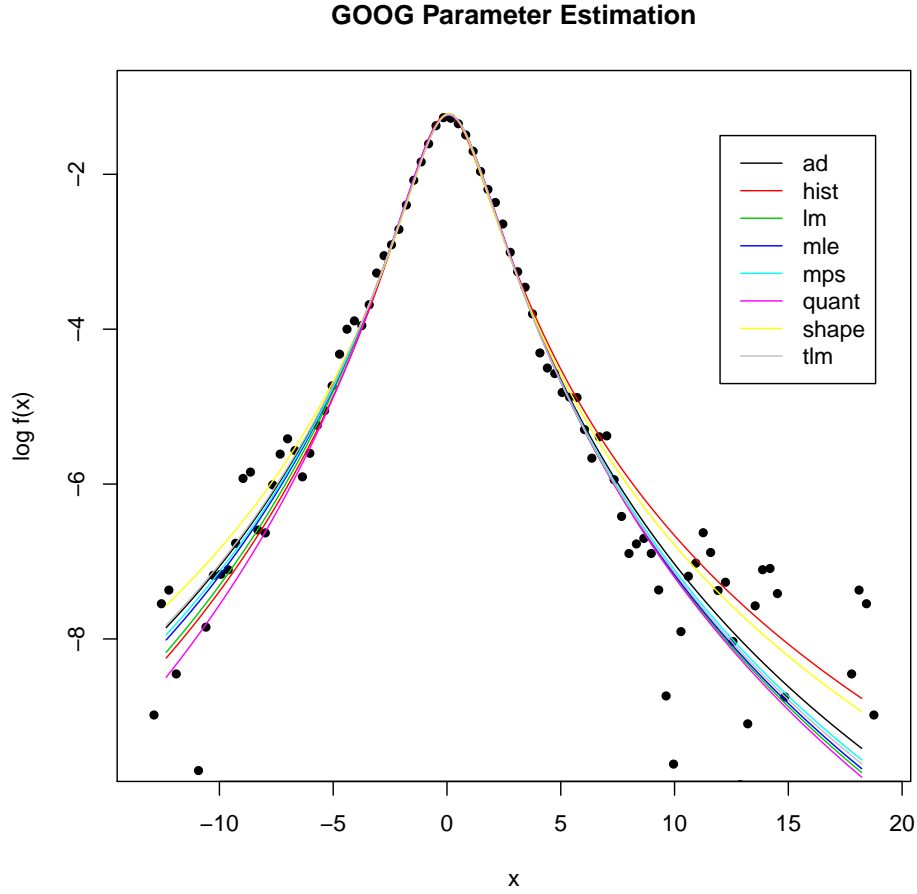


Figure 13: Parameter estimation for the Google equity. The full lines are drawn from the fitted distribution function and the points are taken from a kernel density estimate of the time series. The methods compared are: the Anderson–Darling goodness-of-fit method (ad), the histogram binning method (hist), the linear moments method (lm), the maximum likelihood estimator (mle), the maximum product of spacing estimator (mps), the quantile sequence method (quant), the robust moment matching method (shape), and the trimmed linear moments estimator (tlm) with  $t_1 = t_2 = 1$ .

	sample $(\tilde{\mu}, \tilde{\sigma})$			optim $(\tilde{\mu}, \tilde{\sigma})$		
	MDE	conv.	time [s]	MDE	conv.	time [s]
mle	0.0022 (2.50E-03)	100%	0.15	0.0019 (1.94E-03)	100%	0.43
hist	0.0046 (5.29E-03)	100%	0.01	0.0054 (5.85E-03)	100%	0.02
prob	0.0069 (7.49E-03)	100%	0.14	0.0048 (5.55E-03)	100%	0.34
quant	0.0030 (3.74E-03)	100%	0.02	0.0021 (2.50E-03)	100%	0.03
shape	0.0090 (1.44E-02)	100%	0.01	0.0090 (1.44E-02)	100%	0.01
ad	0.0035 (3.84E-03)	100%	0.14	0.0026 (2.50E-03)	100%	0.36
mps	0.0022 (2.21E-03)	100%	0.10	0.0018 (1.76E-03)	100%	0.19

(a) With no outliers

	sample $(\tilde{\mu}, \tilde{\sigma})$			optim $(\tilde{\mu}, \tilde{\sigma})$		
	MDE	conv.	time [s]	MDE	conv.	time [s]
mle	0.0082 (6.64E-03)	100%	0.16	0.0073 (5.70E-03)	100%	0.37
hist	0.0070 (1.10E-02)	100%	0.01	0.0071 (6.40E-03)	100%	0.02
prob	0.0069 (7.14E-03)	100%	0.14	0.0051 (5.27E-03)	100%	0.35
quant	0.0047 (4.77E-03)	100%	0.01	0.0045 (5.18E-03)	100%	0.03
shape	0.0086 (1.33E-02)	100%	0.01	0.0086 (1.33E-02)	100%	0.01
ad	0.0055 (4.88E-03)	96%	0.20	0.0042 (3.20E-03)	91%	0.40
mps	0.0101 (7.25E-03)	100%	0.09	0.0088 (6.34E-03)	100%	0.19

(b) With 1% of outliers with scale  $d = 3$

	sample $(\tilde{\mu}, \tilde{\sigma})$			optim $(\tilde{\mu}, \tilde{\sigma})$		
	MDE	conv.	time [s]	MDE	conv.	time [s]
mle	0.1391 (1.41E-02)	100%	0.14	0.1300 (1.38E-02)	100%	0.48
hist	0.1968 (8.52E-03)	100%	0.01	0.8170 (2.27E-01)	100%	0.04
prob	0.0585 (1.99E-02)	100%	0.14	0.0578 (1.92E-02)	100%	0.40
quant	0.1259 (8.93E-03)	100%	0.02	0.2061 (5.00E-02)	100%	0.03
shape	0.0400 (2.36E-02)	100%	0.01	0.0400 (2.36E-02)	100%	0.01
ad	0.1174 (1.78E-02)	100%	0.12	0.1155 (1.77E-02)	100%	0.34
mps	0.1418 (1.39E-02)	100%	0.12	0.1329 (1.36E-02)	100%	0.23

(c) With 10% of outliers with scale  $d = 10$

Table 4: Simulation study of the estimators. Median and interquartile of the minimum discrimination information criterion (MDE) is reported. The convergence success rates of the optimization routines, as well as the median computation times are reported. The simulated samples are of length 1000. The number of Monte-Carlo replication is 1000. The set of GLD parameters used to simulate the data are,  $\{\tilde{\mu} = 0, \tilde{\sigma} = 1, \chi = 0, \xi = 0.35\}$

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 5: NASDAQ Symbols. The 66 components of the NASDAQ-100 index that have records from 2001–01–03 to 2011–12–31.

the tails. The shape plot is thus ideal for comparing the fitted shape parameters of a data set between different time series. This section illustrates the use of the shape plot with shape parameters fitted to equities from the NASDAQ-100 index.

The NASDAQ-100 index comprises 100 of the largest US domestic and 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 and wholesale trade, and biotechnology. The listed equity returns are expected to exhibit a wide range of distribution shapes. The GLD is therefore a good candidate for modeling their various distributions. The NASDAQ-100 financial index series financial index series used in this study has records from 2000–01–03 to 2011–12–31<sup>1</sup>. These are listed in Table 5. The log-returns of the adjusted closing prices were used. Note that the survivor bias [Elton et al., 1996] has not been considered in this experiment, because the objective is merely to illustrate the use of the GLD with real time series data.

First, the location and scale parameters were estimated from their sample estimators. The maximum likelihood estimator is then used to fit the shape parameters,  $\chi$  and  $\xi$ . Figure 15 shows the fitted shape parameters. It is interesting to note that the fitted parameters are to the symmetric vertical line at  $\chi = 0$ . However, the fitted shape parameters have values that are well above those that best describe the standard normal distribution; these are represented by a triangle in the shape plot. The GLDs fitted to the NASDAQ data have “fatter” tails than does the normal distribution. This is one of the so-called stylized facts of financial returns.

This example is important because it illustrates how the GLD can model time series with tails that are fatter than those of the normal distribution. The ability of the GLD to model time series with different types of tails could be used in an assets selection process. For example, assets could be classified according to their kurtosis order, which is along the vertical access of the GLD shape plot.

<sup>1</sup>Data downloaded from [finance.yahoo.com](http://finance.yahoo.com)

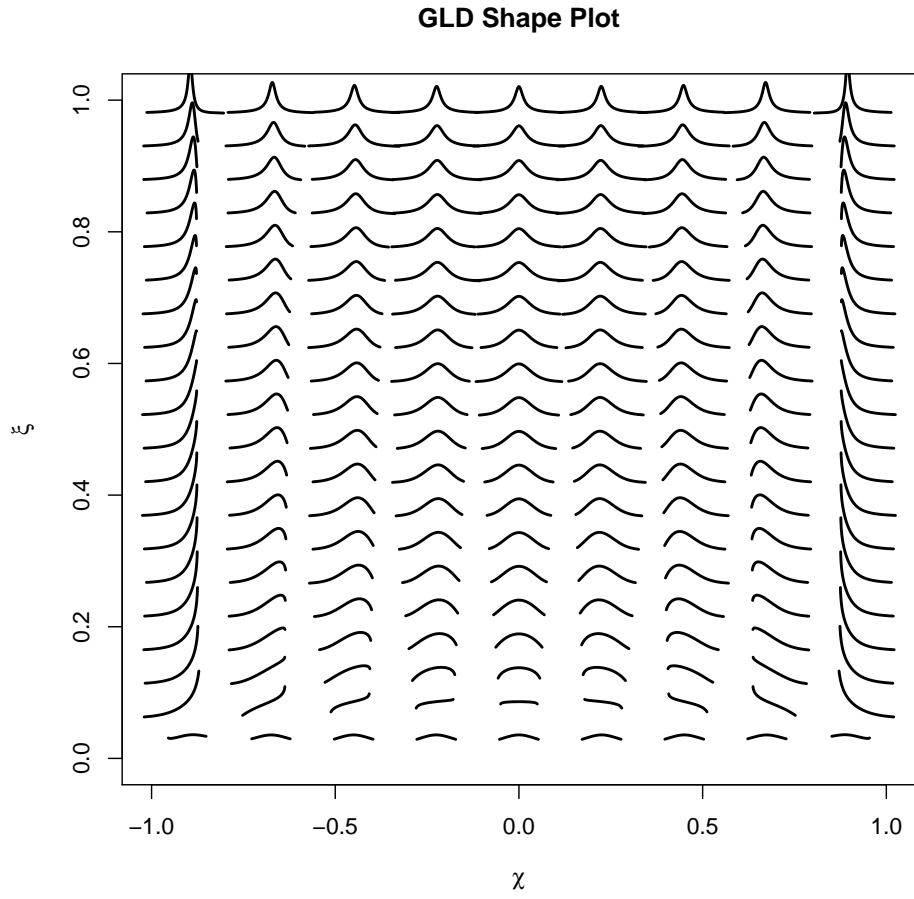


Figure 14: This figure illustrates the different probability density shapes for various steepness and asymmetry parameters. The location and scale of the distribution are, respectively,  $\tilde{\mu} = 0$  and  $\tilde{\sigma} = 1$ .



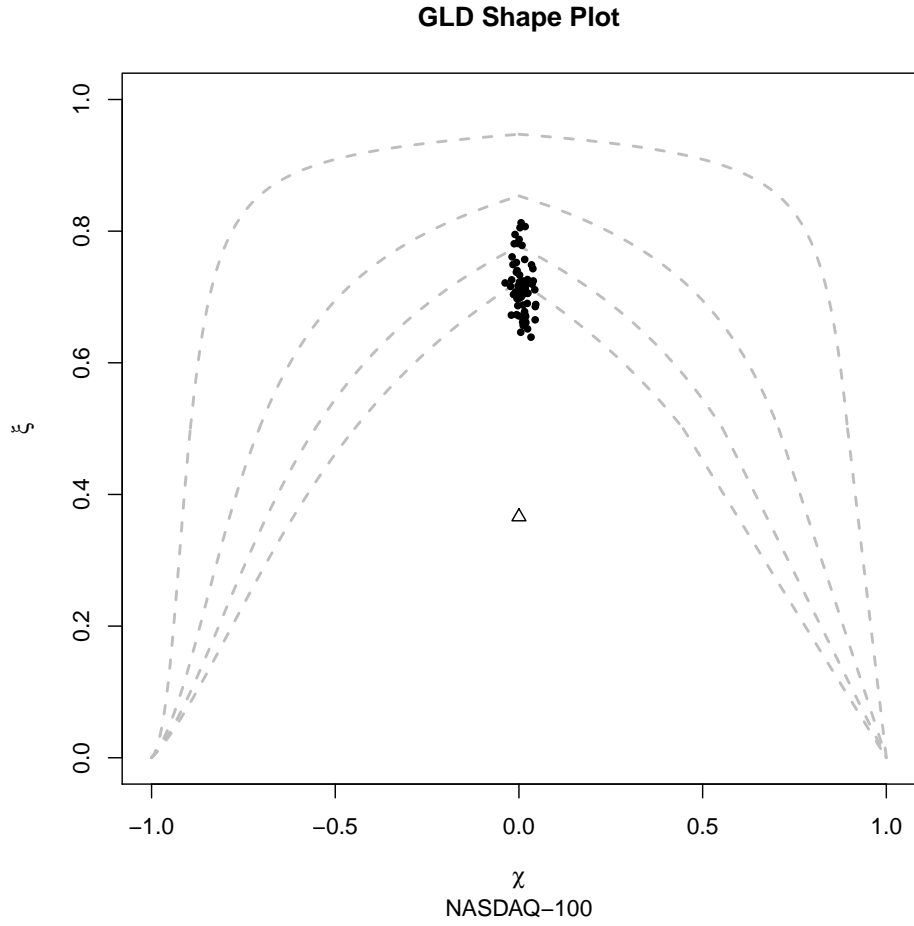


Figure 15: Shape plot of the fitted shape parameters of the 66 components of the NASDAQ-100 Index (Table 5). The dashed lines represent the existence conditions of the moments of the GLD starting from the existence of the first moment from the top to the existence of higher moments. The triangle symbol corresponds to the GLD shape parameters that best approximate the standard normal distribution.

## 5.2 Quantile based risk measures

Two risk measures commonly used in portfolio optimization are the value-at-risk (VaR) and the expected shortfall risk (ES). On one hand, the  $\text{VaR}_\gamma$  is the maximum loss forecast that may happen with probability  $\gamma \in [0, 1]$  given a holding period. On the other hand, the  $\text{ES}_\gamma$  is the averaged  $\text{VaR}_\gamma$  in the confidence interval  $[0, \gamma]$ . These risk measures are thus related to the quantiles of the observed equity price sample distribution. Since the quantile function takes a simple algebraic form, these two risk measures are easily calculated:

$$\text{VaR}_\gamma = Q_{\text{CSW}}(\gamma | \tilde{\mu}, \tilde{\sigma}, \chi, \xi),$$

and

$$\begin{aligned} \text{ES}_\alpha &= \frac{1}{\gamma} \int_0^\gamma \text{VaR}_\nu d\nu \\ &= \frac{1}{\gamma} \int_0^\gamma Q_{\text{CSW}}(\nu | \tilde{\mu}, \tilde{\sigma}, \chi, \xi) d\nu \\ &= \begin{cases} \gamma(B + \tilde{\mu} + A \ln \gamma) + (A - A\gamma) \ln(1 - \gamma) & \text{if } \chi = 0, \xi = \frac{1}{2}, \\ \gamma \tilde{\mu} + A\gamma \left( \frac{1}{2\alpha} + \ln \gamma - 1 \right) + A \frac{(1-\gamma)^{1+2\alpha}-1}{2\alpha+4\alpha^2} + B\gamma & \text{if } \chi \neq 0, \xi = \frac{1}{2}(1 + \chi), \\ \frac{A\gamma(-1+4\beta^2+\gamma^{2\beta})}{2\beta(1+2\beta)} + \gamma(B + \tilde{\mu}) + (A - A\gamma) \ln(1 - \gamma) & \text{if } \chi \neq 0, \xi = \frac{1}{2}(1 - \chi), \\ \frac{A((1-\gamma)^{1+\alpha} - (1-\gamma)^\beta)(1-\gamma)^{-\beta}}{(\alpha-\beta)(1+\alpha-\beta)} + B\gamma \\ + \frac{A\gamma}{\alpha-\beta} - \frac{A\gamma}{\alpha+\beta} + \frac{A\gamma^{1+\alpha+\beta}}{(\alpha+\beta)(1+\alpha+\beta)} + \gamma \tilde{\mu} & \text{otherwise,} \end{cases} \end{aligned}$$

where  $Q_{\text{CSW}}$  is defined in Eq. (9),  $\alpha$  and  $\beta$  are defined in Eq. (7), and

$$\begin{aligned} A &= \tilde{\sigma} [S(3/4 | \chi, \xi) - S(1/4 | \chi, \xi)]^{-1}, \\ B &= -[S(3/4 | \chi, \xi) - S(1/4 | \chi, \xi)]^{-1}. \end{aligned}$$

## 5.3 Apparent scaling law

This section follows the examples of Barndorff-Nielsen and Prause [2001], wherein the normal inverse Gaussian distribution (NIG) has been used to demonstrate the existence of an apparent scaling law for financial returns. By repeating that exercise with the GLD, the ability of the GLD to model both financial returns and the NIG will be demonstrated.

The first so-called scaling law was discovered by Mandelbrot [1963] with cotton prices. Mandelbrot has shown that the averaged daily squared log-returns of cotton prices are self similar. That is, the aggregated returns have the same distribution as the original returns up to a scaling factor. Another well known scaling law, reported by Mueller et al. [1990], is the scaling law of the empirical volatility in high-frequency exchange rate returns. Schnidrig and Würtz

[1995] observed the same scaling law in the US dollar and Deutschmark (USD/DEM) exchange rates. Herein, the volatility is defined as the average of absolute logarithm price changes. The volatility over a duration time,  $\tau$ , is

$$v_\tau = \langle |x_t - x_{t-\tau}| \rangle$$

where  $x_t = \ln \frac{P_t}{P_{t-1}}$  are the logarithm returns and  $P_t$  are the prices at time  $t$ . There is a scaling law

$$v_\tau = \left(\frac{\tau}{T}\right)^H v_T,$$

where  $T$  is an arbitrary time interval and  $H$  is the scaling law index factor. In the last decade, several new scaling laws have been reported. The interested reader may wish to refer to [Glattfelder et al., 2011] for a review of known scaling laws.

Barndorff-Nielsen and Prause [2001] have shown that the apparent scaling of the volatility of high-frequency data is largely due to the semi-heavy-tailedness of financial distributions. The scaling parameter,  $H$ , can be estimated by a linear regression on the logarithm of the scaling relation in the scaling law.

Figure 16 displays the scaling laws of random variates generated by the GLD with location parameter  $\tilde{\mu} = 0$ , scale parameter  $\tilde{\sigma} = 5 \times 10^{-4}$ , and shape parameter  $\chi = 1$ . It is clear that varying the skewness parameter  $\xi$  of the GLD changes the scaling constant (i.e., the regression coefficient in the log-log transform). This is a good illustration of how a semi-heavy fat-tailed distribution can reproduce a scaling law, and of how the GLD parameters can be modified to obtain a desired scaling constant.

Another typical effect of the temporal aggregation of financial data is the Gaussianization of the aggregated financial returns. For high-frequency data, the returns are usually modeled by fat-tailed distributions, since the normal distribution has the wrong shape. However, when the returns are aggregated over long time intervals, the empirical distribution of the aggregated returns is observed to converge onto the normal distribution. Figure 17 illustrates this effect with the USD/DEM exchange rate. The data set consists of the hourly returns of USD/DEM from January 1, 1996 at 11 p.m. to December 31, 1996 at 11 p.m. Figure 17 shows the fitted GLD shape parameters of the 1, 6, 12, 24, and 60 hour aggregated logarithm returns. The fitted shape parameters appear to be converging onto the normal distribution located at approximately ( $\chi = 0$ ,  $\xi = 0.366$ ) as reported in Table 2.

In Fig. 18, the same experiment has been performed with US dollar and Swiss Franc (USD/CHF) exchange rate tick data for the years 2004, 2006, 2008 and 2010<sup>2</sup>. The aggregated returns can be seen to converge in distribution to a normal distribution.

---

<sup>2</sup>The exchange rate tick data was provided by [oanda.com](http://oanda.com)

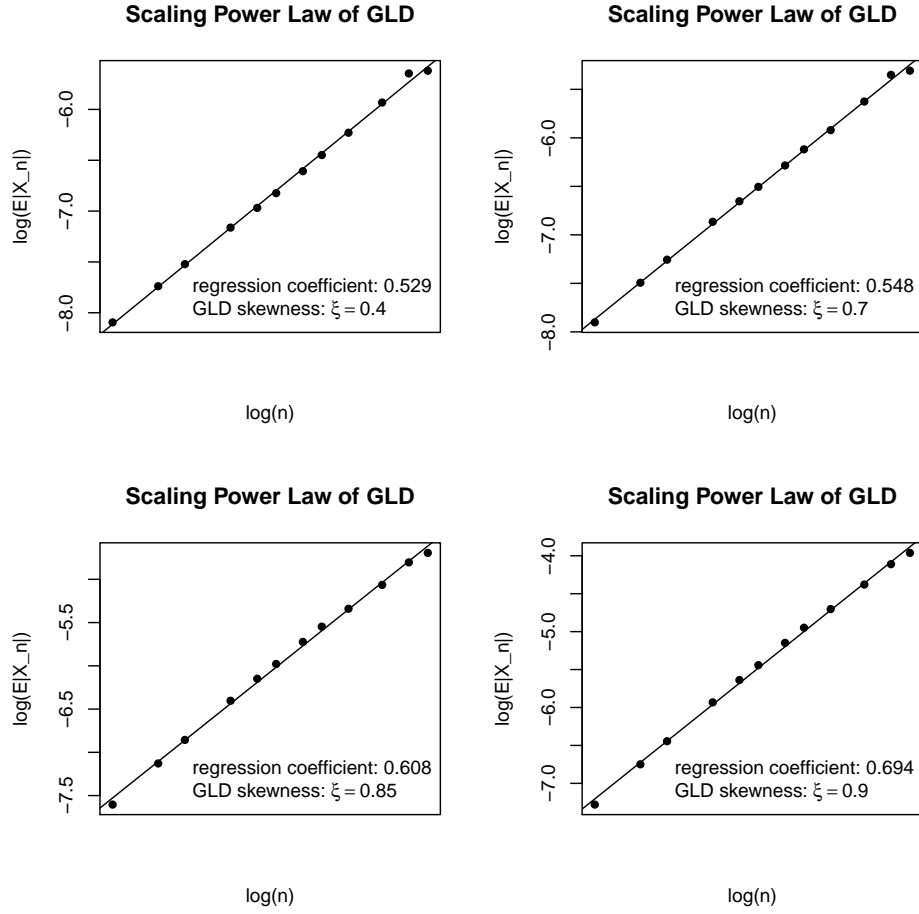


Figure 16: Scaling power law of random variates generated by the GLD with location parameter  $\tilde{\mu} = 0$ , scale parameter  $\tilde{\sigma} = 5 \times 10^{-4}$ , and shape parameter  $\chi = 1$ . The remaining shape parameter is reported in the graphics.

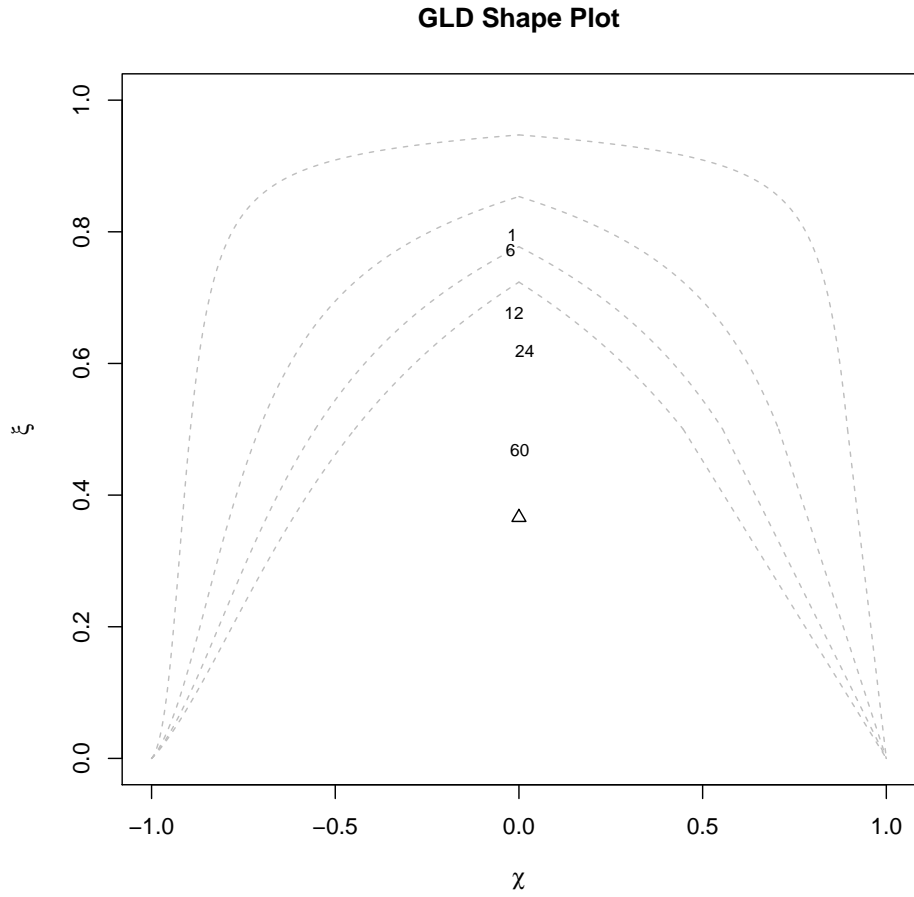


Figure 17: Fitted GLD shape parameters of the 1, 6, 12, 24, and 60 hour aggregated logarithm returns of USD/DEM exchange rate. The triangle corresponds to the shape parameters for which the GLD approximate best the normal distribution (see Table 2)

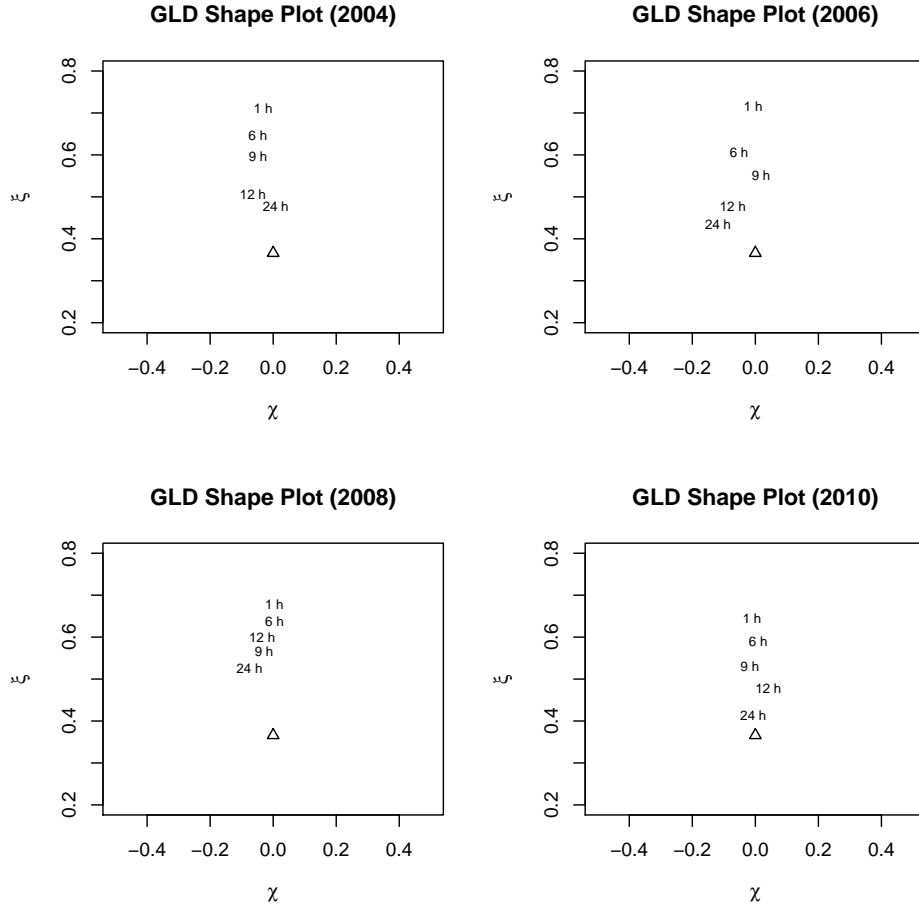


Figure 18: Aggregation of the logarithm returns of USD/CHF exchange rate tick data for the years 2004, 2006, 2008 and 2010. The aggregation level used are 1, 6, 12, 24 hour. The quantile matching method was used to estimate the parameter values of the GLD. The triangle corresponds to the shape parameters for which the GLD approximate best the normal distribution (see Table 2)

## 6 Conclusion

This paper has introduced a new parameterization of the GLD, which provides an intuitive interpretation of its parameters. The median and interquartile range are respectively synonymous with the location and scale parameters of the distribution. The shape parameters describe the asymmetry and steepness of the distribution. This parameterization differs from previous parameterizations, where the asymmetry and steepness of the GLD is expressed in terms of both tail indices  $\lambda_3$  and  $\lambda_4$  of Eqs. (1) and (2). Another advantage of this new parameterization, is that the location and scale parameters can be equated to their sample estimators. This reduces the complexity of the estimator used to fit the remaining two shape parameters. However, this new parameterization comes with the cost of having more intricate expressions for both the conditions of existence for the moments, and the conditions on the shape parameters for the different distribution shapes. Nevertheless, evaluating these expressions remains straightforward. Moreover, the new parameterization enables the use of shape plots that can be used to represent the fitted parameters. Furthermore, value-at-risk, expected shortfall, and the tail indices can be expressed by simple formulae, dependent upon the parameters.

The GLD with unbounded support is an interesting, power-law tailed, alternative to the  $\alpha$ -stable and the Student  $t$ -distribution for purposes of modeling financial returns. The key advantage of the GLD is it can accommodate a large range of distribution shapes. A single distribution can therefore be used to model the data for many different asset classes; e.g., bonds, equities, and alternative instruments. This differs from current practice where different distributions must be employed, corresponding to the different distributional shapes of the data sets. Another point, which is often overlooked in practice, is that the same distribution can be applied to different subsets of the data. For example, the distribution of the returns might completely change from year to year. The GLD would then be able to capture this distributional change.

Due to its flexibility, the GLD is a good candidate for studying structural changes of a data set by Bayesian change-point analysis. This may be of use in the identification of market-change regimes.

The R package **gldist** implements the new GLD parameterization, along with the parameter estimation methods that have been presented in this work.

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

## References

- J. Aldrich. R. A. Fisher and the making of maximum likelihood 1912–1922. *Statistical Science*, 12(3):162–176, Aug. 1997.
- W. H. Asquith. L-moments and TL-moments of the generalized lambda distribution. *Computational Statistics & Data Analysis*, 51(9):4484–4496, May 2007.
- O. E. Barndorff-Nielsen and K. Prause. Apparent scaling. *Finance and Stochastics*, 5:103–113, 2001.
- 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.
- A. Bowley. *Elements of Statistics*. Scribner, New York, 1920.
- R. C. H. Cheng and N. A. K. Amin. Estimating parameters in continuous univariate distributions with a shifted origin. *Journal of the Royal Statistical Society. Series B (Methodological)*, 45(3):394–403, 1983.
- C. J. Corrado. Option pricing based on the generalized lambda distribution. *Journal of Futures Markets*, 21(3):213–236, 2001.
- 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.
- F. Downton. Linear estimates with polynomial coefficients. *Biometrika*, 53(1/2):129–141, 1966.
- E. A. Elamir and A. H. Seheult. Trimmed L-moments. *Computational Statistics & Data Analysis*, 43(3):299–314, 2003.
- E. Elton, M. Gruber, and C. Blake. Survivor bias and mutual fund performance. *Review of Financial Studies*, 9(4):1097–1120, 1996.
- P. Embrechts, C. Klüppelberg, and T. Mikosch. *Modelling extreme events for insurance and finance*. Springer, Berlin, 1997.
- J. J. Filliben. The probability plot correlation coefficient test for normality. *Technometrics*, 17(1):111–117, Feb. 1975.
- R. A. Fisher. On the mathematical foundations of theoretical statistics. *Lond. Phil. Trans. (A)*, 223:1–33, 1922.



- 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.
- 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.
- D. Freedman and P. Diaconis. On the histogram as a density estimator:  $L_2$  theory. *Probability Theory and Related Fields*, 57(4):453–476, Dec. 1981.
- 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.
- W. Gilchrist. *Statistical modelling with quantile functions*. CRC Press, 2000.
- J. Glattfelder, A. Dupuis, and R. Olsen. Patterns in high-frequency FX data: discovery of 12 empirical scaling laws. *Quantitative Finance*, 11(4):599–614, 2011.
- 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):413–426, 1947.
- D. Hogben. Some properties of Tukey’s test for non-additivity. Unpublished Ph.D. thesis, Rutgers-The State University, 1963.
- A. H. Jessen and T. Mikosch. Regularly varying functions. *Publ. Inst. Math., Nouv. Sér.*, 80(94): 171–192, 2006.
- 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): 394–399, 1971.
- Z. Karian and E. Dudewicz. *Fitting statistical distributions: the generalized lambda distribution and generalized bootstrap methods*. Chapman & Hall/CRC, 2000.
- Z. Karian and E. Dudewicz. Comparison of GLD fitting methods: superiority of percentile fits to moments in  $L_2$  norm. *Journal of Iranian Statistical Society*, 2(2):171–187, 2003.
- 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.

- J. Karvanen and A. Nuutinen. Characterizing the generalized lambda distribution by L-moments. *Computational Statistics & Data Analysis*, 52(4):1971–1983, Jan. 2008.
- 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, 2000.
- T. Kim and H. White. On more robust estimation of skewness and kurtosis. *Finance Research Letters*, 1(1):56–73, Mar. 2004.
- 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.
- R. A. R. King and H. L. MacGillivray. A starship estimation method for the generalized  $\lambda$  distributions. *Australian & New Zealand Journal of Statistics*, 41(3):353–374, 1999.
- S. Kullback. *Information theory and statistics*. John Wiley and Sons, New York, 1959.
- A. Lakhany and H. Mausser. Estimating the parameters of the generalized lambda distribution. *Algo Research Quarterly*, 3(3):47–58, 2000.
- A. Luceño. Fitting the generalized Pareto distribution to data using maximum goodness-of-fit estimators. *Computational Statistics & Data Analysis*, 51(2):904–917, Nov. 2006.
- B. Mandelbrot. The variation of certain speculative prices. *The Journal of Business*, 36(4):394–419, 1963.
- J. J. A. Moors. A quantile alternative for kurtosis. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 37(1):25–32, 1988.
- U. A. Mueller, M. M. Dacorogna, R. B. Olsen, O. V. Pictet, M. Schwarz, and C. Morgenegg. Statistical study of foreign exchange rates, empirical evidence of a price change scaling law, and intraday analysis. *Journal of Banking & Finance*, 14(6):1189–1208, 1990.
- 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.
- A. Negiz and A. Çınar. Statistical monitoring of multivariable dynamic processes with state-space models. *AIChE Journal*, 43(8):2002–2020, 1997.
- 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.

- A. Ozturk and R. Dale. Least-squares estimation of the parameters of the generalized lambda-distribution. *Technometrics*, 27(1):81–84, 1985.
- S. Pal. Evaluation of nonnormal process capability indices using generalized lambda distribution. *Quality Engineering*, 17(1):77–85, 2004.
- R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- J. S. Ramberg and B. W. Schmeiser. An approximate method for generating asymmetric random variables. *Commun. ACM*, 17(2):78–82, 1974.
- J. S. Ramberg, P. R. Tadikamalla, E. J. Dudewicz, and E. F. Mykytka. A probability distribution and its uses in fitting data. *Technometrics*, 21(2):201–214, May 1979.
- B. Ranneby. The maximum spacing method. An estimation method related to the maximum likelihood method. *Scandinavian Journal of Statistics*, 11(2):93–112, 1984.
- R. Schnidrig and D. Würtz. Investigation of the volatility and autocorrelation function of the USD/DEM exchange rate on operational time scales. In *High Frequency Data in Finance Conference (HFDF-I)*, Zürich, 1995.
- D. W. Scott. On optimal and data-based histograms. *Biometrika*, 66(3):605–610, 1979.
- S. Shapiro, M. Wilk, and H. Chen. A comparative study of various tests for normality. *Journal of the American Statistical Association*, 63(324):1343–1372, 1968.
- 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.
- 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.
- H. A. Sturges. The choice of a class interval. *Journal of the American Statistical Association*, 21(153):65–66, Mar. 1926.
- S. Su. A discretized approach to flexibly fit generalized lambda distributions to data. *Journal of Modern Applied Statistical Methods*, 4(2):408–424, 2005.
- S. Su. Numerical maximum log likelihood estimation for generalized lambda distributions. *Computational Statistics & Data Analysis*, 51(8):3983–3998, May 2007.

- S. Su. Fitting GLD to data via quantile matching method. In Z. Karian and E. Dudewicz, editors, *Handbook of Fitting Statistical Distributions with R*. CRC Press/Taylor & Francis, 2010.
- A. Tarsitano. Fitting the generalized lambda distribution to income data. In *COMPSTAT'2004 Symposium*, pages 1861–1867. Physica-Verlag/Springer, 2004.
- A. Tarsitano. Comparing estimation methods for the FPLD. *Journal of Probability and Statistics*, 2010.
- 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.
- J. W. Tukey. The future of data analysis. *The Annals of Mathematical Statistics*, 33(1):1–67, Mar. 1962.
- J. Van Dyke. Numerical investigation of the random variable  $y = c(u^\lambda - (1 - u)^\lambda)$ . Unpublished working paper, National Bureau of Standards, 1961.

## About the Authors

*Diethelm Würtz* is Private Lecturer at the “Institute for Theoretical Physics” at the Swiss Federal Institute of Technology (ETH) in Zurich. His research interests are in the field of risk management and stability analysis of financial markets. He teaches computational science and financial engineering. He is senior partner of the ETH spin-off company “Finance Online” and president of the “Rmetrics Association in Zurich”.

*Yohan Chalabi* has a master in Physics from the Swiss Federal Institute of Technology in Lausanne. He is a PhD student in the Econophysics group at ETH Zurich at the Institute for Theoretical Physics. Yohan is a maintainer of the Rmetrics packages and the R/Rmetrics software environment.

## Acknowledgement

The work presented in this article was partly supported by grants given by ETH Zurich and Rmetrics Association Zurich.