The Continuous Hidden Threshold Mixed Skew-Symmetric Distribution

Mohammed Bouaddi and Rachid Belhachemi and Mohamed Douch

American University in Cairo, American University in Cairo, Royal Military College of Canada

15 December 2015

The Continuous Hidden Threshold Mixed Skew-Symmetric Distribution

Mohammed Bouaddi,* Rachid Belhachemi† and Mohamed Douch ‡

Abstract

This paper explores a way to construct a new family of univariate probability distributions where the parameters of the distribution capture the dependence between the variable of interest and the continuous latent state variable (the regime). The distribution nests two well known families of distributions, namely, the skew normal family of Azzalini (1985) and a mixture of two Arnold et al. (1993) distribution. We provide a stochastic representation of the distribution which enables the user to easily simulate the data from the underlying distribution using generated uniform and normal variates. We also derive the moment generating function and the moments. The distribution comprises eight free parameters that make it very flexible. This flexibility allows the user to capture many stylized facts about the data such as the regime dependence, the asymmetry and fat tails as well as thin tails.

Keywords: Continuous Hidden threshold, Mixture Distribution, Skew-Symmetric distribution, Split Distribution.

*Corresponding Author, American University in Cairo, Department of Economics, Email: m.bouaddi@aucegypt.edu
†Xi'an Jiaotong-Liverpool University, Department of Mathematical Science, Email: rachid.belhachemi@xjtlu.edu.cn
‡Royal Military College of Canada, Department of Management and Economics, Email: mdouch@mdouch.net

1 Introduction

Over the last three decades, extensive research papers have focused on the construction of asymmetric family of distributions which include the normal distribution as a particular case, that are flexible and able to capture a wide range of skewness and kurtosis relative to the normal distribution. For instance, skewed distributions are particularly useful in modelling empirical stock returns which are known to exhibit negative skewness and excess kurtosis.

Univariate skew-symmetric distributions have been studied by several authors. Azzalini (1985,1986) introduced the skew-normal (SN) distribution as a continuous extension of the normal distribution which accommodates asymmetry. A random variable X has a skew normal distribution with parameter λ if its density function is given by,

$$f(x|\lambda) = 2\phi(x)\Phi(\lambda x),\ x, \lambda \in \mathbf{R},$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ denote, respectively, the density and distribution functions of the standard normal distribution, and $\Phi(\lambda x)$ is the skewing function. The special case, where $\lambda = 0$, gives the standard normal distribution. Henze (1986) obtained a stochastic representation for the SN and used it to obtain its odd moments. A comprehensive treatment of the skew normal family of distributions can be found in Genton (2004), Arellano-Valle et al. (2004, 2005) and Gupta et al. (2002).

Another setting in which skewed-normal distributions arise is discussed in Arnold et al. (1993, 2000). Authors consider the distribution of the truncated bivariate normal random variable (X, Y) where X is observed and Y is a hidden truncation. The marginal density of X is obtained and the resulting distribution of X is skew-normal. It was also shown that their general family of distributions contains as a special case the skew-normal distribution of Azzalini (1985).

There also exists another type of general method that is used to transform symmetric distribution into a particular mixture. This class consists of truncated mixtures that are known as split distributions or two-piece distributions. This family is presented by Fernandez and Steel (1998) and generalized by Arellano-Valle et al. (2005).[1] Its generalized form is

$$f(x\,|\alpha) = \pi(\alpha)\frac{2g\left(\dfrac{x}{f_1(\alpha)}\right)}{f_1(\alpha)}I_{\{x\leqslant 0\}} + (1-\pi(\alpha))\frac{2g\left(\dfrac{x}{f_2(\alpha)}\right)}{f_2(\alpha)}\left(1 - I_{\{x\leqslant 0\}}\right)$$

where α is a real, $\pi(\alpha) = \dfrac{f_1(\alpha)}{f_1(\alpha)+f_2(\alpha)}$, $I_{\{x\leqslant 0\}}$ is an indicator function, $g(.)$ is a symmetric density around the origin in its standard form and $f_1(\alpha)$ and $f_2(\alpha)$ are known positive functions that govern the asymmetry and the behavior in the tails of the distribution.

The aim of this paper is to introduce a new family of univariate probability distributions capable of capturing a wide range of values of skewness and kurtosis. Our eight parameters distribution is a mixture of two asymmetric densities making it more flexible than its competitors. The most important contribution to the literature is the inclusion of a latent state variable with a continuum of states, unlike the traditional mixture distributions where the state variable is discrete with few number of states. This new class of distributions will henceforth be referred to as the hidden-threshold-skew-normal (HTSN).

This paper is outlined as follows. In section 2, the HTSN distribution is introduced. We give the stochastic representation of the proposed family of distributions which allows us to simulate data from HTSN distribution by only generating samples from the uniform and the normal distributions. Moreover, the moments generating function and formulas for centred moments are derived. In section 3, we use the HTSN distribution to model the physical distribution of US market returns and the height of Australian athletes. Our results show that the family of HTSN distributions outperforms the family of Skew-distributions introduced by Arnold et al. (1993), the mixture of two hidden truncated normal distributions, the Skew-Generalized Normal Distribution (SGN) discussed in Arellano-Valle et al. (2004), the mixture of two Normals and Split-Normal Distributions. In addition, we present the discrete version of the HTSN with Markov Switching Dynamic and an estimation of this version of HTSN is given using maximum likelihood. Section 4 concludes the paper with a final discussion.

[1] Similar approach can be found in Fernandez et al. (1995), Fernandez and Steel (1998), Mudholkar and Hutson (2000), and Jones (2006).

2 The continuous hidden threshold distribution

2.1 Definition

Definition 1 *The random variable x follows a hidden threshold distribution if its probability density function is defined by*

$$f(x) = \pi\phi\left(\frac{x-\mu_x}{\sigma_{x_1}}\right)\frac{\Phi\left(-\frac{\sigma_{x_1}\left(1+\sigma_{x_1}^{-2}\sigma_{\tau x_1}\right)}{\sqrt{\sigma_{\tau_1}^2-\sigma_{x_1}^{-2}\sigma_{\tau x_1}^2}}\frac{x-\mu_x}{\sigma_{x_1}}-\frac{\mu_\tau-\mu_x}{\sqrt{\sigma_{\tau_1}^2-\sigma_{x_1}^{-2}\sigma_{\tau x_1}^2}}\right)}{\Phi\left(-\frac{\mu_x-\mu_\tau}{\sqrt{\sigma_{\tau_1}^2+\sigma_{x_1}^2-2\sigma_{\tau x_1}}}\right)}$$
$$+\left(1-\pi\right)\phi\left(\frac{x-\mu_x}{\sigma_{x_2}}\right)\frac{\Phi\left(\frac{\sigma_{x_2}\left(1-\sigma_{x_2}^{-2}\sigma_{\tau x_2}\right)}{\sqrt{\sigma_{\tau_2}^2-\sigma_{x_2}^{-2}\sigma_{\tau x_2}^2}}\frac{x-\mu_x}{\sigma_{x_2}}+\frac{\mu_x-\mu_\tau}{\sqrt{\sigma_{\tau_2}^2-\sigma_{x_2}^{-2}\sigma_{\tau x_2}^2}}\right)}{\Phi\left(\frac{\mu_x-\mu_\tau}{\sqrt{\sigma_{\tau_2}^2+\sigma_{x_2}^2-2\sigma_{\tau x_2}}}\right)}. \quad (1)$$

where μ_x, μ_τ, σ_{x_1}, σ_{x_2}, σ_{τ_1}, σ_{τ_2}, $\sigma_{\tau x_1}$, $\sigma_{\tau x_2}$ are the parameters that govern the location, the scale and the shape of the distribution. The mixing probability is given by

$$\pi = \frac{\Phi\left(-\frac{\mu_x-\mu_\tau}{\sqrt{\sigma_{\tau_1}^2+\sigma_{x_1}^2-2\sigma_{\tau x_1}}}\right)}{\Phi\left(-\frac{\mu_x-\mu_\tau}{\sqrt{\sigma_{\tau_1}^2+\sigma_{x_1}^2-2\sigma_{\tau x_1}}}\right)+\Phi\left(\frac{\mu_x-\mu_\tau}{\sqrt{\sigma_{\tau_2}^2+\sigma_{x_2}^2-2\sigma_{\tau x_2}}}\right)} \quad (2)$$

We will show below that the distribution (*1*) is the marginal of x derived from (*5*).

For the derivation see Appendix 1■

We now introduce some general notation, which is used throughout the reminder of this paper. We set $\lambda_{1i} = \frac{\sigma_{xi}\left(1-\sigma_{xi}^{-2}\sigma_{\tau xi}\right)}{\sqrt{\sigma_{\tau i}^2-\sigma_{xi}^{-2}\sigma_{\tau xi}^2}}$, $\lambda_{0i} = \frac{\mu_x-\mu_\tau}{\sqrt{\sigma_{\tau i}^2-\sigma_{xi}^{-2}\sigma_{\tau xi}^2}}$ and $\Delta_i = \frac{\mu_x-\mu_\tau}{\sqrt{\sigma_{\tau i}^2+\sigma_{xi}^2-2\sigma_{\tau xi}}} = \frac{\lambda_{0i}}{\sqrt{1+\lambda_{1i}^2}}$, $(i=1,2)$, then a re-parametrization of (*1*) is

$$f(x) = \pi\phi\left(\frac{x-\mu_x}{\sigma_{x_1}}\right)\frac{\Phi\left(-\lambda_{01}-\lambda_{11}\frac{x-\mu_x}{\sigma_{x_1}}\right)}{\Phi\left(-\Delta_1\right)}$$
$$+\left(1-\pi\right)\phi\left(\frac{x-\mu_x}{\sigma_{x_2}}\right)\frac{\Phi\left(\lambda_{02}+\lambda_{12}\frac{x-\mu_x}{\sigma_{x_2}}\right)}{\Phi\left(\Delta_2\right)}, \quad (3)$$

with mixing probability,

$$\pi = \frac{\Phi\left(-\Delta_1\right)}{\Phi\left(-\Delta_1\right) + \Phi\left(\Delta_2\right)}. \tag{4}$$

If we set $\mu_x = \mu_\tau$, $\sigma_{x_1}^{-2}\sigma_{\tau x_1} = \sigma_{x_2}^{-2}\sigma_{\tau x_2} = 1$, and $\sigma_{x_1} = \sigma_{x_2}$ we get the normal distribution with mean μ_x and standard deviation σ_{x_1}.

If π is allowed to be a free parameter, i.e. independent of the other parameters then we obtain a mixture of two Arnold et al. (1993) distribution with mixing probability π.

Also if $\mu_x = \mu_\tau = 0$, we get a mixture of two Azzalini's Skew-Normal distributions with skewness parameters $-\lambda_{11}$ and λ_{12}, scaling parameters σ_{x_1} and σ_{x_2}, with mixing probability $\pi = \frac{1}{2}$.

If in addition $\sigma_{x_1} = \sigma_{x_2} = 1$, we have a mixture of two Azzalini's standard skew-normal distributions with skewness parameters $-\dfrac{1-\sigma_{\tau x_1}}{\sqrt{\sigma_{\tau_1}^2 - \sigma_{\tau x_1}^2}}$ and $\dfrac{1-\sigma_{\tau x_2}}{\sqrt{\sigma_{\tau_2}^2 - \sigma_{\tau x_2}^2}}$. Moreover, if we set $\sigma_{\tau x_1} = \sigma_{\tau x_2} = 0$ we get a mixture of two Azzalini's skew-normal distributions with location parameter μ_x, scaling parameters σ_{x_1} and σ_{x_2} and, skewness parameters $-\frac{\sigma_{x1}}{\sigma_{\tau 1}}$ and $\frac{\sigma_{x2}}{\sigma_{\tau 1}}$ with mixing probability

$$\pi = \frac{\Phi\left(-\dfrac{\mu_x - \mu_\tau}{\sqrt{\sigma_{\tau_1}^2 + \sigma_{x_1}^2}}\right)}{\Phi\left(-\dfrac{\mu_\tau - \mu_x}{\sqrt{\sigma_{\tau_1}^2 + \sigma_{x_1}^2}}\right) + \Phi\left(\dfrac{\mu_x - \mu_\tau}{\sqrt{\sigma_{\tau_2}^2 + \sigma_{x_2}^2}}\right)}.$$

On the other hand if we set $\sigma_{x_1}^{-2}\sigma_{\tau x_1} = \sigma_{x_2}^{-2}\sigma_{\tau x_2} = 1$ we get a mixture of two normals with mixing probability

$$\pi = \frac{\Phi\left(-\dfrac{\mu_x - \mu_\tau}{\sqrt{\sigma_{\tau_1}^2 + 3\sigma_{x_1}^2}}\right)}{\Phi\left(-\dfrac{\mu_x - \mu_\tau}{\sqrt{\sigma_{\tau_1}^2 + 3\sigma_{x_1}^2}}\right) + \Phi\left(\dfrac{\mu_x - \mu_\tau}{\sqrt{\sigma_{\tau_2}^2 + 3\sigma_{x_2}^2}}\right)}.$$

2.2 Derivation and stochastic representation

2.2.1 Derivation

Consider the bivariate distribution who's density is given by

$$f(Z) = c\left\{\frac{\exp\left(-\frac{(Z-\mu)'\,\Omega_1^{-1}\,(Z-\mu)}{2}\right)}{2\pi\left|\Omega_1\right|^{\frac{1}{2}}}I + \frac{\exp\left(-\frac{(Z-\mu)'\,\Omega_2^{-1}\,(Z-\mu)}{2}\right)}{2\pi\left|\Omega_2\right|^{\frac{1}{2}}}(1-I)\right\}, \tag{5}$$

where $Z = (x, \tau)'$, x is the observable random variable, τ is a latent random variable i.e., the hidden random threshold, $I = 1$ if $x \leqslant \tau$ and $I = 0$ otherwise, $\mu = (\mu_x, \mu_\tau)'$ is a vector of location parameters of the distribution, Ω_1 and Ω_2 are 2×2 symmetric and positive definite scaling matrices written as

$$\Omega_i = \begin{pmatrix} \sigma_{xi}^2 & \sigma_{\tau xi} \\ \sigma_{\tau xi} & \sigma_{\tau i}^2 \end{pmatrix}, \quad i = 1, 2,$$

and c is the normalizing parameter of the distribution as can be shown in appendix 1 to be,

$$c = \frac{1}{\Phi\left(-\Delta_1\right) + \Phi\left(\Delta_2\right)}. \tag{6}$$

The marginal distribution of the observable x is obtained by integrating out the latent variable τ. The following lemma gives the form of this marginal distribution.

Lemma 1 *The marginal distribution of the observable x is given by (1).*

For the Proof see Appendix 1■

2.2.2 Stochastic representation

In the proposition below we give a stochastic representation of the distribution (*1*). The proposed simple representation allows for easy simulation of random variables from (*1*).

Proposition 1 (Stochastic representation). *Let λ_{1i}, λ_{0i} and Δ_i be as above $(i = 1, 2)$, u and v be two independent standard normal variables and η is uniformly distributed random variable in $[0, 1]$ independent from u and v, where u is truncated below at Δ_1 if $\eta \leqslant \pi$ and above at Δ_2 otherwise. In addition, let*

$$z = \begin{cases} -\frac{\lambda_{11}}{\sqrt{1+\lambda_{11}^2}}u + \frac{1}{\sqrt{1+\lambda_{11}^2}}v, & \textit{if } \eta \leqslant \pi \\ -\frac{\lambda_{12}}{\sqrt{1+\lambda_{12}^2}}u + \frac{1}{\sqrt{1+\lambda_{12}^2}}v, & \text{otherwise,} \end{cases} \tag{7}$$

and

$$x = \begin{cases} \sigma_{x_1} z + \mu_x, & \textit{if } \eta \leqslant \pi \\ \sigma_{x_2} z + \mu_x, & \text{otherwise,} \end{cases} \tag{8}$$

then x has a distribution with a density function (1).

For the Proof see Appendix 2■

2.2.3 The marginal moments of x

The moments of (1) are given in the following proposition.

Proposition 2 Suppose *x has a distribution with a density function (1). Let* $\mu^* = \begin{pmatrix} \mu_x^* \\ \mu_\tau^* \end{pmatrix} = \begin{pmatrix} \mu_x \\ \mu_\tau - \mu_x \end{pmatrix}$, $\Omega_i^* = \begin{pmatrix} \sigma_{xi}^{*2} & \sigma_{\tau xi}^* \\ \sigma_{\tau xi}^* & \sigma_{\tau i}^{*2} \end{pmatrix} = \begin{pmatrix} \sigma_{xi}^2 & \sigma_{\tau xi} - \sigma_{xi}^2 \\ \sigma_{\tau xi} - \sigma_{xi}^2 & \sigma_{\tau i}^2 + \sigma_{xi}^2 - 2\sigma_{\tau xi} \end{pmatrix}$ *for* $i = 1, 2$, and $h_1 = -\frac{\mu_\tau^*}{\sigma_{\tau 1}^*}$ and $h_2 = -\frac{\mu_\tau^*}{\sigma_{\tau 2}^*}$. *The non-central moment of (1) of order K is given by,*

$$m_k^K(x) = c(I_1^K(x) + I_2^K(x)), \tag{9}$$

where

$$I_1^K(x) = \sum_{k=0}^{K} \binom{K}{k} [\mu_x^* - E(x)]^{K-k} \sum_{j=0}^{k} \binom{k}{j} \left(\sigma_{\tau_1}^{*-2} \sigma_{\tau x_1}^*\right)^{k-j}$$
$$\left(\sigma_{x_1}^{*2} - \sigma_{\tau_1}^{*-2} \sigma_{\tau x_1}^{*2}\right)^{\frac{j}{2}} I_j \sum_{i=0}^{k-j} \binom{k-j}{i} \mu_\tau^{*k-j-i} (\sigma_{\tau_1}^*)^i I_i^*, \tag{10}$$

and

$$I_2^K(x) = \sum_{k=0}^{K} \binom{K}{k} \left[\mu_x^* - E(x)\right]^{K-k} \sum_{j=0}^{k} \binom{k}{j} \left(\sigma_{\tau_2}^{*-2}\sigma_{\tau x_2}^*\right)^{k-j}$$
$$\left(\sigma_{x_2}^{*2} - \sigma_{\tau_2}^{*-2}\sigma_{\tau x_2}^{*2}\right)^{\frac{j}{2}} \sum_{i=0}^{k-j} \binom{k-j}{i} \mu_\tau^{*k-j-i} (\sigma_{\tau_2}^*)^i I_i^{**}. \qquad (11)$$

where

$$I_k^* = \frac{1}{\sqrt{2\pi}} 2^{\frac{k-1}{2}} \Gamma(\frac{k+1}{2}) \left\{1 - sign(h_1)(-1)^{kI_{h_1<0}} \gamma(\frac{k+1}{2}, \frac{h_1^2}{2})\right\},$$

and

$$I_k^{**} = \frac{1}{\sqrt{2\pi}} 2^{\frac{k-1}{2}} \Gamma(\frac{k+1}{2}) \left\{(-1)^k + (-1)^{kI_{h_2<0}} \gamma(\frac{k+1}{2}, \frac{h_2^2}{2})\right\},$$

and

$$I_k = \begin{cases} 0, & \text{if } k \text{ is odd} \\ \frac{k!}{(\frac{k}{2})! 2^{\frac{k}{2}}}, & \text{if } k \text{ is even.} \end{cases}$$

For the Proof see Appendix 3■

The four first moments are given in appendix 3.

2.3 The moment generating function and some properties

The moment generating function of (5) is given in the following theorem.

Theorem 1 *The moment generating function of (5) is*

$$M(\theta) = \pi \exp\left(\theta'\mu + \frac{\theta'\Omega_1\theta}{2}\right) \frac{\Phi\left(\frac{\mu_{\tau_1}^*}{\sqrt{\sigma_{\tau_1}^2 + \sigma_{x_1}^2 - 2\sigma_{\tau x_1}}}\right)}{\Phi\left(-\Delta_1\right)}$$
$$+ (1-\pi)\exp\left(\theta'\mu + \frac{\theta'\Omega_2\theta}{2}\right) \frac{\Phi\left(-\frac{\mu_{\tau_2}^*}{\sqrt{\sigma_{\tau_2}^2 + \sigma_{x_2}^2 - 2\sigma_{\tau x_2}}}\right)}{\Phi\left(\Delta_2\right)}, \qquad (12)$$

where $\mu_{\tau i}^* = \mu_\tau - \mu_x + \theta_x\left(\sigma_{\tau x i} - \sigma_{x_1}^2\right) + \theta_\tau\left(\sigma_{\tau i}^2 - \sigma_{\tau x i}\right)$ *for* $i = 1, 2$ *and the mixing probability* π *is given*

by (*2*).

For the Proof see Appendix 4■

The following lemma gives the moment generating function of the marginal density of x in (*1*).

Lemma 2 *The moment generating function of (1) is*

$$M(\theta_x) = \pi \exp\left(\theta_x\mu_x + \frac{\theta_x^2\sigma_{x_1}^2}{2}\right) \frac{\Phi\left(\frac{\mu_\tau - \mu_x + \theta_x\left(\sigma_{\tau x_1} - \sigma_{x_1}^2\right)}{\sqrt{\sigma_{\tau_1}^2 + \sigma_{x_1}^2 - 2\sigma_{\tau x_1}}}\right)}{\Phi\left(-\Delta_1\right)} + (1-\pi)\exp\left(\theta_x\mu_x + \frac{\theta_x^2\sigma_{x2}^2}{2}\right) \frac{\Phi\left(-\frac{\mu_\tau - \mu_x + \theta_x\left(\sigma_{\tau x_2} - \sigma_{x_2}^2\right)}{\sqrt{\sigma_{\tau_2}^2 + \sigma_{x_2}^2 - 2\sigma_{\tau x_2}}}\right)}{\Phi\left(\Delta_2\right)}. \tag{13}$$

For the proof we just set $\theta_\tau = 0$ in (12).

The four first moments are given in appendix 3.

3 Applications

3.1 Maximum likelihood estimation

Let $X_1, X_2, ..., X_n$ be a random sample from $HTSN(\mu_x, \mu_\tau, \sigma_{x_1}, \sigma_{x_2}, \sigma_{\tau_1}, \sigma_{\tau_2}, \sigma_{\tau x_1}, \sigma_{\tau x_2})$ so that the likelihood function is given by

$$L(\theta\,|x_1, x_2, ..., x_n) = \prod_{i=1}^{n} f(x_i\,|\theta) \tag{14}$$

where for $i = 1, ..., n$, $f(x_i\,|\theta)$ is given by *(1)*. The maximum likelihood estimator of θ is obtained by maximizing (14), a task that has to be accomplished numerically. We employ the BFGS algorithm optimization method based on two datasets which are described below to estimate the parameters by numerically maximizing the log-likelihood function (14) with respect to the parameter vector $\theta = (\mu_x, \mu_\tau, \sigma_{x_1}, \sigma_{x_2}, \sigma_{\tau_1}, \sigma_{\tau_2}, \sigma_{\tau x_1}, \sigma_{\tau x_2})'$. (14) is maximized using the instruction Maximize in RATS and choosing initial values for the parameters as $\mu_x = 174.594$ and $\sigma_{x_1} = 8.24$. The latter values are the sample mean and the sample standard deviation for the Australian athletes dataset. In a similar fashion, we choose as starting values $\mu_x = 0.028$ and $\sigma_{x_1} = 1.07$ for the market excess returns. In terms of computational time, the BFGS algorithm converges rather quickly.

3.2 Application to two datasets

We apply the HTSN to model the physical distribution of two real datasets. The first dataset consists of daily market excess return of the US stock market covering the period, July 1, 1926 to June 28, 2013.[2] The second set concerns the heights (in centimeters) of 100 Australian female athletes available from the Australian Institute of Sport (AIS dataset) extensively used in the literature by Azzalini (1986) and Arellano-Valle et al. (2004).[3]

Summary statistics of the AIS data are given in Table 1 and in Table 9 for the market excess return of the US stock market. These summarizes suggest leptokurtic densities for both examples with negative skewness in all cases.

For performance purposes, we also estimate five competing distributions namely, the hidden truncation normal, mixture of two hidden truncated normal distributions, the Skew-Generalized Normal distribution (SGN), the mixture of two normals and Split-Normal distributions (densities of these mixtures are given bellow).

Mixture of two hidden truncation normal distribution (MHTN)

$$f\left(x\right)=\omega\frac{\frac{1}{\sigma_1}\phi\left(\frac{x-\mu_1}{\sigma_1}\right)\Phi\left(\lambda_1+\lambda_2\frac{x-\mu_1}{\sigma_1}\right)}{\Phi\left(\frac{\lambda_1}{\sqrt{1+\lambda_2^2}}\right)}+\left(1-\omega\right)\frac{\frac{1}{\sigma_2}\phi\left(\frac{x-\mu_2}{\sigma_2}\right)\Phi\left(\lambda_3+\lambda_4\frac{x-\mu_2}{\sigma_2}\right)}{\Phi\left(\frac{\lambda_3}{\sqrt{1+\lambda_4^2}}\right)}$$

where $0<\omega<1$.

Hidden truncation normal distribution (HTN)

$$f\left(x\right)=\frac{\frac{1}{\sigma}\phi\left(\frac{x-\mu}{\sigma}\right)\Phi\left(\lambda_1+\lambda_2\frac{x-\mu}{\sigma}\right)}{\Phi\left(\frac{\lambda_1}{\sqrt{1+\lambda_2^2}}\right)}$$

where $\sigma>0$, λ_1, $\lambda 2$ and μ are all real numbers.

[2] US stock market returns are available at: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

[3] AIS dataset is available at http://azzalini. stat.unipd.it/SN/index.html.

Skew-Generalized normal distribution (SGN)

$$f(x) = \frac{2}{\sigma}\phi\left(\frac{x-\mu}{\sigma}\right)\Phi\left(\frac{\lambda_1(x-\mu)}{\sqrt{\sigma^2+\lambda_2(x-\mu)^2}}\right)$$

where $\sigma > 0$, λ_1, μ and $\lambda_2 \geqslant 0$ are all real numbers .

The mixture of two normals (MN)

$$\begin{aligned} f(x) &= \omega\frac{1}{\sqrt{2\pi}\sigma_{x_1}}\exp\left\{-\frac{1}{2}\left(\frac{x-\mu_{x_1}}{\sigma_{x_1}}\right)^2\right\} \\ &\quad + (1-\omega)\frac{1}{\sqrt{2\pi}\sigma_{x_2}}\exp\left\{-\frac{1}{2}\left(\frac{x-\mu_{x_2}}{\sigma_{x_2}}\right)^2\right\} \end{aligned}$$

where $0 < \omega < 1$.

Split normal distribution (SN)

$$\begin{aligned} f(x) &= \frac{\sqrt{2}}{\sqrt{\pi}\left(\sigma_{x_1}+\sigma_{x_2}\right)}\left\{\exp\left[-\frac{1}{2}\left(\frac{x-\mu_x}{\sigma_{x_1}}\right)^2\right]I_{x<\mu_x}\right. \\ &\quad \left. + \exp\left[-\frac{1}{2}\left(\frac{x-\mu_x}{\sigma_{x2}}\right)^2\right]\left(1-I_{x<\mu_x}\right)\right\} \end{aligned}$$

where

$$I_{x<\mu_x} = \begin{cases} 1 & \text{if } x < \mu_x \\ 0 & \text{Otherwise} \end{cases}.$$

Figure 1 shows plots of the HTSN and standard normal distributions for different parameters values. We note that the HTSN distribution nests several density shapes starting from symmetric heavy tails to asymmetric heavy tailed distributions as shown in figure $2-4$ along with particular parameter values. Figures $5-6$ are histograms for both datasets. The superimposed densities are obtained from fitting the HTSN. MHTN, SGN and MN using maximum likelihood estimation. The best fit of the HTSN over the other distributions are also illustrated in figures $5-6$, which show that HTSN density is a better fit than the other densities. It is also worth noting that the fitted HTSN captures all the skewness and kurtosis present in the data.

Tables $2-5$ show our results using AIS dataset. According to the BIC information criterion reported at the bottom of Tables $2-7$, we conclude that the HTSN model provides the best fit compared to the other distributions using this dataset. Tables $10-15$ provide results using US market excess returns which show by using the same criteria (BIC and AIC) that HTSN outperforms all the five competing distributions. These results are interpreted as strong evidence in favor of the HTSN distribution.

3.3 Discrete version of HTSN distribution

From (5) it is clear that the parameters $\sigma_{\tau x_1}$ and $\sigma_{\tau x_2}$ in (1) capture the dependence between the observable x and the latent regime (threshold) τ in the bad state and the good state respectively, while μ_τ is the location of the threshold. This property of the coefficient makes the distribution (1) more tractable since it departs from the traditional regime switching models in two ways. On one hand, the regime or state variable is a continuous process handling a good updating of the distribution if the regime changes. On the other hand, regimes of the distribution are also identifiable and hence the distribution doesn't suffer from the problem of label switching unlike the case of discrete mixtures. We now introduce a discrete version of HTSN with a switching regime dynamic.

Henceforth a Markov regime switching model will be abbreviated as (MS). The following description follows closely that of Hamilton (1993).

We assume to have observed x_t but not the state S_t. We consider a two-state, first-order Markov process and assume that the state variable is governed by the Markov chain:

$$P = \begin{pmatrix} P(S_{t+1}=1\,|S_t=1) & P(S_{t+1}=1\,|S_t=2) \\ P(S_{t+1}=2\,|S_t=1) & P(S_{t+1}=2\,|S_t=2) \end{pmatrix} = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix}$$

These transition probabilities are restricted so that $p_{12} = 1 - p_{11}$, this follows because starting in regime 1, you can only switch to either regime 2 or stay in 1, etc.

In order to estimate the parameters of an MS model with this uncertainty, we must compute probabilities associated with each possible regime. An estimation of the parameter vector $\theta = (\mu_x, \mu_\tau, \sigma_{x_1}, \sigma_{x_2}, \sigma_{\tau_1}, \sigma_{\tau_2}, \sigma_{\tau x_1}, \sigma_{\tau x_2}, p_{11}, p_{22})$ in the MS model is carried out using maximum likelihood as described below. The basic assumption made here is the existence of the regime variables $S_t = 1, 2$, which for each time t selects one of the following distributions

$$f_j(x) = \phi\left(\frac{x-\mu_x}{\sigma_{xj}}\right)\frac{\Phi\left(-\frac{\sigma_{xj}(1-\sigma_{xj}^{-2}\sigma_{\tau xj})}{\sqrt{\sigma_{\tau j}^2-\sigma_{xj}^{-2}\sigma_{\tau xj}^2}}\frac{x-\mu_x}{\sigma_{xj}} - \frac{\mu_x-\mu_\tau}{\sqrt{\sigma_{\tau j}^2-\sigma_{xj}^{-2}\sigma_{\tau xj}^2}}\right)}{\Phi\left(\pm\Delta_j\right)}, \;\; j = 1,2$$

which is then observed, i.e. the conditional density of x_t given that $S_t = j, \theta$ is equal to f_j.The log-likelihood of the model is then given by,

$$\ln L = \sum_{t=1}^{T} \ln \sum_{j=1}^{2} \left(f_j\left(x_t \left| S_t = j, \theta\right.\right) \Pr(S_t = j)\right) \tag{15}$$

Where T is the number of observations. (15) may be interpreted as the weighted average of the likelihood in each state, where the weights are given by the state's probabilities. Since $Pr(S_t = j)$ is not observed, equation (15) cannot be directly used. Instead we apply Hamilton's (1993) method to calculate filtered probabilities for each state based on the available information. The estimation of MS model is obtained by finding the set of parameters that maximizes (15). We apply this methodology to the market excess return and AIS datasets and display the results of this estimation in tables 8 and 16.

We note that convergence of the discrete HTSN proved to be slower compared to the continuous HTSN. The two main drawbacks of the discrete HTSN model are, first, an updating of the probability of each state is required with the arrival of new information, unlike the continuous HTSN where the process is updated continuously in time. Second, since the states are unobservable, estimates based on forecasts of the state in the following period are inconsistent as we have came across in the estimation procedure. This inconsistencies are mainly due to the nonidentifiability of the two regimes, in contrast to the continuous HTSN where the regimes are identifiable.

4 Conclusion

In this paper we propose a new family of distributions which we referred to as hidden-threshold-skew-normal (HTSN). The most important contribution to the literature is the inclusion of a latent state variable with a continuum of states unlike the traditional mixture distributions where the state variable is discrete with few number of states. The new family of distributions is regime dependent. The distribution contains eight parameters which makes it more flexible than its competitors. A wide range of shapes of HTSN are obtained. The distribution has a mixture interpretation. The information criteria shows that the HTSN distribution outperforms all the proposed competitors, including the split normal, the hidden truncation normal and the mixture of two normals with different location and scale parameters.

References

Arellano-Valle, R. B., H. W. Gomez and F. A. Quintana, (2004). A new classof skew-normal distributions. Communications in Statistics : Theory and Methods, 33(7), 1465−1480.

Arellano-Valle, R. B., H. W. Gomez and F. A. Quintana, (2005). Statistical inference for a general class of asymmetric distributions. Journal of Statistical Planning and Inference, 128, 427−443.

Arnold, B.C., R.J. Beaver, R.A. Groeneveld and W.Q. Meeker, (1993). The nontruncated marginal of a truncated bivariate normal distribution. Psychometrika, 58, 47−478.

Arnold, B. C. and R.J. Beaver, (2000). Hidden truncation models. Sankhya Series A, 62, 23−35.

Azzalini, A., (1985). A class of distribution which includes the normal ones. Scandinavian Journal of Statistics. 12, pp. 171−178.

Azzalini, A. (1986). Further results on a class of distributions which includes the normal ones. Statistica., 46, 199−208.

Fernandez, C. J. O. and M.F.J. Steel, (1995). Modeling and inference with v-spherical distributions. Journal of the American Statistical Association, 90, 1331−1340.

Fernandez, C. J. O. and M.F.J. Steel, (1998). On Bayesian modelling of fat tails and skewness. Journal of the American Statistical Association, 93, 359−371.

Genton, M. G., (2004). Skew-elliptical distributions and their applications: a journey beyond normality. Chapman & Hall/CRC.

Gupta .K., F.C. Chang and W.C. Huang, (2002). Some skew-symmetric models. Random Oper. Stochastic Equations 10, 133−140.

Hamilton, James D. (1993). Estimation, Inference and Forecasting of Time Series Subject to Changes in Regime, in G.S. Maddala, C.R. Rao and H.D. Vinod. (eds.), Handbook of Statistics Vol. 11, Amsterdam: North Holland.

Henze, N., (1986). A probabilistic representation of the skew-normal distribution. Scandinavian Journal of Statistics, 13, 271−275.

Jones, M.C., (2006). A note on rescalings, reparametrizations and classes of distributions. Journal of Statistical Planning and Inference, 136, 3730−3733.

Jones, M.C. and M.J. Faddy, (2003). A skew extension of the t-distribution, with applications. Journal of the Royal Statistical Society, Series B, 65, 159−174.

Ma Y. and M.G. Genton, (2004). A flexible class of skew-symmetric distributions. Scandinavian Journal of Statistics, 31, 459−468.

Mudholkar G.S. and A. D. Hutson, (2000). The epsilon-skew-normal distribution for analysing near-normal data. Journal of Statistical Planning and Inference, 83, 291−309.

Table 1: Descriptive Statistics for the AIS dataset

Sample Mean	Standard Deviation	Skewness	Kurtosis
174.5940	8.2422	−0.5684	4.3212

Table 2: Parameter Estimates for the AIS dataset under HTSN

Parameter	$\hat{\mu}_x$	$\hat{\mu}_\tau$	$\hat{\sigma}_{x_1}$	$\hat{\sigma}_{\tau_1}$	$\hat{\sigma}_{\tau x_1}$	$\hat{\sigma}_{x_2}$	$\hat{\sigma}_{\tau_2}$	$\hat{\sigma}_{\tau x_2}$
Value	173.8034	171.7584	10.8978	5548.9526	−16128.212	6.5308	3.2168	19.9747
Std Error	3.4980	2.9936	2.0894	176746.6703	531285.505	1.7692	3.0566	21.0419
T-Stat	49.6860	57.3751	5.2156	0.0314	−0.0304	3.6914	1.0524	0.9493
Log-likelihood				−336.779				
BIC				7.104				
AIC				7.096				

Table 3: Parameter Estimates for the AIS dataset under MHTN

Parameter	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}_1$	$\hat{\sigma}_2$	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\lambda}_3$	$\hat{\lambda}_4$	$\widehat{\omega}$
Value	163.2777	164.5975	6.4540	8.6878	−29.8276	0.0668	−116.1166	−645.4920	0.9222
Std Error	0.1349	0.05646	0.42550	0.3699	1.8586	0.000741	6.6866	8.9949	0.0303
T-Stat	1210.1854	2915.3633	15.1680	23.4845	−16.0483	90.1539	−17.3656	−71.7620	30.4220
Log-likelihood					−347.568				
BIC					7.366				
AIC					7.131				

Table 4: Parameter Estimates for the AIS dataset under HTN

Parameter	$\hat{\mu}_x$	$\hat{\sigma}_x$	$\hat{\lambda}_1$	$\hat{\lambda}_2$
Value	174.4617896	8.2010669	−1.4901255	0.0083191
Std Error	14.0486411	0.4730751	11.7003650	0.9954514
T-Stat	12.41841	17.33566	−0.12736	0.00836
Log-likelihood		−352.318		
BIC		7.231		
AIC		7.126		

Table 5: Parameter Estimates for the AIS dataset under GSN

Parameter	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\lambda}_1$	$\hat{\lambda}_2$
Value	170.3204	9.2476	4.3802	24.1780
Std Error	0.9776	0.6561	3.9173	46.3275
T-Stat	174.2266	14.0952	1.1182	0.5219
Log-likelihood		-347.2392		
BIC		7.129		
AIC		7.025		

Table 6: Parameter Estimates for the AIS dataset under MN

Parameter	$\hat{\mu}_{x_1}$	$\hat{\mu}_{x_2}$	$\hat{\sigma}_{x_1}$	$\hat{\sigma}_{x_2}$	$\hat{\omega}$
Value	177.022	6.3517	214.982	60.854	0.1810
Std Error	1.5273	0.5519	3.6510	0.6482	0.0069
T-Stat	115.904	11.509	58.883	93.888	26.2996
Log-likelihood			-350.844		
BIC			7.240		
AIC			7.110		

Table 7: Parameter Estimates for the AIS dataset under SN

Parameter	$\hat{\mu}_x$	$\hat{\sigma}_{x_1}$	$\hat{\sigma}_{x_2}$
Value	177.022	9.6983	6.4635
Std Error	1.5273	1.1198	0.9996
T-Stat	115.904	8.6600	6.4662
Log-likelihood		-350.844	
BIC		7.155	
AIC		7.076	

Table 8: Parameter Estimates for the AIS dataset under discrete HTSN

Parameter	$\hat{\mu}_x$	$\hat{\mu}_\tau$	$\hat{\sigma}_{x_1}$	$\hat{\sigma}_{\tau_1}$	$\hat{\sigma}_{\tau x_1}$	$\hat{\sigma}_{x_2}$	$\hat{\sigma}_{\tau_2}$	$\hat{\sigma}_{\tau x_2}$	p_{11}	p_{22}
Value	177.381	177.279	9.911	1.5687	13.3414	6.4076	0.595	-3.813	0.819	0.275
Std Error	0.6863	0.0373	0.817	0.0697	0.6763	0.0394	0.0352	-15.393	0.044	0.000
T-Stat	258.458	4757.75	12.13	22.518	19.727	162.58	16.893	0.2477	18.76	737013
Log-likelihood					-337.772					
BIC					7.216					
AIC					6.955					

Table 9: Descriptive Statistics for the US Market Excess Returns dataset

Sample Mean	Standard Deviation	Skewness	Kurtosis
0.027923	1.0704233	-0.1333	16.6038

Table 10: Parameter Estimates for US Market Excess Returns dataset under HTSN

Parameter	$\hat{\mu}_x$	$\hat{\mu}_\tau$	$\hat{\sigma}_{x_1}$	$\hat{\sigma}_{\tau_1}$	$\hat{\sigma}_{\tau x_1}$	$\hat{\sigma}_{x_2}$	$\hat{\sigma}_{\tau_2}$	$\hat{\sigma}_{\tau x_2}$
Value	0.5344	0.0785	2.2068	2.1827	4.7515	0.7933	343.99	213.78
Std Error	0.0174	0.0042	0.0057	0.0052	0.0197	0.0146	0.1186	8.2648
T-Stat	30.758	18.551	386.58	415.17	241.58	54.331	2898.99	25.866
Log-likelihood				-30256.19				
BIC					2.634			
AIC					2.631			

Table 11: Parameter Estimates for the US Market Excess Returns dataset under MHTN

Parameter	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}_1$	$\hat{\sigma}_2$	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\lambda}_3$	$\hat{\lambda}_4$	$\widehat{\omega}$
Value	-0.0186	0.0424	2.1504	0.6086	-1.7627	-0.0246	-1.5634	0.0172	0.1810
Std Error	0.1907	0.0315	0.0369	0.0069	1.0982	0.0257	0.6494	0.0206	0.0071
T-Stat	-0.0974	1.3471	58.144	88.171	-1.6052	-0.9567	-2.4536	0.8339	25.324
Log-likelihood				-30290.39					
BIC					2.637				
AIC					2.634				

Table 12: Parameter Estimates for the US Market Excess Returns dataset under HTN

Parameter	$\hat{\mu}_x$	$\hat{\sigma}_x$	$\hat{\lambda}_1$	$\hat{\lambda}_2$
Value	0.0220	1.0704	-0.0983	0.0064
Std Error	0.0034	0.0051	0.2600	0.0026
T-Stat	6.5148	209.54	-0.3782	2.4213
Log-likelihood		-34207.77		
BIC			2.976	
AIC			2.974	

Table 13: Parameter Estimates for the US Market Excess Returns dataset under GSN

Parameter	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\lambda}_1$	$\hat{\lambda}_2$
Value	-0.6898	1.2888	3.5417	6.9373
Std Error	0.0087	0.0077	0.1089	0.5461
T-Stat	-78.925	166.937	32.498	12.703
Log-likelihood		-32105.26		
BIC		2.793		
AIC		2.792		

Table 14: Parameter Estimates for the US Market Excess Returns dataset under MN

Parameter	$\hat{\mu}_{x_1}$	$\hat{\mu}_{x_2}$	$\hat{\sigma}_{x_1}$	$\hat{\sigma}_{x_2}$	$\hat{\omega}$
Value	-0.1330	0.0635	2.1498	0.6085	0.1810
Std Error	0.0345	0.0052	0.0375	0.0068	0.0070
T-Stat	-3.8476	12.206	57.313	90.042	25.730
Log-likelihood		-30290.39			
BIC			2.635		
AIC			2.633		

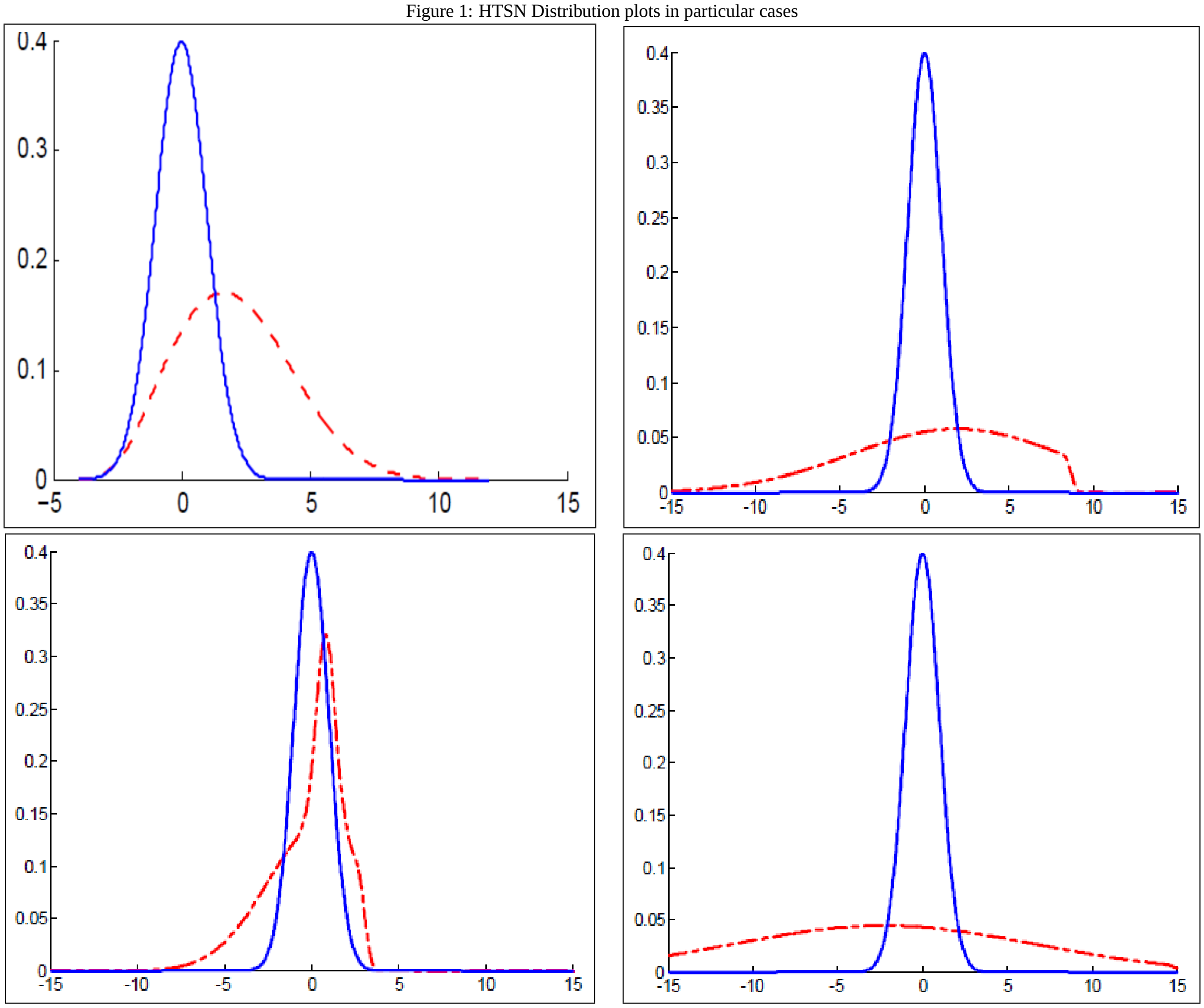
Table 15: Parameter Estimates for the US Market Excess Returns dataset under SN

Parameter	$\hat{\mu}_x$	$\hat{\sigma}_{x_1}$	$\hat{\sigma}_{x_2}$
Value	0.0669	1.1003	1.0394
Std Error	0.0103	0.0075	0.0077
T-Stat	6.4751	147.67	135.37
Log-likelihood		-34195.09	
BIC		2.974	
AIC		2.973	

Table 16: Parameter Estimates for the US Market Excess Returns dataset under discrete HTSN

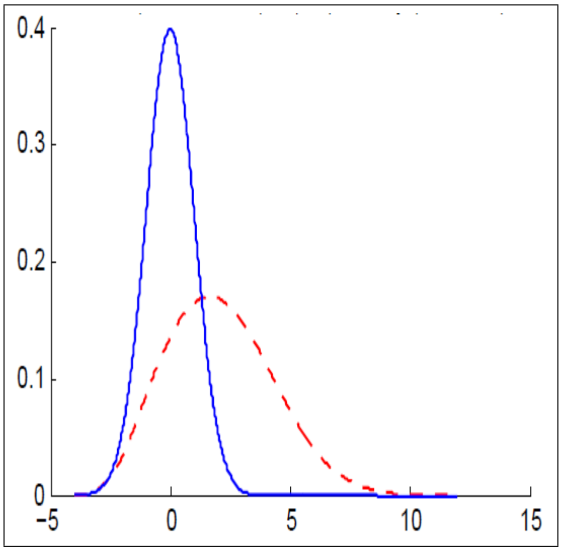
Parameter	$\hat{\mu}_x$	$\hat{\mu}_\tau$	$\hat{\sigma}_{x_1}$	$\hat{\sigma}_{\tau_1}$	$\hat{\sigma}_{\tau x_1}$	$\hat{\sigma}_{x_2}$	$\hat{\sigma}_{\tau_2}$	$\hat{\sigma}_{\tau x_2}$	p_{11}	p_{22}
Value	-1.2347	-5.4433	0.6872	1.5599	0.8338	2.2020	385.38	-548.48	0.9869	0.0410
Std Error	0.0896	0.0601	0.0098	0.0289	0.0178	0.0530	46.712	55.444	0.0011	0.0030
T-Stat	-13.786	-90.560	69.810	53.854	46.759	41.516	8.250	-9.8926	874.52	11.640
Log-likelihood					-28642.69					
BIC					2.494					
AIC					2.491					

Figure 1: HTSN Distribution plots in particular cases



Examples of density shapes using different parameters values. Solid line represents Standard Normal density and the dashed line is for the HTSN.

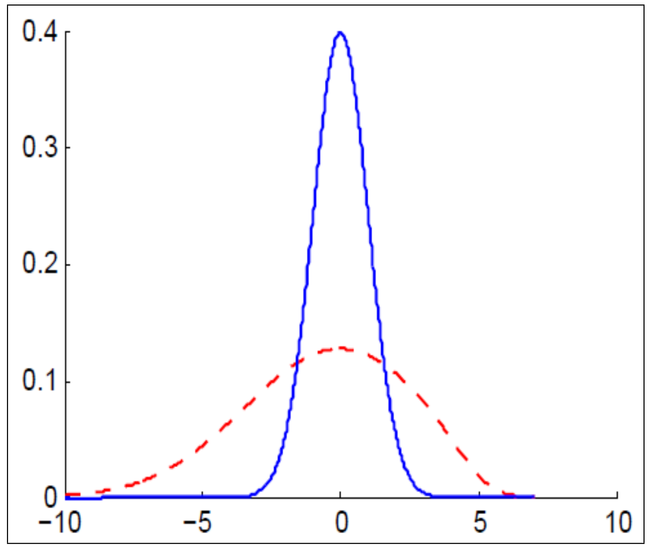
Figure 2: An Example of (right-skewed) heavy tailed HTSN (dashed lines), N(0,1) density (solid line)



$\mu_x = 0.0,\ \sigma_{x_1} = 1.0,\ \sigma_{\tau_1} = 1.0,\ \sigma_{x\tau_1} = 0.8$

$\mu_\tau = 12.0,\ \sigma_{x_2} = 8.0,\ \sigma_{\tau_2} = 3.0,\ \sigma_{x\tau_2} = -0.5$

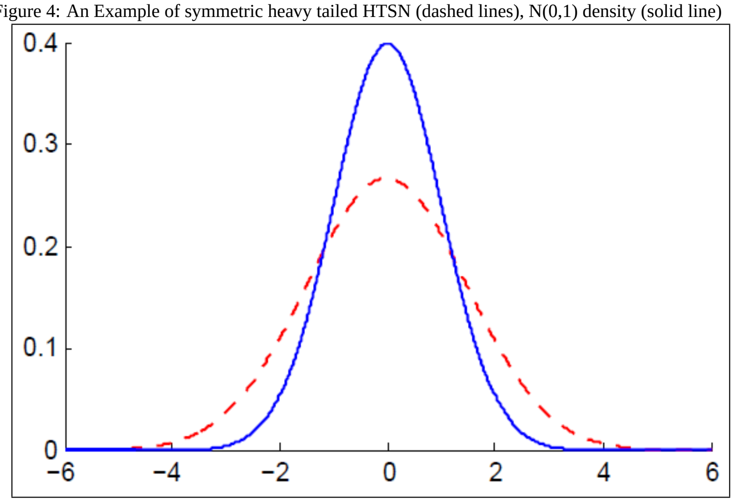
Figure 3: An Example of (left-skewed) heavy tailed HTSN (dashed lines), N(0,1) density (solid line)



$\mu_x = 0.0,\ \sigma_{x_1} = 5.0,\ \sigma_{\tau_1} = 1.0,\ \sigma_{x\tau_1} = 0.5$

$\mu_\tau = -7.0,\ \sigma_{x_2} = 1.0,\ \sigma_{\tau_2} = 2.0,\ \sigma_{x\tau_2} = 0.9$

Figure 4: An Example of symmetric heavy tailed HTSN (dashed lines), N(0,1) density (solid line)

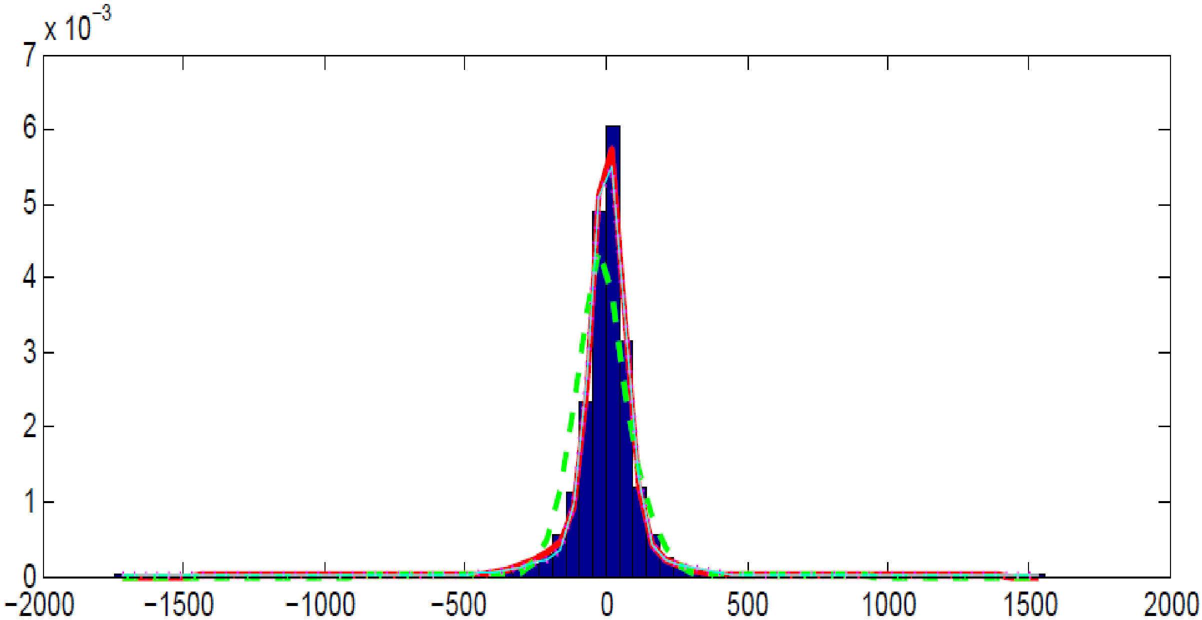


$\mu_x = 0.0,\ \sigma_{x_1} = 1.0,\ \sigma_{\tau_1} = 1.0,\ \sigma_{x\tau_1} = 0.5$

$\mu_\tau = -2.5,\ \sigma_{x_2} = 1.0,\ \sigma_{\tau_2} = 9.0,\ \sigma_{x\tau_2} = -0.9$

Figure 5: Histogram of percentage excess market returns

The lines represents distributions fitted using maximum likelihood estimation

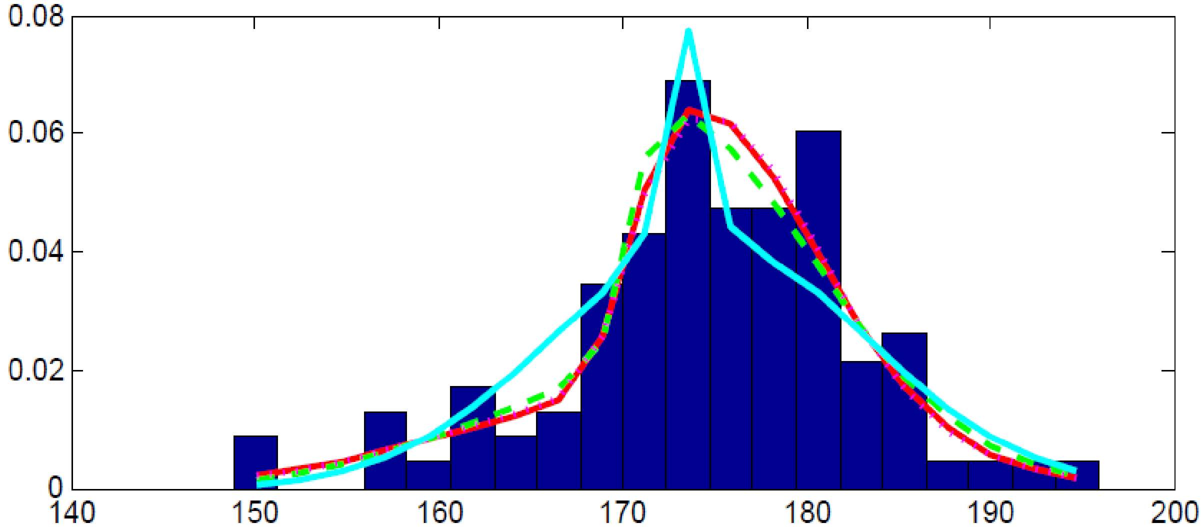


HTSN (solid lines), MTHN (dotted line), SGN (dashed line), MN (dashed-dotted)

which in this case is indistinguishable from MTHN and HTSN

Figure 6: Histogram of heights of 100 Australian athletes

The lines represents distributions fitted using maximum likelihood estimation



HTSN (solid lines), MTHN (dotted line), SGN (dashed line), MN (dashed-dotted) which in this case is indistinguishable from HTSN

Appendix 1

Consider $f(x)$ as the probability density function of a random variable x, we have that

$$\int_{-\infty}^{+\infty} f(x)dx = c\left(I_1 + I_2\right),$$

where

$$\begin{cases} I_1 = \int_{x}^{+\infty} \frac{1}{2\pi\left|\Omega_1\right|^{\frac{1}{2}}} \exp\left\{-\frac{(Z-\mu)'\,\Omega_1^{-1}\,(Z-\mu)}{2}\right\} d\tau \\ I_2 = \int_{-\infty}^{x} \frac{1}{2\pi\left|\Omega_2\right|^{\frac{1}{2}}} \exp\left\{-\frac{(Z-\mu)'\,\Omega_2^{-1}\,(Z-\mu)}{2}\right\} d\tau \end{cases},$$

and

$$c = \frac{1}{\Phi\left(-\Delta_1\right) + \Phi\left(\Delta_2\right)}.$$

Let $Y = Z - \mu$ and

$$\Omega_1 = \begin{pmatrix} \sigma_{\tau_1}^2 & \sigma_{\tau x_1} \\ \sigma_{\tau x_1} & \sigma_{x_1}^2 \end{pmatrix}$$

then,

$$Y\Omega_1^{-1}Y = \omega_{11}y_\tau^2 + 2\omega_{12}y_x y_\tau + \omega_{22}y_x.$$

Noting that

$$\Omega_1^{-1}\Omega_1 = \begin{pmatrix} \omega_{11} & \omega_{12} \\ \omega_{12} & \omega_{22} \end{pmatrix} \begin{pmatrix} \sigma_{\tau_1}^2 & \sigma_{\tau x_1} \\ \sigma_{\tau x_1} & \sigma_{x_1}^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

then

$$\omega_{11}\sigma_{\tau_1}^2 + \omega_{12}\sigma_{\tau x_1} = 1, \tag{16}$$

$$\omega_{11}\sigma_{\tau x_1} + \omega_{12}\sigma_{x_1}^2 = 0, \tag{17}$$

$$\omega_{12}\sigma_{\tau_1}^2 + \omega_{22}\sigma_{\tau x_1} = 0, \tag{18}$$

and

$$\omega_{12}\sigma_{\tau x1} + \omega_{22}\sigma_{x1}^2 = 1, \tag{19}$$

from (16-19) we deduce

$$\omega_{12} = -\sigma_{x_1}^{-2}\omega_{11}\sigma_{\tau x_1},$$

$$\omega_{22} = \sigma_{x_1}^{-2}\left(1 - \omega_{12}\sigma_{\tau x_1}\right),$$

and

$$\begin{aligned}\omega_{11} &= \sigma_{\tau_1}^{-2}\left(1 - \omega_{12}\sigma_{\tau x_1}\right) \\ &= \sigma_{\tau_1}^{-2}\left(1 + \sigma_{x_1}^{-2}\omega_{11}\sigma_{\tau x_1}^{2}\right) \\ &= \left(\sigma_{\tau_1}^{2} - \sigma_{x_1}^{-2}\sigma_{\tau x_1}^{2}\right)^{-1}.\end{aligned}$$

Thus

$$Y\Omega_1^{-1}Y = \omega_{11}y_\tau^2 + 2\omega_{12}y_x y_\tau + \omega_{22}y_x.$$

$$\begin{aligned}Y'\Omega_1^{-1}Y &= \omega_{11}y_\tau^2 + 2\omega_{12}y_x y_\tau + \frac{y_x^2}{\sigma_{x_1}^2} \\ &+ \omega_{12}^2\omega_{11}^{-1}y_x^2 \\ &= \frac{y_x^2}{\sigma_{x1}^2} + \omega_{11}\left(y_\tau + \varpi\right)^2 \\ &= \frac{y_x^2}{\sigma_{x_1}^2} + \frac{\left(y_\tau + \varpi\right)^2}{\left(\sigma_{\tau_1}^2 - \sigma_{x_1}^{-2}\sigma_{\tau x_1}^2\right)},\end{aligned} \tag{20}$$

where $\varpi = (\omega_{12}y_x)/\omega_{11} = -\sigma_{x_1}^{-2}\sigma_{\tau x_1}y_x = -\sigma_{x_1}^{-2}\sigma_{\tau x_1}\left(x - \mu_x\right)$. This implies that,

$$\begin{aligned}I_1 &= \int_x^{+\infty} \frac{1}{2\pi\left|\Omega_1\right|^{\frac{1}{2}}}\exp\left\{-\frac{\left(Z-\mu\right)'\Omega_1^{-1}\left(Z-\mu\right)}{2}\right\}d\tau \\ &= \frac{1}{(2\pi)^{\frac{1}{2}}\left|\Omega_1\right|^{\frac{1}{2}}}\exp\left\{-\frac{\left(x-\mu_x\right)^2}{2\sigma_{x1}^2}\right\}\int_x^{+\infty}\frac{1}{(2\pi)^{\frac{1}{2}}}\exp\left\{-\frac{\left(\tau-\mu_\tau+\varpi\right)^2}{2\left(\sigma_{\tau_1}^2-\sigma_{x_1}^{-2}\sigma_{\tau x_1}^2\right)}\right\}d\tau,\end{aligned}$$

where $u = \dfrac{\tau - \mu_\tau + \varpi}{\left(\sigma_{\tau_1}^2 - \sigma_{x_1}^{-2}\sigma_{\tau x_1}^2\right)^{\frac{1}{2}}}$. It follows that

$$\begin{aligned}
I_1 &= \frac{\left(\sigma_{\tau_1}^2 - \sigma_{x_1}^{-2}\sigma_{\tau x_1}^2\right)^{\frac{1}{2}}}{(2\pi)^{\frac{1}{2}}\left|\Omega_1\right|^{\frac{1}{2}}} \exp\left\{-\frac{(x-\mu_x)^2}{2\sigma_{x_1}^2}\right\} \int_{\frac{x-\mu_\tau+\varpi}{\sqrt{\sigma_{\tau 1}^2 - \sigma_{x1}^{-2}\sigma_{\tau x1}^2}}}^{+\infty} \frac{1}{(2\pi)^{\frac{1}{2}}} \exp\left\{-\frac{u^2}{2}\right\} du \\
&= \phi(\frac{x-\mu_x}{\sigma_{x_1}})\Phi\left(-\frac{x-\mu_\tau+\varpi}{\left(\sigma_{\tau_1}^2 - \sigma_{x_1}^{-2}\sigma_{\tau x_1}^2\right)^{\frac{1}{2}}}\right) \\
&= \phi(\frac{x-\mu_x}{\sigma_{x_1}})\Phi\left(-\lambda_{01} - \lambda_{11}\frac{x-\mu_x}{\sigma_{x_1}}\right).
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
I_2 &= \int_{-\infty}^{x} \frac{1}{2\pi\left|\Omega_2\right|^{\frac{1}{2}}} \exp\left\{-\frac{(Z-\mu)'\,\Omega_2^{-1}\,(Z-\mu)}{2}\right\} d\tau \\
&= \phi(\frac{x-\mu_x}{\sigma_{x_2}})\Phi\left(\lambda_{02} + \lambda_{12}\frac{x-\mu_x}{\sigma_{x_2}}\right),
\end{aligned}$$

and the result follows.

End of the proof■

Appendix 2 Proof of proposition 1

Let $\lambda_{1i} = \dfrac{\sigma_{xi}\left(1+\sigma_{xi}^{-2}\sigma_{\tau xi}\right)}{\sqrt{\sigma_{\tau i}^2 - \sigma_{xi}^{-2}\sigma_{\tau xi}^2}}$, $\lambda_{0i} = \dfrac{\mu_x - \mu_\tau}{\sqrt{\sigma_{\tau i}^2 - \sigma_{xi}^{-2}\sigma_{\tau xi}^2}}$ and $\Delta_i = \dfrac{\mu_\tau - \mu_x}{\sqrt{\sigma_{\tau i}^2 + \sigma_{xi}^2 - 2\sigma_{\tau xi}}} = \frac{\lambda_{0i}}{\sqrt{1+\lambda_{1i}^2}}$, $(i = 1, 2)$. u and v are two independent standard normal variables and η is a uniformly distributed random variable in $[0, 1]$, independent from u and v, where u is truncated below at Δ_1 if $\eta \leqslant \pi$ and above at Δ_2 otherwise, in addition let

$$z = -\frac{\lambda_{1i}}{\sqrt{1+\lambda_{1i}^2}}u + \frac{1}{\sqrt{1+\lambda_{1i}^2}}v$$

then the joint density of u and v given that $\eta \leqslant \pi$ is,

$$f\left(u, v\,|\eta \leqslant \pi\right) = \frac{1}{2\pi}\frac{\exp\left(-\frac{v^2}{2} - \frac{u^2}{2}\right)}{\Phi\left(-\Delta_1\right)} I_{u\,\geqslant\,\Delta_1}$$

and

$$v = \sqrt{1+\lambda_{11}^2}z + \lambda_{11}u.$$

We then have,

$$f\left(z,u\left|\eta\leqslant\pi\right.\right)=\frac{\sqrt{1+\lambda_{11}^{2}}}{2\pi\Phi\left(-\Delta_{1}\right)}\exp\left(-\frac{\left(\sqrt{1+\lambda_{11}^{2}}z+\lambda_{11}u\right)^{2}}{2}-\frac{u^{2}}{2}\right)I_{u\,\geqslant\,\Delta_{1}}$$

$$=\frac{\sqrt{1+\lambda_{11}^{2}}}{2\pi\Phi\left(-\Delta_{1}\right)}\exp\left(-\frac{z^{2}}{2}\right)\exp\left(-\frac{\left(1+\lambda_{11}^{2}\right)\left(u+\frac{\lambda_{11}}{\sqrt{1+\lambda_{11}^{2}}}z\right)^{2}}{2}\right)I_{u\geqslant\Delta_{1}}$$

the marginal density of z is

$$f\left(z\left|\eta\leqslant\pi\right.\right)=\frac{\sqrt{1+\lambda_{11}^{2}}}{2\pi\Phi\left(-\Delta_{1}\right)}\exp\left(-\frac{z^{2}}{2}\right)\int_{\Delta_{1}}^{+\infty}\exp\left(-\frac{\left(1+\lambda_{11}^{2}\right)\left(u+\frac{\lambda_{11}}{\sqrt{1+\lambda_{11}^{2}}}z\right)^{2}}{2}\right)du.$$

Let $h=\sqrt{1+\lambda_{11}^{2}}\left(u+\frac{\lambda_{11}}{\sqrt{1+\lambda_{11}^{2}}}z\right)$ then,

$$f\left(z\left|\eta\leqslant\pi\right.\right)=\frac{1}{\sqrt{2\pi}\Phi\left(-\Delta_{1}\right)}\exp\left(-\frac{z^{2}}{2}\right)\int_{\lambda_{01}+\lambda_{11}z}^{+\infty}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{h^{2}}{2}\right)dh$$

$$=\frac{1}{\sqrt{2\pi}\Phi\left(-\Delta_{1}\right)}\exp\left(-\frac{z^{2}}{2}\right)\Phi\left(-\lambda_{01}-\lambda_{11}z\right)$$

Let $x=\sigma_{x_1}z+\mu_x$ then

$$f\left(x\left|\eta\leqslant\pi\right.\right)=\phi(\frac{x-\mu_x}{\sigma_{x_1}})\frac{\Phi\left(-\lambda_{01}-\lambda_{11}\frac{x-\mu_x}{\sigma_{x_1}}\right)}{\Phi\left(-\Delta_{1}\right)}$$

where

$$\Phi\left(-\lambda_{01}-\lambda_{11}z\right)=\int_{\lambda_{01}+\lambda_{11}z}^{+\infty}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{h^{2}}{2}\right)dh.$$

We can similarly show that,

$$f\left(x\left|\eta>\pi\right.\right)=\phi(\frac{x-\mu_x}{\sigma_{x_2}})\frac{\Phi\left(\lambda_{02}+\lambda_{12}\frac{x-\mu_x}{\sigma_{x_2}}\right)}{\Phi\left(\Delta_{2}\right)}$$

then it follows that

$$\begin{aligned} f(x) &= \pi f\left(x\left|\eta \leqslant \pi\right.\right) + (1-\pi) f\left(x\left|\eta > \pi\right.\right) \\ &= \pi\phi(\frac{x-\mu_x}{\sigma_{x_1}})\frac{\Phi\left(-\lambda_{01}-\lambda_{11}\dfrac{x-\mu_x}{\sigma_{x1}}\right)}{\Phi\left(-\Delta_1\right)} \\ &\quad +(1-\pi)\phi(\frac{x-\mu_x}{\sigma_{x_2}})\frac{\Phi\left(\lambda_{02}+\lambda_{12}\dfrac{x-\mu_x}{\sigma_{x_2}}\right)}{\Phi\left(\Delta_2\right)} \end{aligned}$$

which is just the density function given by (1).

End of the proof■

Appendix 3

The central moment of order K of (1)

$$\begin{aligned} m^K(x) &= \int \left[x - E(x)\right]^K f(Z|_{\mu,\Omega_1,\Omega_2,})dZ, \\ &= c\int_{-\infty}^{+\infty}\int_{x}^{+\infty}\left[x-E(x)\right]^K \frac{\exp\left\{-\dfrac{(Z-\mu)'\,\Omega_1^{-1}\,(Z-\mu)}{2}\right\}}{2\pi\left|\Omega_1\right|^{\frac{1}{2}}}d\tau dx \\ &+ c\int_{-\infty}^{+\infty}\int_{-\infty}^{x}\left[x-E(x)\right]^K \frac{\exp\left\{-\dfrac{(Z-\mu)'\,\Omega_2^{-1}\,(Z-\mu)}{2}\right\}}{2\pi\left|\Omega_2\right|^{\frac{1}{2}}}d\tau dx, \\ &= c(I_1^K(x) + I_2^K(x)), \end{aligned}$$

$$I_1^K(x) = \int_{-\infty}^{+\infty}\int_{x}^{+\infty}\left[x-E(x)\right]^K \frac{\exp\left\{-\dfrac{(Z-\mu)'\,\Omega_1^{-1}\,(Z-\mu)}{2}\right\}}{2\pi\left|\Omega_1\right|^{\frac{1}{2}}}d\tau dx,$$

$$Y = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} Z.$$

$$Z = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} Y.$$

$$
\begin{aligned}
I_1^K(x) &= \int_{-\infty}^{+\infty}\int_0^{+\infty}[y_x - E(x)]^K \frac{\exp\left\{-\frac{(\Upsilon Y-\mu)'\,\Omega_1^{-1}\,(\Upsilon Y-\mu)}{2}\right\}}{2\pi\left|\Omega_1\right|^{\frac{1}{2}}}dy_\tau dy_x \\
&= \int_{-\infty}^{+\infty}\int_0^{+\infty}[y_x - E(x)]^K \frac{\exp\left\{-\frac{(Y-\mu^*)'\,\Omega_1^{*-1}\,(\Upsilon Y-\mu^*)}{2}\right\}}{2\pi\left|\Omega_1\right|^{\frac{1}{2}}}dy_\tau dy_x,
\end{aligned}
$$

where

$$
\Upsilon = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix},
$$

$$
Z = \Upsilon Y,
$$

$$
\mu^* = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \qquad \mu = \begin{pmatrix} \mu_x \\ \mu_\tau - \mu_x \end{pmatrix},
$$

$$
\begin{aligned}
\Omega_1^* &= \begin{pmatrix} \sigma_{xi}^{*2} & \sigma_{\tau xi}^* \\ \sigma_{\tau xi}^* & \sigma_{\tau i}^{*2} \end{pmatrix} \\
&= \left(\Upsilon'\Omega_1^{-1}\Upsilon\right)^{-1} \\
&= \Upsilon^{-1}\Omega_1\Upsilon'^{-1} \\
&= \begin{pmatrix} \sigma_{x_1}^2 & \sigma_{\tau x_1} - \sigma_{x_1}^2 \\ \sigma_{\tau x_1} - \sigma_{x_1}^2 & \sigma_{\tau_1}^2 + \sigma_{x_1}^2 - 2\sigma_{\tau x_1} \end{pmatrix},
\end{aligned}
$$

$$
\begin{aligned}
(\Upsilon Y-\mu)'\,\Omega_1^{-1}\,(\Upsilon Y-\mu) &= (Y-\mu^*)'\,\Omega_1^{*-1}\,(Y-\mu^*) \\
&= \frac{(y_\tau - \mu_\tau^*)^2}{\sigma_{\tau_1}^{*2}} + \frac{(y_x - \mu_x^* + \overline{\omega})^2}{\sigma_{x_1}^{*2} - \sigma_{\tau_1}^{*-2}\sigma_{\tau x_1}^{*2}},
\end{aligned}
$$

where $\overline{\omega} = -\sigma_{\tau_1}^{*-2}\sigma_{\tau x_1}^{*}y_\tau$. It follows that

$$
I_1^K(x) = \int_{-\infty}^{+\infty}\int_0^{+\infty}[y_x - E(x)]^K \frac{\exp\left\{-\frac{(y_\tau-\mu_\tau^*)^2}{2\sigma_{\tau_1}^{*2}}\right\}\exp\left\{-\frac{(y_x-\mu_x^*+\overline{\omega})^2}{2\left(\sigma_{x_1}^{*2}-\sigma_{\tau_1}^{*-2}\sigma_{\tau x_1}^{*2}\right)}\right\}}{2\pi\left|\Omega_1\right|^{\frac{1}{2}}}dy_\tau dy_x.
$$

Let $h_x = y_x - \mu_x^*$ and $h_\tau = y_\tau - \mu_\tau^*$ then $\overline{\omega} = -\sigma_{\tau_1}^{*-2}\sigma_{\tau x_1}^* \left(h_\tau + \mu_\tau^*\right)$ and

$$I_1^K(x) = \int_{-\infty}^{+\infty}\int_{-\mu_\tau^*}^{+\infty} \left[h_x + \mu_x^* - E(x)\right]^K \frac{\exp\left\{-\frac{h_\tau^2}{2\sigma_{\tau_1}^{*2}}\right\}\exp\left\{-\frac{(h_x+\overline{\omega})^2}{2\left(\sigma_{x_1}^{*2} - \sigma_{\tau_1}^{*-2}\sigma_{\tau x_1}^{*2}\right)}\right\}}{2\pi\left|\Omega_1\right|^{\frac{1}{2}}} dh_\tau dh_x$$

$$= \sum_{k=0}^{K}\binom{K}{k}\left[\mu_x^* - E(x)\right]^{K-k}\int_{-\infty}^{+\infty}\int_{-\mu_\tau^*}^{+\infty} h_x^k \frac{\exp\left\{-\frac{h_\tau^2}{2\sigma_{\tau_1}^{*2}}\right\}\exp\left\{-\frac{(h_x+\overline{\omega})^2}{2\left(\sigma_{x_1}^{*2} - \sigma_{\tau_1}^{*-2}\sigma_{\tau x_1}^{*2}\right)}\right\}}{2\pi\left|\Omega_1\right|^{\frac{1}{2}}} dh_\tau dh_x,$$

Now let $u_x = h_x + \overline{\omega}$ and $u_\tau = h_\tau$ then

$$I_1^K(x) = \frac{1}{\left|\Omega_1\right|^{\frac{1}{2}}}\sum_{k=0}^{K}\binom{K}{k}\left[\mu_x^* - E(x)\right]^{K-k}\sum_{j=0}^{k}\binom{k}{j}\left(\sigma_{\tau_1}^{*-2}\sigma_{\tau x_1}^*\right)^{k-j}$$
$$\int_{-\infty}^{+\infty} u_x^j \frac{\exp\left\{-\frac{u_x^2}{2\left(\sigma_{x_1}^{*2} - \sigma_{\tau_1}^{*-2}\sigma_{\tau x_1}^{*2}\right)}\right\}}{\sqrt{2\pi}} du_x \int_{-\mu_\tau^*}^{+\infty}\left(u_\tau + \mu_\tau^*\right)^{k-j}\frac{\exp\left\{-\frac{u_\tau^2}{2\sigma_{\tau_1}^{*2}}\right\}}{\sqrt{2\pi}} du_\tau,$$

$$I_1^K(x) = \frac{1}{\left|\Omega_1\right|^{\frac{1}{2}}}\sum_{k=0}^{K}\binom{K}{k}\left[\mu_x^* - E(x)\right]^{K-k}\sum_{j=0}^{k}\binom{k}{j}\left(\sigma_{\tau_1}^{*-2}\sigma_{\tau x_1}^*\right)^{k-j}$$
$$\int_{-\infty}^{+\infty} u_x^j \frac{\exp\left\{-\frac{u_x^2}{2\left(\sigma_{x_1}^{*2} - \sigma_{\tau_1}^{*-2}\sigma_{\tau x_1}^{*2}\right)}\right\}}{\sqrt{2\pi}} du_x \sum_{i=0}^{k-j}\binom{k-j}{i}\mu_\tau^{*k-j}\int_{-\mu_\tau^*}^{+\infty} u_\tau^i \frac{\exp\left\{-\frac{u_\tau^2}{2\sigma_{\tau_1}^{*2}}\right\}}{\sqrt{2\pi}} du_\tau,$$

Let $v_x = \dfrac{u_x}{\sqrt{\sigma_{x_1}^{*2} - \sigma_{\tau_1}^{*-2}\sigma_{\tau x_1}^{*2}}}$ and $v_\tau = \dfrac{u_\tau}{\sigma_{\tau_1}^*}$ then

$$I_1^K(x) = \frac{1}{\left|\Omega_1\right|^{\frac{1}{2}}}\sum_{k=0}^{K}\binom{K}{k}\left[\mu_x^* - E(x)\right]^{K-k}\sum_{j=0}^{k}\binom{k}{j}\left(\sigma_{\tau_1}^{*-2}\sigma_{\tau x_1}^*\right)^{k-j}$$
$$\left(\sigma_{x_1}^{*2} - \sigma_{\tau_1}^{*-2}\sigma_{\tau x_1}^{*2}\right)^{\frac{j+1}{2}}\left(\sigma_{\tau 1}^{*2}\right)^{\frac{k-j+1}{2}}\int_{-\infty}^{+\infty} v_x^j \frac{\exp\left\{-\frac{v_x^2}{2}\right\}}{\sqrt{2\pi}} dv_x \sum_{i=0}^{k-j}\binom{k-j}{i}\mu_\tau^{*k-j-i}\left(\sigma_{\tau_1}^*\right)^{i+1}\int_{-\frac{\mu_\tau^*}{\sigma_{\tau 1}^*}}^{+\infty} v_\tau^i \frac{\exp\left\{-\frac{v_\tau^2}{2}\right\}}{\sqrt{2\pi}} dv_\tau$$
$$= \frac{\left(\sigma_{x_1}^{*2} - \sigma_{\tau_1}^{*-2}\sigma_{\tau x_1}^{*2}\right)^{\frac{1}{2}}\left(\sigma_{\tau_1}^{*2}\right)^{\frac{1}{2}}}{\left|\Omega_1\right|^{\frac{1}{2}}}\sum_{k=0}^{K}\binom{K}{k}\left[\mu_x^* - E(x)\right]^{K-k}\sum_{j=0}^{k}\binom{k}{j}\left(\sigma_{\tau_1}^{*-2}\sigma_{\tau x_1}^*\right)^{k-j}$$
$$\left(\sigma_{x_1}^{*2} - \sigma_{\tau_1}^{*-2}\sigma_{\tau x_1}^{*2}\right)^{\frac{j}{2}} I_j \sum_{i=0}^{k-j}\binom{k-j}{i}\mu_\tau^{*k-j-i}\left(\sigma_{\tau_1}^*\right)^i I_i^*,$$

where

$$I_j = \int_{-\infty}^{+\infty} v_x^j \frac{\exp\left\{-\frac{v_x^2}{2}\right\}}{\sqrt{2\pi}} dv_x. \tag{21}$$

and

$$I_k^* = \int_{-\frac{\mu_\tau^*}{\sigma_{\tau 1}^*}}^{+\infty} v^k \frac{\exp\left\{-\frac{v^2}{2}\right\}}{\sqrt{2\pi}} dv \tag{22}$$

$$= \int_0^{+\infty} v^k \frac{\exp\left\{-\frac{v^2}{2}\right\}}{\sqrt{2\pi}} dv + \int_{-\frac{\mu_\tau^*}{\sigma_{\tau 1}^*}}^{0} v_\tau^k \frac{\exp\left\{-\frac{v^2}{2}\right\}}{\sqrt{2\pi}} dv$$

From the standard normal properties, we have

$$I_j = \begin{cases} 0, & \text{if } j \text{ is odd} \\ \frac{j!}{(\frac{j}{2})! 2^{\frac{j}{2}}}, & \text{if } j \text{ is even} \end{cases}.$$

Let $y = \frac{v^2}{2}$ then from the first part of (22) we have,

$$\int_0^{+\infty} v^k \frac{\exp\left\{-\frac{v^2}{2}\right\}}{\sqrt{2\pi}} dv = \frac{2^{\frac{k-1}{2}}}{\sqrt{2\pi}} \int_0^{+\infty} y^{\frac{k-1}{2}} \exp(-y) dy$$

$$= \frac{2^{\frac{k-1}{2}}}{\sqrt{2\pi}} \Gamma(\frac{k+1}{2})$$

where

$$\Gamma(a) = \int_0^{+\infty} y^{a-1} \exp(-y) dy.$$

The second part of (22) where we make use of $y = \frac{v^2}{2}$, to get

$$\int_{-\frac{\mu_\tau^*}{\sigma_{\tau 1}^*}}^{0} v_\tau^k \frac{\exp\left\{-\frac{v^2}{2}\right\}}{\sqrt{2\pi}} dv = \frac{2^{\frac{k-1}{2}}}{\sqrt{2\pi}} \begin{cases} -\int_0^{\frac{h_1^2}{2}} y^{\frac{k-1}{2}} \exp(-y) dy & \text{if } h_1 > 0 \\ (-1)^k \int_0^{\frac{h_1^2}{2}} y^{\frac{k-1}{2}} \exp(-y) dy & \text{otherwise} \end{cases}$$

$$= \frac{2^{\frac{k-1}{2}}}{\sqrt{2\pi}} \Gamma(\frac{k+1}{2}) sign(h_1) (-1)^{k I_{h_1 < 0}} \gamma(\frac{k+1}{2}, \frac{h_1^2}{2})$$

where $h_1 = -\frac{\mu_\tau^*}{\sigma_{\tau_1}^*}$ and

$$\gamma(a, x) = \frac{\int_0^x y^{a-1} \exp(-y) dy}{\Gamma(a)}.$$

Therefore,

$$I_k^* = \frac{1}{\sqrt{2\pi}} 2^{\frac{k-1}{2}} \Gamma(\frac{k+1}{2}) \left\{1 - sign(h_1)(-1)^{k I_{h_1 < 0}} \gamma(\frac{k+1}{2}, \frac{h_1^2}{2})\right\}.$$

It can be shown, using the same procedure that,

$$I_k^{**} = \frac{1}{\sqrt{2\pi}} 2^{\frac{k-1}{2}} \Gamma(\frac{k+1}{2}) \left\{ (-1)^k + (-1)^{(k+1) I_{h_2 < 0}} \gamma(\frac{k+1}{2}, \frac{h_2^2}{2}) \right\}.$$

Using (21), (**??**) and the fact that $|\Omega_1| = |\Omega_1^*| = \left(\sigma_{x1}^{*2} - \sigma_{\tau 1}^{*-2} \sigma_{\tau x 1}^{*2}\right)^{\frac{1}{2}} \left(\sigma_{\tau 1}^{*2}\right)^{\frac{1}{2}}$, we get

$$I_1^K(x) = \sum_{k=0}^{K} \binom{K}{k} \left[\mu_x^* - E(x)\right]^{K-k} \sum_{j=0}^{k} \binom{k}{j} \left(\sigma_{\tau_1}^{*-2} \sigma_{\tau x_1}^*\right)^{k-j}$$
$$\left(\sigma_{x_1}^{*2} - \sigma_{\tau_1}^{*-2} \sigma_{\tau x_1}^{*2}\right)^{\frac{j}{2}} I_j \sum_{i=0}^{k-j} \binom{k-j}{i} \mu_\tau^{*k-j-i} (\sigma_{\tau_1}^*)^i I_i^*.$$

Similarly we can show that

$$I_2^K(x) = \sum_{k=0}^{K} \binom{K}{k} \left[\mu_x^* - E(x)\right]^{K-k} \sum_{j=0}^{k} \binom{k}{j} \left(\sigma_{\tau_2}^{*-2} \sigma_{\tau x_2}^*\right)^{k-j}$$
$$\left(\sigma_{x_2}^{*2} - \sigma_{\tau_2}^{*-2} \sigma_{\tau x_2}^{*2}\right)^{\frac{j}{2}} I_j \sum_{i=0}^{k-j} \binom{k-j}{i} \mu_\tau^{*k-j-i} (\sigma_{\tau_2}^*)^i I_i^{**}.$$

It follows that the four moments are given by,

$$E(x) = \mu = \mu_x^* + \frac{1}{I_0^* + I_0^{**}} (\sigma_{\tau_1}^{*-2} \sigma_{\tau x_1}^* (\sigma_\tau^* I_1^* + \mu_\tau^* I_0^*) + \sigma_{\tau_2}^{*-2} \sigma_{\tau x_2}^* (\sigma_{\tau_2}^* I_1^{**} + \mu_\tau^* I_0^{**})$$

$$\begin{aligned}
var(x) &= E\left((x-\mu)^2\right) = c[(\mu_x^* - \mu)^2 (I_0^* + I_0^{**}) + 2(\mu_x^* - \mu)\left(\sigma_{\tau_1}^{*-2} \sigma_{\tau x_1}^* (\sigma_{\tau_1}^* I_1^* + \mu_\tau^* I_0^*) + \sigma_{\tau_2}^{*-2} \sigma_{\tau x_2}^* (\sigma_{\tau_2}^* I_1^{**} + \mu_\tau^* I_0^{**})\right) \\
&\quad + (\sigma_{\tau_1}^{*-2} \sigma_{\tau x_1}^*)^2 \left(\sigma_{\tau_1}^{*2} I_2^* + 2\mu_\tau^* \sigma_{\tau_1}^* I_1^* + \mu_\tau^{*2} I_0^*\right) + (\sigma_{\tau_2}^{*-2} \sigma_{\tau x_2}^*)^2 \left(\sigma_{\tau_2}^{*2} I_2^{**} + 2\mu_\tau^* \sigma_{\tau_2}^* I_1^{**} + \mu_\tau^{*2} I_0^{**}\right) \\
&\quad + (\sigma_{x_1}^{*2} - \sigma_{\tau_1}^{*-2} \sigma_{\tau x_1}^{*2}) I_0^* + (\sigma_{x_2}^{*2} - \sigma_{\tau_2}^{*-2} \sigma_{\tau x_2}^{*2}) I_0^{**}]
\end{aligned}$$

$$\begin{aligned}
E\left((x-\mu)^3\right) &= c[(\mu_x^* - \mu)^3 (I_0^* + I_0^{**}) + 3(\mu_x^* - \mu)^2 \left((\sigma_{\tau_1}^{*-2} \sigma_{\tau x_1}^* (\sigma_{\tau_1}^* I_1^* + \mu_\tau^* I_0^*) + \sigma_{\tau_2}^{*-2} \sigma_{\tau x_2}^* (\sigma_{\tau_2}^* I_1^{**} + \mu_\tau^* I_0^{**})\right) \\
&\quad + 3(\mu_x^* - \mu)\left\{(\sigma_{\tau_1}^{*-2} \sigma_{\tau x_1}^*)^2 \left(\sigma_{\tau_1}^{*2} I_2^* + 2\mu_\tau^* \sigma_{\tau_1}^* I_1^* + \mu_\tau^{*2} I_0^*\right) + (\sigma_{\tau_2}^{*-2} \sigma_{\tau x_2}^*)^2 [\sigma_{\tau_2}^{*2} I_2^{**} + 2\mu_\tau^* \sigma_{\tau_2}^* I_1^{**}\right. \\
&\quad \left. + \mu_\tau^{*2} I_0^{**}] + (\sigma_{x_1}^{*2} - \sigma_{\tau_1}^{*-2} \sigma_{\tau x_1}^{*2}) I_0^* + (\sigma_{x_2}^{*2} - \sigma_{\tau_2}^{*-2} \sigma_{\tau x_2}^{*2}) I_0^{**}\right\} + (\sigma_{\tau_1}^{*-2} \sigma_{\tau x_1}^*)^3 [\sigma_{\tau_1}^{*3} I_3^* + 3\mu_\tau^* \sigma_{\tau_1}^{*2} I_2^* \\
&\quad + 3\mu_\tau^{*2} \sigma_{\tau_1}^* I_1^* + \mu_\tau^{*3} I_0^*] + (\sigma_{\tau_2}^{*-2} \sigma_{\tau x_2}^*)^3 \left(\sigma_{\tau 2}^{*3} I_3^{**} + 3\mu_\tau^* \sigma_{\tau 2}^{*2} I_2^{**} + 3\mu_\tau^{*2} \sigma_{\tau_2}^* I_1^{**} + \mu_\tau^{*3} I_0^{**}\right) \\
&\quad + 3\sigma_{\tau_1}^{*-2} \sigma_{\tau x_1}^* \left(\sigma_{x_1}^{*2} - \sigma_{\tau_1}^{*-2} \sigma_{\tau x_1}^{*2}\right) \left(\sigma_{\tau_1}^* I_1^* + \mu_\tau^* I_0^*\right) + 3\sigma_{\tau_2}^{*-2} \sigma_{\tau x_2}^* \left(\sigma_{x_2}^{*2} - \sigma_{\tau_2}^{*-2} \sigma_{\tau x_2}^{*2}\right) \left(\sigma_{\tau_2}^* I_1^{**} + \mu_\tau^* I_0^{**}\right)]
\end{aligned}$$

$$
\begin{aligned}
E\left((x-\mu)^4\right) &= c[(\mu_x^*-\mu)^4\left(I_0^*+I_0^{**}\right)+4\left(\mu_x^*-\mu\right)^3\left((\sigma_{\tau_1}^{*-2}\sigma_{\tau x_1}^*(\sigma_{\tau_1}^* I_1^*+\mu_\tau^* I_0^*)+\sigma_{\tau_2}^{*-2}\sigma_{\tau x_2}^*(\sigma_{\tau_2}^* I_1^{**}+\mu_\tau^* I_0^{**})\right)\\
&+6\left(\mu_x^*-\mu\right)^2\left\{(\sigma_{\tau_1}^{*-2}\sigma_{\tau x_1}^*)^2(\sigma_{\tau_1}^{*2}I_2^*+2\mu_\tau^*\sigma_{\tau_1}^* I_1^*+\mu_\tau^{*2}I_0^*)+(\sigma_{\tau_2}^{*-2}\sigma_{\tau x_2}^*)^2(\sigma_{\tau_2}^{*2}I_2^{**}+2\mu_\tau^*\sigma_{\tau_2}^* I_1^{**}+\mu_\tau^{*2}I_0^{**}\right)\\
&+(\sigma_{x_1}^{*2}-\sigma_{\tau_1}^{*-2}\sigma_{\tau x_1}^{*2})I_0^*+(\sigma_{x_2}^{*2}-\sigma_{\tau_2}^{*-2}\sigma_{\tau x_2}^{*2})I_0^{**}\}+4\left(\mu_x^*-\mu\right)[(\sigma_{\tau_1}^{*-2}\sigma_{\tau x_1}^*)^3(\sigma_{\tau_1}^{*3}I_3^*+3\mu_\tau^*\sigma_{\tau_1}^{*2}I_2^*\\
&+3\mu_\tau^{*2}\sigma_{\tau_1}^* I_1^*+\mu_\tau^{*3}I_0^*)+(\sigma_{\tau_2}^{*-2}\sigma_{\tau x_2}^*)^3\left(\sigma_{\tau_2}^{*3}I_3^{**}+3\mu_\tau^*\sigma_{\tau 2}^{*2}I_2^{**}+3\mu_\tau^{*2}\sigma_{\tau_2}^* I_1^{**}+\mu_\tau^{*3}I_0^{**}\right)\\
&+3\sigma_{\tau_1}^{*-2}\sigma_{\tau x_1}^*\left(\sigma_{x_1}^{*2}-\sigma_{\tau_1}^{*-2}\sigma_{\tau x_1}^{*2}\right)\left(\sigma_{\tau_1}^* I_1^*+\mu_\tau^* I_0^*\right)+3\sigma_{\tau_2}^{*-2}\sigma_{\tau x_2}^*\left(\sigma_{x_2}^{*2}-\sigma_{\tau_2}^{*-2}\sigma_{\tau x_2}^{*2}\right)\left(\sigma_{\tau_2}^* I_1^{**}+\mu_\tau^* I_0^{**}\right)]\\
&+(\sigma_{\tau_1}^{*-2}\sigma_{\tau x_1}^*)^4\left(\sigma_{\tau_1}^{*4}I_4^*+4\mu_\tau^*\sigma_{\tau_1}^{*3}I_3^*+6\mu_\tau^{*2}\sigma_{\tau_1}^{*2}I_2^*+4\mu_\tau^{*3}\sigma_{\tau_1}^* I_1^*+\mu_\tau^{*4}I_0^*\right)+(\sigma_{\tau_2}^{*-2}\sigma_{\tau x_2}^*)^4(\sigma_{\tau 2}^{*4}I_4^{**}\\
&+4\mu_\tau^*\sigma_{\tau_2}^{*3}I_3^{**}+6\mu_\tau^{*2}\sigma_{\tau_2}^{*2}I_2^{**}+4\mu_\tau^{*3}\sigma_{\tau_2}^* I_1^{**}+\mu_\tau^{*4}I_0^{**})+6(\sigma_{\tau_1}^{*-2}\sigma_{\tau x_1}^*)^2\left(\sigma_{x_1}^{*2}-\sigma_{\tau_1}^{*-2}\sigma_{\tau x_1}^{*2}\right)(\sigma_{\tau_1}^{*2}I_2^*\\
&+2\mu_\tau^*\sigma_{\tau_1}^* I_1^*+\mu_\tau^{*2}I_0^*)+6(\sigma_{\tau_2}^{*-2}\sigma_{\tau x_2}^*)^2\left(\sigma_{x_2}^{*2}-\sigma_{\tau_2}^{*-2}\sigma_{\tau x_2}^{*2}\right)\left(\sigma_{\tau_2}^{*2}I_2^{**}+2\mu_\tau^*\sigma_{\tau_2}^* I_1^{**}+\mu_\tau^{*2}I_0^{**}\right)\\
&+3(\sigma_{x_1}^{*2}-\sigma_{\tau_1}^{*-2}\sigma_{\tau x_1}^{*2})^2 I_0^*+3(\sigma_{x_2}^{*2}-\sigma_{\tau_2}^{*-2}\sigma_{\tau x_2}^{*2})^2 I_0^{**}]
\end{aligned}
$$

where,

$$
\begin{aligned}
I_0^* &= \frac{1}{2}\left\{1-sign(h_1)\gamma(\frac{1}{2},\frac{h_1^2}{2})\right\},\\
I_1^* &= \frac{1}{\sqrt{2\pi}}\left\{1-sign(h_1)(-1)^{I_{h_1<0}}\gamma(\frac{1}{2},\frac{h_1^2}{2})\right\},\\
I_2^* &= \frac{1}{2}\left\{1-sign(h_1)\gamma(\frac{3}{2},\frac{h_1^2}{2})\right\},\\
I_3^* &= \frac{2}{\sqrt{2\pi}}\left\{1-sign(h_1)(-1)^{3I_{h_1<0}}\gamma(2,\frac{h_1^2}{2})\right\}\\
I_4^* &= \frac{3}{2}\left\{1-sign(h_1)\gamma(\frac{5}{2},\frac{h_1^2}{2})\right\}\\
I_0^{**} &= \frac{1}{2}\left\{1+\gamma(\frac{1}{2},\frac{h_2^2}{2})\right\}\\
I_1^{**} &= \frac{1}{\sqrt{2\pi}}\left\{-1+(-1)^{I_{h_2<0}}\gamma(1,\frac{h_2^2}{2})\right\}\\
I_2^{**} &= \frac{1}{2}\left\{1+\gamma(\frac{3}{2},\frac{h_2^2}{2})\right\}\\
I_3^{**} &= \frac{2}{\sqrt{2\pi}}\left\{-1+(-1)^{3I_{h_2<0}}\gamma(2,\frac{h_2^2}{2})\right\}\\
I_4^{**} &= \frac{3}{2}\left\{1+\gamma(\frac{5}{2},\frac{h_2^2}{2})\right\}
\end{aligned}
$$

End of the proof■

Appendix 4

The moment generating function of (5) is given by

$$\begin{aligned}
M(\theta) &= \int \exp(\theta' Z) f(Z|_{\mu,\Omega_1,\Omega_2,}) dZ, \\
&= c\int_{-\infty}^{+\infty}\int_{x}^{+\infty} \exp(\theta' Z) \frac{\exp\left\{-\dfrac{(Z-\mu)'\,\Omega_1^{-1}\,(Z-\mu)}{2}\right\}}{2\pi\left|\Omega_1\right|^{\frac{1}{2}}} d\tau dx \\
&+ c\int_{-\infty}^{+\infty}\int_{-\infty}^{x} \exp(\theta' Z) \frac{\exp\left\{-\dfrac{(Z-\mu)'\,\Omega_2^{-1}\,(Z-\mu)}{2}\right\}}{2\pi\left|\Omega_2\right|^{\frac{1}{2}}} d\tau dx, \\
&= c(I_1(\theta)+I_2(\theta)),
\end{aligned}$$

where

$$\left\{\begin{array}{c}
I_1(\theta) = \int_{-\infty}^{+\infty}\int_{x}^{+\infty} \exp(\theta' Z) \dfrac{\exp\left\{-\dfrac{(Z-\mu)'\,\Omega_1^{-1}\,(Z-\mu)}{2}\right\}}{2\pi\left|\Omega_1\right|^{\frac{1}{2}}} d\tau dx \\
I_2(\theta) = \int_{-\infty}^{+\infty}\int_{-\infty}^{x} \exp(\theta' Z) \dfrac{\exp\left\{-\dfrac{(Z-\mu)'\,\Omega_2^{-1}\,(Z-\mu)}{2}\right\}}{2\pi\left|\Omega_2\right|^{\frac{1}{2}}} d\tau dx
\end{array}\right. .$$

and $\theta = \left(\theta_\tau, \theta_x'\right)'$.

Now let

$$\Upsilon = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

$$Z = \Upsilon Y.$$

then,

$$Y = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} Z,$$

$$\begin{aligned}
\mu^* &= \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} (\mu + \Omega_1 \theta) \\
&= \begin{pmatrix} \mu_x + \theta_x \sigma_{x1}^2 + \theta_\tau \sigma_{\tau x1} \\ \mu_\tau - \mu_x + \theta_x\left(\sigma_{\tau x1} - \sigma_{x1}^2\right) + \theta_\tau\left(\sigma_{\tau 1}^2 - \sigma_{\tau x1}\right) \end{pmatrix},
\end{aligned}$$

$$
\begin{aligned}
\Omega_1^* &= \left(\Upsilon'\Omega_1^{-1}\Upsilon\right)^{-1} \\
&= \Upsilon^{-1}\Omega_1\Upsilon'^{-1} \\
&= \begin{pmatrix} \sigma_{x1}^2 & \sigma_{\tau x1}-\sigma_{x1}^2 \\ \sigma_{\tau x1}-\sigma_{x1}^2 & \sigma_{\tau 1}^2+\sigma_{x1}^2-2\sigma_{\tau x1} \end{pmatrix},
\end{aligned}
$$

$$
\begin{aligned}
\left(\Upsilon Y-\mu\right)'\Omega_1^{-1}\left(\Upsilon Y-\mu\right) &= \left(\Upsilon Y-\mu-\Omega_1\theta\right)'\Omega_1^{-1}\left(\Upsilon Y-\mu-\Omega_1\theta\right) \\
&- 2\theta'\mu-\theta'\Omega_1\theta+2\theta'\Upsilon Y \\
&= \left(Y-\mu^*\right)'\Omega_1^{*-1}\left(Y-\mu^*\right)-2\theta'\mu \\
&- \theta'\Omega_1\theta+2\theta'\Upsilon Y \\
&= \frac{\left(y_{\tau+1}-\mu_\tau^*\right)^2}{\sigma_{\tau 1}^{*2}}+\frac{\left(y_{x+1}-\mu_x^*+\varpi\right)^2}{\sigma_{x1}^{*2}-\sigma_{\tau 1}^{*-2}\sigma_{\tau x1}^{*2}}-2\theta'\mu \\
&- \theta'\Omega_1\theta+2\theta'\Upsilon Y,
\end{aligned}
$$

and

$$
\begin{aligned}
I_1(\theta) &= \int_{-\infty}^{+\infty}\int_x^{+\infty}\exp(\theta' Z)\frac{\exp\left\{-\dfrac{\left(Z-\mu\right)'\Omega_1^{-1}\left(Z-\mu\right)}{2}\right\}}{2\pi\left|\Omega_1\right|^{\frac{1}{2}}}d\tau dx \\
&= \int_{-\infty}^{+\infty}\int_0^{+\infty}\frac{\exp\left\{\theta'\Upsilon Y-\dfrac{\left(\Upsilon Y-\mu\right)'\Omega_1^{-1}\left(\Upsilon Y-\mu\right)}{2}\right\}}{2\pi\left|\Omega_1\right|^{\frac{1}{2}}}dy_{\tau+1}dy_{x+1} \\
&= \exp\left(\theta'\mu+\frac{\theta'\Omega_1\theta}{2}\right)\int_{-\infty}^{+\infty}\int_0^{+\infty}\frac{\exp\left\{-\dfrac{\left(y_{\tau+1}-\mu_\tau^*\right)^2}{2\sigma_{\tau 1}^{*2}}-\dfrac{\left(y_{x+1}-\mu_x^*+\varpi\right)^2}{2\left(\sigma_{x1}^{*2}-\sigma_{\tau 1}^{*-2}\sigma_{\tau x1}^{*2}\right)}\right\}}{2\pi\left|\Omega_1\right|^{\frac{1}{2}}}dy_{\tau+1}dy_{x+1} \\
&= \exp\left(\theta'\mu+\frac{\theta'\Omega_1\theta}{2}\right)\Phi\left(\frac{\mu_\tau^*}{\sigma_{\tau 1}^*}\right).
\end{aligned}
$$

where $\varpi=-\sigma_{\tau_1}^{*-2}\sigma_{\tau x_1}^*y_{\tau+1}$, $\mu_\tau^*=\mu_\tau-\mu_x+\theta_x\left(\sigma_{\tau x_1}-\sigma_{x_1}^2\right)+\theta_\tau\left(\sigma_{\tau_1}^2-\sigma_{\tau x_1}\right)$ and $\sigma_{\tau_1}^*=\sqrt{\sigma_{\tau_1}^2+\sigma_{x_1}^2-2\sigma_{\tau x_1}}$. The last equality follows from the fact that $|\Omega_1|=|\Omega_1^*|$ and from the properties of the normal distribution and where $\Phi\left(.\right)$ is the cumulative of the standard normal distribution.

Similarly, it can be shown that

$$
I_2(\theta)=\exp\left(\theta'\mu+\frac{\theta'\Omega_2\theta}{2}\right)\Phi\left(-\frac{\mu_\tau^*}{\sigma_{\tau_2}^*}\right),
$$

where $\sigma^*_{\tau_2} = \sqrt{\sigma^2_{\tau_2} + \sigma^2_{x_2} - 2\sigma_{\tau x_2}}$.

Therefore, the moment generating function is given by

$$M(\theta) = c\left\{\exp\left(\theta'\mu + \frac{\theta'\Omega_1\theta}{2}\right)\Phi\left(\frac{\mu^*_\tau}{\sigma^*_{\tau_1}}\right) + \exp\left(\theta'\mu + \frac{\theta'\Omega_2\theta}{2}\right)\Phi\left(-\frac{\mu^*_\tau}{\sigma^*_{\tau_2}}\right)\right\},$$

By the properties of the moment generating function we have

$$M(\theta = 0) = c(I_1(\theta = 0) + I_2(\theta = 0)) = 1,$$

where

$$I_1(\theta = 0) = \Phi\left(-\Delta_1\right),$$

and

$$I_2(\theta = 0) = \Phi\left(\Delta_2\right).$$

Thus

$$c = \frac{1}{\Phi\left(-\Delta_1\right) + \Phi\left(\Delta_2\right)}.$$

End of the proof■