Scientific
Research
Publishing

Generalized Method of Moments and Generalized Estimating Functions Using Characteristic Function

Andrew Luong

École d'actuariat, Université Laval, Ste Foy, Québec, Canada
Email: andrew.luong@act.ulaval.ca

Abstract

GMM inference procedures based on the square of the modulus of the model characteristic function are developed using sample moments selected using estimating function theory and bypassing the use of empirical characteristic function of other GMM procedures in the literature. The procedures are relatively simple to implement and are less simulation-oriented than simulated methods of inferences yet have the potential of good efficiencies for models with densities without closed form. The procedures also yield better estimators than method of moment estimators for models with more than three parameters as higher order sample moments tend to be unstable.

1. Introduction and an Overview of GMM Procedures Based on Empirical Characteristic Function

1.1. Introduction

In many applied fields, data analysts often have to use distributions with density functions having complicated forms. They are often expressed using mean of series representations but model characteristic functions are simpler and have closed form expressions. For actuarial sciences, the compound Poisson distributions are classical examples and for finance, the stable distributions fall into the same category. These are infinitely divisible and many infinitely divisible distributions share the same property of having much simpler characteristic functions

than density functions. We shall examine in more details using the Generalized Normal Laplace (GNL) distribution which is obtained by adding a normal component to the GAL random variable hence can be viewed as created by a convolution operation. The GNL distribution was introduced by Reed [1] and we shall use it to motivate inferences procedures based on characteristic functions instead of densities. Both the GNL and GAL distributions provide better fit to log returns data in finance. The density of the GNL distribution is more complicated than the density of the GAL distribution. The book by Kotz *et al.* [2] gives a very comprehensive account of the GAL distribution. Obviously, if these distributions with properties just mentioned are used for modelling, we still want to be able to estimate the parameters and perform tests for validating the models used. Often maximum likelihood (ML) procedures are difficult to implement due to the lack of closed form for the density functions for the models being used and even the ML estimators are available, when they are used with the Pearson chi-square statistics in general do not lead to distribution free statistics further complicate ML procedures. Therefore, it is natural that we aim at a unified approach to estimation and testing. Inferences developed in this paper will be unified using GMM approach but with the use of estimating function theory to select sample moments using moment conditions extracted from model characteristic function or more precisely the square of its modulus for constructing the GMM objective function. Subsequently, estimation and testing can be carried out. Before giving more details of the developed GMM procedures of this paper and how they differ from GMM procedures in the literature and the advantages of the new procedures, we shall give more details about the GNL distribution where the use of characteristic function appears to be more natural than the use of the model density. The GMM methods developed are also less simulation intensive than simulated methods which appear in the paper by Luong and Bilodeau [3] and faster in computing time for implementing.

The GNL generalizes the GAL distribution, the density of the GAL can be obtained in closed form but depend on Bessel functions, see Kotz *et al.* [2] (page 189), Luong [4] and since the GAL distribution can also be obtained from the distribution of the difference of two gamma random variables, we shall consider first the characteristic function of the gamma distribution in example 1 and subsequently in example 2 and example 3, we shall consider respectively the characteristic function of the GAL and GNL distributions.

First, recall that the characteristic function $\phi(s)$ of a random variable X is a complex function defined as

$$\phi(s) = E\left(e^{isX}\right)$$

and it can be expressed as

$$\phi(s) = Re\phi(s) + iIm\phi(s)$$

with the real and imaginary parts of $\phi(s)$ given respectively by $Re\phi(s)$ and $Im\phi(s)$. We can also use polar forms instead of algebraical forms to express

complex numbers or functions.

The modulus of $\phi(s)$ is defined as

$$\left|\phi(s)\right| = \left[\left(Re\phi(s)\right)^2 + \left(Im\phi(s)\right)^2\right]^{1/2}$$

and the argument $\omega(s)$ of $\phi(s)$ is defined as $\omega(s) = \arctan\dfrac{Im\phi(s)}{Re\phi(s)}$. This allows to express $\phi(s) = \left|\phi(s)\right|e^{i\omega(s)}$ and depending on situations, the polar form of $\phi(s)$ can be simpler to handle than its algebraical form as illustrated in example 1 which gives the characteristic function of the gamma random variable using both representations for the gamma distribution.

Example 1 (characteristic function of the gamma distribution)

It is well known than the characteristic function of the Gamma distribution in algebraical form is given by

$$\phi_{\boldsymbol{\gamma}}(s) = (1 - i\beta s)^{-\rho}, \quad \boldsymbol{\gamma} = (\rho, \beta)'$$

with β being the scale parameter and ρ being the shape parameter, $\beta > 0$ and $\rho > 0$. Now before giving the polar form of $\phi(s)$, we give the polar form of $z(s) = 1 - i\beta s$ first and using properties of the modulus as well as properties of the argument of a complex number we then give the polar form for $\phi_{\boldsymbol{\gamma}}(s)$. The modulus of $z(s)$ is denoted by

$$\left|z(s)\right| = \left[1 + \beta^2 s^2\right]^{1/2}$$

and the argument of $\arg(z(s)) = \arctan(-\beta s)$ and since the function $\arctan(x)$ is odd, $\arg(z(s)) = -\arctan(\beta s)$. This allows the representation in polar form for $z(s)$.

Using properties of the modulus, since $\phi(s) = (z(s))^{-\rho}$, so $\left|\phi(s)\right| = \left|z(s)\right|^{-\rho}$. With $z(s) = \left[1 + \beta^2 s^2\right]^{1/2} e^{i(-\arctan(\beta s))}$ and since $\phi(s) = (z(s))^{-\rho}$, it is easy to see that the characteristic function of the gamma distribution is given by $\phi_{\boldsymbol{\gamma}}(s) = \left[1 + \beta^2 s^2\right]^{-\rho/2} e^{i(\rho\arctan(\beta s))}$, $\boldsymbol{\gamma} = (\beta, \rho)'$ using polar form.

Using the characteristic function of the gamma distribution in polar form, we can find the characteristic function of the GAL distribution which can be considered as the difference of two independent gamma random variables.

Example 2 (characteristic function of the GAL distribution)

Among many representations in distribution of the GAL distribution, the one which makes use of two independent random gamma random variables allows the following representation for the GAL random variable X, see proposition 4.1.3 given by Kotz et al. [2] (p 183), $X =^d \theta + \dfrac{\sigma}{\kappa\sqrt{2}}G_1 - \dfrac{\sigma\kappa}{\sqrt{2}}G_2$ with $=^d$ being an equality in distribution, G_1 and G_2 are independent and identically distributed as G which follows a gamma distribution with scale parameter equals to one and shape parameter being ρ, the parameter θ is a location parameter with $-\infty < \theta < \infty$. The parameter κ controls the skewness of the GAL distribution, $\kappa > 0$ and if $\kappa = 1$, the distribution is symmetric. The parameter σ

is a scale parameter with $\sigma > 0$. Using the representation with gamma random variables it is easy to see that by letting

$$\frac{1}{\alpha} = \frac{\sigma}{\kappa\sqrt{2}}, \frac{1}{\beta} = \frac{\sigma\kappa}{\sqrt{2}} \quad \text{and} \quad X =^d \theta + \frac{1}{\alpha}G_1 - \frac{1}{\beta}G_2, \quad \kappa = 1 \text{ if } \alpha = \beta.$$

The characteristic function $\phi(s)$ for the GAL distribution in polar form is given by

$$\phi_{\boldsymbol{\gamma}}(s) = \left(1 + \frac{s^2}{\alpha^2}\right)^{-\rho/2} \left(1 + \frac{s^2}{\beta^2}\right)^{-\rho/2} e^{i(\theta + \rho\omega_1(s) + \rho\omega_2(s))}, \quad \boldsymbol{\gamma} = (\theta, \alpha, \beta, \rho)'$$

$\omega_1(s) = \arctan\left(\frac{s}{\alpha}\right)$, $\omega_2(s) = \arctan\left(-\frac{s}{\beta}\right)$, using the characteristic function of

the gamma distribution in polar form as given by example 1. Instead of using, replace it by $\rho\mu$ then

$$\phi_{\boldsymbol{\gamma}}(s) = \left(1 + \frac{s^2}{\alpha^2}\right)^{-\rho/2} \left(1 + \frac{s^2}{\beta^2}\right)^{-\rho/2} e^{i\rho(\mu + \omega_1(s) + \rho\omega_2(s))}, \quad \boldsymbol{\gamma} = (\mu, \alpha, \beta, \rho)'$$

Using this parametrization, it is easier to connect with the GNL distribution with the representation of the GNL random variable as the convolution of a normal random variable with a GAL random variable. The GAL is symmetric if $\alpha = \beta$ and its characteristic function can be further simplified and reduced to

$$\phi_{\boldsymbol{\gamma}}(s) = \left(1 + \frac{s^2}{\alpha^2}\right)^{-\rho} e^{i\rho\mu} \quad \text{or} \quad \phi_{\boldsymbol{\gamma}}(s) = \left(1 + \frac{s^2}{\alpha^2}\right)^{-\rho} e^{i\theta}.$$

Observe that often it is relatively simple to find characteristic function of a distribution of a convolution of two independent random variables using characteristic functions of the component independent random variables and also the characteristic function of the GAL distribution does not depend on the Bessel functions and is much simpler than its density despite the density of the GAL density has closed form expression. The GNL random variable X can be created by adding an independent normal random variable to a GAL random variable and allows the following representation as introduced by Reed [1] (p 475),

$$X =^d \sigma\sqrt{\rho}Z + \rho\mu + \frac{1}{\alpha}G_1 - \frac{1}{\beta}G_2$$

or equivalently,

$$X =^d \sigma\sqrt{\rho}Z + \theta + \frac{1}{\alpha}G_1 - \frac{1}{\beta}G_2.$$

Z is a standard normal random variable and independent of G_1 and G_2 with G_1 and G_2 are as defined as in example 2. Since the characteristic function for the standard normal random variable is $e^{\frac{1}{2}s^2}$ and the characteristic function of the GAL distribution is already obtained, the polar form of the GNL distribution can be also obtained and it is given in the following example.

Example 3 (characteristic function of the GNL distribution)

From the representation of the GNL random variable it is easy to see that the characteristic function of the GNL distribution in algebraical form is

$$\phi_{\gamma}(s) = e^{\rho i \mu s - \rho \sigma^2 \frac{s^2}{2}} \left(\frac{1}{1 - is/\alpha}\right)^{\rho} \left(\frac{1}{1 + is/\beta}\right)^{\rho}$$

which is given by Reed [1] (p 474) and in polar form

$$\phi_{\gamma}(s) = e^{-\rho \sigma^2 \frac{s^2}{2}} \left(1 + \frac{s^2}{\alpha^2}\right)^{-\rho/2} \left(1 + \frac{s^2}{\beta^2}\right)^{-\rho/2} e^{i\rho(\mu + \omega_1(s) + \rho \omega_2(s))}$$

with $\omega_1(s)$ and $\omega_2(s)$ are as defined in example 2.

Using the modulus of $\phi_{\gamma}(s)$,

$$\left|\phi_{\gamma}(s)\right| = e^{-\rho \sigma^2 \frac{s^2}{2}} \left(1 + \frac{s^2}{\alpha^2}\right)^{-\rho/2} \left(1 + \frac{s^2}{\beta^2}\right)^{-\rho/2},$$

we also have

$$\phi_{\gamma}(s) = \left|\phi_{\gamma}(s)\right| e^{i\rho(\mu + \omega_1(s) + \rho \omega_2(s))}, \quad \gamma = (\mu, \alpha, \beta, \sigma^2, \rho)'$$

As for the GAL distribution, if $\alpha = \beta$, the GNL distribution is symmetric and its characteristic function is further simplified and is given by

$$\phi_{\gamma}(s) = e^{-\rho \sigma^2 \frac{s^2}{2}} \left(1 + \frac{s^2}{\alpha^2}\right)^{-\rho} e^{i\theta} \quad \text{with} \quad \theta = \rho \mu$$

Reed [1] using the characteristic function also established expressions for the k-th cumulants $c_k, k = 1, 2, \cdots$ with

$$c_1 = \rho\left(\mu + \frac{1}{\alpha} - \frac{1}{\beta}\right), c_2 = \rho\left(\sigma^2 + \frac{1}{\alpha^2} + \frac{1}{\beta^2}\right),$$

$$c_r = \rho(r-1)!\left(\frac{1}{\alpha^2} + (-1)^r \frac{1}{\beta^r}\right), r > 3$$

The GNL distribution provides a better fit to log returns data than the GAL distribution and both these distributions provide much better fit to log returns data than the normal distribution. In addition, all integer moments exist for these distributions and they are also infinitely divisible like the normal distribution which makes them being good alternatives to the normal distribution. From the characteristic function of the GNL distribution, it is easy to see that the real and imaginary part of the characteristic function are given respectively as

$$Re\phi_{\gamma}(s) = \left|\phi_{\gamma}(s)\right| \cos\left(\rho\left(\mu + \omega_1(s) + \rho \omega_2(s)\right)\right)$$

and

$$Im\phi_{\gamma}(s) = \left|\phi_{\gamma}(s)\right| \sin\left(\rho\left(\mu + \omega_1(s) + \rho \omega_2(s)\right)\right).$$

1.2. Empirical Characteristic Function and GMM Procedures in the Literature

For inferences, we assume that we have a random sample of size n which con-

sists of $X_1, \cdots, X_n$ of independent and identically distributed continuous random variables and they are distributed as X, with common characteristic function $\phi_\gamma(s)$, γ is a p by 1 vector of parameters of interests with $\gamma = (\gamma_1, \cdots, \gamma_p)'$, γ_0 is the vector of the true parameters with $\gamma_0 \in \Omega$, the parameter space is assumed to be compact. The number of parameters in the model is p. In fact, most inferences procedures based on characteristic function proposed in the literature are still valid if X has a discontinuity point with mass attributed at the origin such as in the cases of the compound distributions. If X is discrete, it is often preferred to work with probability generating function rather characteristic function and for related procedures using probability generating function, see Luong [5].

Commonly proposed GMM procedures in the literature are based on the empirical characteristic function which is the counterpart of the theoretical one and it is defined as

$$\phi_n(s) = Re\phi_n(s) + iIm\phi_n(s)$$

with the real and imaginary parts given respectively by

$$Re\phi_n(s) = \frac{1}{n}\sum_{i=1}^{n}\cos(sX_i) \quad \text{and} \quad Im\phi_n(s) = \frac{1}{n}\sum_{i=1}^{n}\cos(sX_i).$$

For example, the K-L procedures proposed by Feuerverger and McDunnough [6] p 20-23) can be viewed as equivalent to GMM procedures based on $2k$ sample moments of the following forms with the first k sample moments which make use respectively the chosen points $s_1, \cdots, s_k$ from the real part of empirical characteristic function and the real part of the model characteristic function

$$\boldsymbol{g}_1(\boldsymbol{\gamma}) = \left(\frac{1}{n}\sum_{i=1}^{n}\cos(s_1 X_i) - Re\phi_\gamma(s_1), \cdots, \frac{1}{n}\sum_{i=1}^{n}\cos(s_k X_i) - Re\phi_\gamma(s_k)\right)'$$

and the rest of moments are similarly formed but based on the imaginary part of the empirical characteristic function and the imaginary part of the model characteristic function,

$$\boldsymbol{g}_2(\boldsymbol{\gamma}) = \left(\frac{1}{n}\sum_{i=1}^{n}\sin(s_1 X_i) - Im\phi_\gamma(s_1), \cdots, \frac{1}{n}\sum_{i=1}^{n}\sin(s_k X_i) - Im\phi_\gamma(s_k)\right)'.$$

By letting

$$\boldsymbol{g}(\boldsymbol{\gamma}) = \begin{pmatrix} \boldsymbol{g}_1(\boldsymbol{\gamma}) \\ \boldsymbol{g}_2(\boldsymbol{\gamma}) \end{pmatrix}$$

and define S to be the limit covariance matrix of the vector $\sqrt{n}\boldsymbol{g}(\boldsymbol{\gamma})$ under the true parameter γ_0 when $n \to \infty$ and let $\hat{\boldsymbol{S}}$ be a preliminary consistent estimate of $\boldsymbol{S}$ and from which we can obtain a preliminary consistent estimate $\hat{\boldsymbol{S}}^{-1}$ for the inverse of S then the related GMM objective function $Q(\boldsymbol{\gamma})$ can be formed, *i.e.*, $Q(\boldsymbol{\gamma}) = \boldsymbol{g}'(\boldsymbol{\gamma})\hat{\boldsymbol{S}}^{-1}\boldsymbol{g}(\boldsymbol{\gamma})$ and minimizing $Q(\boldsymbol{\gamma})$ will give the vector of K-L estimators.

The following expectation properties are quite obvious and the elements of the covariance matrix for $\boldsymbol{g}(\boldsymbol{\gamma})$ can be found explicitly using the following

identities which are established using properties of trigonometric functions, we have:

$$E\left(Re\phi_n\left(s\right)\right)=Re\phi_\gamma\left(s\right),\quad E\left(Im\phi_n\left(s\right)\right)=Im\phi_\gamma\left(s\right),$$

$$Cov\left(Re\phi_n\left(s\right),Re\phi_n\left(t\right)\right)=\left(Re\phi_\gamma\left(s+t\right)+Re\phi_\gamma\left(t-s\right)-2Re\phi_\gamma\left(s\right)Re\phi_\gamma\left(t\right)\right)\big/2n,$$

$$Cov\left(Im\phi_n\left(s\right),Im\phi_n\left(t\right)\right)=\left(Re\phi_\gamma\left(t-s\right)+Re\phi_\gamma\left(t+s\right)-2Im\phi_\gamma\left(s\right)Im\phi_\gamma\left(t\right)\right)\big/2n,$$

$$Cov\left(Re\phi_n\left(t\right),Im\phi_n\left(s\right)\right)=\left(Im\phi_\gamma\left(t+s\right)-Im\phi_\gamma\left(t-s\right)-2Im\phi_\gamma\left(s\right)Re\phi_\gamma\left(t\right)\right)\big/2n \quad (1)$$

The above identities are results of Proposition 3.1 given by Groparu-Cojocaru and Doray [7] (p 1992) or results from Koutrouvelis [8] (p 919).

Observe that for the K-L procedures or GMM procedures based on the above $2k$ sample moments, we need to fix the points $s_1,\cdots,s_k$ where we can make use of the real and imaginary part of the model characteristic function $\phi_\gamma\left(s\right)$ and there is still a lack of general criteria on how to choose these points, see discussions by Tran [9] but it is recommended that these points are equally spaced, *i.e.*, consider $s_j=j\tau$ and the optimum choice for τ has the property that $\tau\to0$ when $k\to\infty$.

Koutrouvelis [8] (p 920) has shown that in general, since the variances of $\cos\left(sX\right)$ and $\sin\left(sX\right)$ have the following properties

$$V\left(\cos\left(sX\right)\right)=\sqrt{n}V\left(Re\phi_n\left(s\right)\right)\to0\quad\text{and}\quad V\left(\sin\left(sX\right)\right)=\sqrt{n}V\left(Im\phi_n\left(s\right)\right)\to0$$

as $s\to0$.

and argued that we should select points in the range of $\left(0,\pi\right)$ as points near 0 that we need to focus when extracting information from the model characteristic functions. Despite that the K-L procedures have good potentials for generating good efficiencies for estimators but it is often numerical difficult to implement, as the studies of Groparu-Cojocaru and Doray [7] (p 1996) have shown that in practice we need at least $k\geq10$ which means that at least 20 sample moments are needed for the procedures to have good efficiency and in these situations, the matrices S and $\hat{S}$ are often nearly singular and inverting such large matrix often create difficulties and we shall see that GMM procedures proposed in this paper with the use of theory of estimation function to select sample moments will only need a number of sample moments which is less than 10 in general instead of at least 20. In addition, the number of points from the model characteristic function used to construct sample moments also goes to infinity as the sample size $n\to\infty$.

The proposed GMM procedures with the selection of the sample moments based on estimating function theory will be developed in the next section. With the original sample, we also transform it to a sample of *n* observations which are still independent and we work with the original sample and the transformed sample to construct moment conditions.

Carrasco and Florens [10] have introduced GMM methods with a continuum of moment conditions and Carrasco and Kotchoni [11] have used the empirical characteristic function and developed GMM procedures based on objective

functions which match the empirical characteristic function with the model counterpart using points which belong to a continuum interval. Using a continuum of moment conditions is a solution for the arbitrariness choice of selecting points of the characteristic functions to extract information but the procedures might be difficult to implement for practitioners meanwhile our procedures remain simple and closer to the classical GMM procedures with a finite number of moments but we shall use estimation theory to select sample moments and the number of points will be selected equally spaced in the interval $(0, \pi)$ and the number of points will go to infinity as $n \to \infty$.

In fact, the points s_i of our procedures are selected with

$$s_i = \frac{\pi}{n}\left(i - \frac{1}{2}\right), i = 1, \cdots, n$$

and observe that the spacing used is $\frac{\pi}{n} \to 0$, as $n \to \infty$ and also observe that the spacing mimics the behavior of the optimum spacing and numerically it bypasses the difficulties of having to find explicitly the value of the optimum spacing by minimizing the determinant of the asymptotic covariance of the K-L estimators if the K-L procedures are used.

For the proposed methods, we need the additional assumption that the first four integer moments of the model distribution exist but in practical situations, this assumption is often met.

The proposed procedures make use of sample moments which focus on extracting information from the square of the modulus of the characteristic function $\left|\phi_\gamma(s)\right|^2$ using the points $s_i = \frac{\pi}{n}\left(i - \frac{1}{2}\right), i = 1, \cdots, n$ and clearly there will be as many points as the sample size.

For model, with a location parameter μ, the modulus $\left|\phi_\gamma(s)\right|$ and consequently $\left|\phi_\gamma(s)\right|^2$ will not depend on the location parameter μ and we need another two sample moments beside the sample moments which make use of $\left|\phi_\gamma(s)\right|^2$ to take care of this situation. The example given below will help to clarify the problem that we might encounter when the modulus $\left|\phi_\gamma(s)\right|$ or the square of the modulus $\left|\phi_\gamma(s)\right|^2$ is used for inferences.

For the normal distribution with the vector of parameters $\gamma = \left(\mu, \sigma^2\right)'$, the characteristic function is

$$\phi_\gamma(s) = e^{i\mu t - \frac{\sigma^2 t^2}{2}}$$

and its modulus is

$$\left|\phi_\gamma(s)\right| = e^{-\frac{\sigma^2 t^2}{2}}$$

and the square of the modulus is

$$\left|\phi_\gamma(s)\right|^2 = e^{-\sigma^2 t^2},$$

The location parameter μ is missing in $\left|\phi_\gamma(s)\right|$ and consequently, in $\left|\phi_\gamma(s)\right|^2$.

This is to illustrate that there might be one parameter being left out if inferences procedures are solely based on $\left|\phi_{\boldsymbol{\gamma}}(s)\right|$ or $\left|\phi_{\boldsymbol{\gamma}}(s)\right|^2$. This also means that for GMM procedures which make use of sample moments formed using $\left|\phi_{\boldsymbol{\gamma}}(s)\right|$ or $\left|\phi_{\boldsymbol{\gamma}}(s)\right|^2$, there should also be other sample moments to take into account the parameters being left out and if there are parameters being left out, it only affects the location parameter of the model in general, so if we use two additional sample moments which take care of the parameters being left out beside the sample moments which make use of $\left|\phi_{\boldsymbol{\gamma}}(s)\right|^2$, the GMM procedures will be viable. As mentioned often at most there is one parameter in the model which is not included in $\left|\phi_{\boldsymbol{\gamma}}(s)\right|^2$ so the proposed will make use of additional moments which are based on the mean and variance of the model distribution beside the moments based on

$$\left|\phi_{\boldsymbol{\gamma}}(s)\right|^2 \quad \text{using} \quad s_i = \frac{\pi}{n}\left(i - \frac{1}{2}\right), i = 1, \cdots, n$$

We hope to achieve good efficiency yet preserve simplicity by not using more than ten sample moments, this achieved by using the theory of estimating function for building sample moments which make use of

$$\left|\phi_{\boldsymbol{\gamma}}(s)\right|^2 \quad \text{using} \quad s_i = \frac{\pi}{n}\left(i - \frac{1}{2}\right), i = 1, \cdots, n$$

Therefore, it is relatively simple to implement and all can be done within the classical context of GMM procedures without having to rely on a continuum of moment conditions which the practitioners might find difficult to implement. The use of theory of estimating function appears to be new and not included in proposed GMM procedures in the literature which focused on the use of the empirical characteristic function. The new procedures also make use of transformed observations besides the original observations.

The paper is organized as follows. Section 1 introduces the commonly used GMM procedures which are based on empirical characteristic function, the approach taken here does not use the empirical characteristic function and relies on estimating function theory to select sample moments based on the square of the modulus of the model characteristic function. The new GMM procedures are introduced in Section 2.1 with the choice of selected sample moments aiming to provide efficiency for GMM estimation. In Section 2.2 the chi-square test for moment restrictions which can be interpreted as goodness-of-fit is presented. In Section 3, illustrations for implementing the methods using the GNL distribution and normal distribution, the methods appear to be relatively simple to implement yet being very efficient based on the limited studies and appear to be better alternatives the method of moments (MOM) in general.

2. The Proposed GMM Procedures Based on Theory of Estimating Functions

2.1. Estimation

The theory of GMM procedures are well established in the literature, see Martin

et al. [12], Hayashi [13], Hamilton [14] but based on the assumption that sample moments are already selected. In this paper, we focus on how to select moments for model with characteristic function being simple and has closed form but the model density is complicated and we do not make use of the classical empirical characteristic function as other GMM procedures being proposed in the literature. Here, we focus on the square of the modulus of the model characteristic function to build sample moments and since the modulus might not include all the parameters of the model such as the case when there is a location parameter, we shall also include two moments which focus on the model distribution mean and variance to complete the set of sample moments and for practical applications, we do not need more than ten sample moments for the use of the proposed GMM procedures.

We shall define the sample moments focusing on $\left|\phi_{\boldsymbol{\gamma}}(s)\right|^2$. Let us consider the basic estimating functions

$$h(x_i, s_i; \boldsymbol{\gamma}) = Re\phi_{\boldsymbol{\gamma}}(s_i)\left(\cos(s_i X_i) - Re\phi_{\boldsymbol{\gamma}}(s_i)\right)$$
$$+ Im\phi_{\boldsymbol{\gamma}}(s_i)\left(\sin(s_i X_i) - Im\phi_{\boldsymbol{\gamma}}(s_i)\right), i = 1, \cdots, n$$

Clearly, the basic estimating functions are unbiased, *i.e.*,

$$E_{\boldsymbol{\gamma}}\left(h(x_i, s_i; \boldsymbol{\gamma})\right) = 0, i = 1, \cdots, n$$

Using $\left|\phi_{\boldsymbol{\gamma}}(s)\right|^2 = \left(Re\left|\phi_{\boldsymbol{\gamma}}(s)\right|^2\right)^2 + \left(Im\phi_{\boldsymbol{\gamma}}(s)\right)^2$, we can also express

$$h(x_i, s_i; \boldsymbol{\gamma}) = Re\phi_{\boldsymbol{\gamma}}(s_i)\left(\cos(s_i X_i)\right) + Im\phi_{\boldsymbol{\gamma}}(s_i)\left(\sin(s_i X_i)\right)$$
$$- \left|\phi_{\boldsymbol{\gamma}}(s_i)\right|^2, i = 1, \cdots, n \tag{2}$$

Now we can construct the optimum estimating functions for estimating $\boldsymbol{\gamma}$ or more precisely for parameters which appear in $\left|\phi_{\boldsymbol{\gamma}}(s)\right|^2$ using results of Godambe and Thompson [15] (p 139) or Morton [16] (p 229). The optimum estimating functions which are linear combinations of elements of the set $\left\{h(x_i, s_i; \boldsymbol{\gamma}), i = 1, \cdots, n\right\}$ and with $\boldsymbol{\gamma} = (\gamma_1, \cdots, \gamma_p)'$ can be expressed as

$$\frac{1}{n}\sum_{i=1}^{n} h(x_i, s_i; \boldsymbol{\gamma}) \frac{E_{\boldsymbol{\gamma}}\left(\dfrac{\partial h(x_i, s_i; \boldsymbol{\gamma})}{\partial \gamma_j}\right)}{v_{\boldsymbol{\gamma}}\left(h(x_i, s_i; \boldsymbol{\gamma})\right)}, j = 1, \cdots, p \tag{3}$$

and $v_{\boldsymbol{\gamma}}\left(h(x_i, s_i; \boldsymbol{\gamma})\right)$ denote the variance of $h(x_i, s_i; \boldsymbol{\gamma})$ and can be obtained explicitly, see expression (1).

We would like to make a few remarks here. First note that it is easy to show that

$$E_{\boldsymbol{\gamma}}\left(\frac{\partial h(x_i, s_i; \boldsymbol{\gamma})}{\partial \gamma_j}\right) = -\frac{1}{2}\frac{\partial \left|\phi_{\boldsymbol{\gamma}}(s)\right|^2}{\partial \gamma_j}$$

using $E_{\boldsymbol{\gamma}}\left(\cos(sX)\right) = Re\phi_{\boldsymbol{\gamma}}(s)$ and $E_{\boldsymbol{\gamma}}\left(\sin(sX)\right) = Im\phi_{\boldsymbol{\gamma}}(s)$ and clearly if there is one parameter of the model says γ_l which does not appear in $\left|\phi_{\boldsymbol{\gamma}}(s)\right|^2$ then $E_{\boldsymbol{\gamma}}\left(\dfrac{\partial h(x_i, s_i; \boldsymbol{\gamma})}{\partial \gamma_j}\right) = 0$ and there is no optimum estimating function for this pa-

rameter and if we want to estimate all the parameters, we need an extra estimating function. We use the following notation for the vector of optimum estimating function discarding the ones with $E_{\gamma}\left(\dfrac{\partial h\left(x_i, s_i; \gamma\right)}{\partial \gamma_j}\right) = 0$. The vector of optimum estimating functions for parameters included in $\left|\phi_{\gamma}\left(s\right)\right|^2$ adopting the convention discarding those with $E_{\gamma}\left(\dfrac{\partial h\left(x_i, s_i; \gamma\right)}{\partial \gamma_j}\right) = 0$ is given by

$$\frac{1}{n}\sum_{i=1}^{n} h\left(x_i, s_i; \gamma\right) \frac{E_{\gamma}\left(\dfrac{\partial h\left(x_i, s_i; \gamma\right)}{\partial \gamma_1}\right)}{v_{\gamma}\left(h\left(x_i, s_i; \gamma\right)\right)}$$

where we partition the vector γ into two components, $\gamma = \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix}$ with the property that all the parameters appear in $\left|\phi_{\gamma}\left(s\right)\right|^2$ form the vector γ_1 and all the remaining parameters are included in the vector γ_2. In general, if $\gamma \neq \gamma_1$ in then γ_2 is reduced to a scalar. Therefore, the vector of optimum estimating function in general is either a vector of p elements or $p-1$ elements and consequently when these estimating functions are converted to sample moments, we shall have either p or $p-1$ sample moments.

We shall let $g_1\left(\gamma\right)$ be vector the sample moments which make use of points of $\left|\phi_{\gamma}\left(s\right)\right|^2$ as

$$g_1\left(\gamma\right) = \frac{1}{n}\sum_{i=1}^{n} h\left(x_i, s_i; \gamma\right) \frac{E_{\gamma}\left(\dfrac{\partial h\left(x_i, s_i; \gamma\right)}{\partial \gamma_1}\right)}{v_{\gamma}\left(h\left(x_i, s_i; \gamma\right)\right)},$$

$v_{\gamma}\left(h\left(x_i, s_i; \gamma\right)\right)$ can be obtained using the real and imaginary parts of the model characteristic function and from the definition of $h\left(x_i, s_i; \gamma\right)$, it then follows that

$$\begin{aligned} v_{\gamma}\left(h\left(x_i, s_i; \gamma\right)\right) &= \left(Re\phi_{\gamma}\left(s_i\right)\right)^2 var\left(\cos\left(s_i X_i\right)\right) + \left(Im\phi_{\gamma}\left(s_i\right)\right)^2 var\left(\sin\left(s_i X_i\right)\right) \\ &\quad + 2\left(Re\phi_{\gamma}\left(s_i\right)\right)\left(Im\phi_{\gamma}\left(s_i\right)\right) cov\left(\cos\left(s_i X_i\right), \sin\left(s_i X_i\right)\right) \end{aligned} \tag{4}$$

Using the identities as given by expression (1), the variance of $\cos\left(s_i X_i\right)$ is

$$var\left(\cos\left(s_i X_i\right)\right) = \left(Re\phi_{\gamma}\left(2s_i\right) + 1 - 2\left(Re\phi_{\gamma}\left(s_i\right)\right)^2\right)\Big/2 \tag{5}$$

and the variance of $\sin\left(s_i X_i\right)$ and the covariance $cov\left(\cos\left(s_i X_i\right), \sin\left(s_i X_i\right)\right)$ are given respectively by

$$var\left(\sin\left(s_i X_i\right)\right) = \frac{\left(1 - Re\phi_{\gamma}\left(2s_i\right) - 2\left(Im\phi_{\gamma}\left(s_i\right)\right)^2\right)}{2}, \tag{6}$$

$$cov\left(\cos\left(s_i X_i\right), \sin\left(s_i X_i\right)\right) = \left(Im\phi_{\gamma}\left(2s_i\right) - 2\left(Re\phi_{\gamma}\left(s_i\right)\right)\left(Im\phi_{\gamma}\left(s_i\right)\right)\right)\Big/2. \tag{7}$$

These variances and covariance terms can also be obtained using results given by Groparu-Cojocaru and Doray [7] (p 1993). Asymptotic properties of estimators obtained by solving estimating equations have been given by Yuan and Jennrich

[17] but we emphasize GMM estimation in this paper.

Now define two additional sample moments $g_2(\gamma)$ and $g_3(\gamma)$

$$g_2(\gamma) = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - E_{\gamma}(X)\right),$$

The location parameter if it belongs to the model but is not included in γ_1, it will appear in $E_{\gamma}(X)$ which is the mean of the model distribution and it can be obtained by differentiating the model characteristic function and for GMM procedures we prefer to have the number of sample moments exceeding the number of parameters in the model, so we also consider the following sample moment which makes use of the variance of the model distribution $V_{\gamma}(X)$ and it can also be obtained by differentiating twice the model characteristic function, *i.e.*,

$$g_3(\gamma) = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - V_{\gamma}(X)\right).$$

The vector sample moments for the developed GMM procedures is given by

$$g(\gamma) = \begin{pmatrix} g_1(\gamma) \\ g_2(\gamma) \\ g_3(\gamma) \end{pmatrix}$$

and notice $g_1(\gamma)$ makes use of the transformed observations $s_i X_i, i = 1, \cdots, n$ but $g_2(\gamma)$ and $g_3(\gamma)$ make use of the original sample observations $X_i, i = 1, \cdots, n$.

The vector of the proposed GMM estimators $\hat{\gamma}$ is obtained by minimizing the criterion function,

$$Q(\gamma) = g'(\gamma)\hat{S}^{-1}g(\gamma)$$

with $\hat{S}^{-1}$ being a positive definite matrix with probability one and it will be defined subsequently after the definition of S and its inverse S^{-1}, $\hat{S}^{-1}$ is a consistent estimate of S^{-1}.

For finding elements of the matrix, we can first express the components of the vector of sample moments as

$$g(\gamma) = \begin{pmatrix} g_1(\gamma) \\ g_2(\gamma) \\ g_3(\gamma) \end{pmatrix}$$

with

$$g_1(\gamma) = \frac{1}{n}\sum_{i=1}^{n}m_1(x_i; s_i; \gamma), \quad m_1(x_i; s_i; \gamma) = h(x_i, s_i; \gamma)\frac{E_{\gamma}\left(\frac{\partial h(x_i, s_i; \gamma)}{\partial \gamma_1}\right)}{v_{\gamma}\left(h(x_i, s_i; \gamma)\right)},$$

$$g_2(\gamma) = \frac{1}{n}\sum_{i=1}^{n}m_2(x_i; \gamma), \quad m_2(x_i; \gamma) = X_i - E_{\gamma}(X),$$

$$g_3(\gamma) = \frac{1}{n}\sum_{i=1}^{n}m_3(x_i; \gamma), \quad m_3(x_i; \gamma) = \left(X_i - E_{\gamma}(X)\right)^2 - V_{\gamma}(X),$$

and let

$$\boldsymbol{m}\left(x_i;s_i,\boldsymbol{\gamma}\right) = \begin{pmatrix} \boldsymbol{m}_1\left(x_i;s_i,\boldsymbol{\gamma}\right) \\ m_2\left(x_i;s_i,\boldsymbol{\gamma}\right) \\ m_3\left(x_i;s_i,\boldsymbol{\gamma}\right) \end{pmatrix}.$$

The matrix $\boldsymbol{S}$ is defined as

$$\boldsymbol{S} = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} E_{\boldsymbol{\gamma}_0}\left(\boldsymbol{m}\left(x_i;s_i,\boldsymbol{\gamma}_0\right)\boldsymbol{m}'\left(x_i;s_i,\boldsymbol{\gamma}_0\right)\right).$$

Now if we have a preliminary consistent estimate $\boldsymbol{\gamma}^{(0)} = \begin{pmatrix} \boldsymbol{\gamma}_1^{(0)} \\ \boldsymbol{\gamma}_2^{(0)} \end{pmatrix}$ for the vector

$\boldsymbol{\gamma}_0$ then an estimate for $\boldsymbol{S}$ which is $\hat{\boldsymbol{S}}$ can be defined with

$$\hat{\boldsymbol{S}} = \frac{1}{n}\sum_{i=1}^{n}\left(\boldsymbol{m}\left(x_i;s_i,\boldsymbol{\gamma}^{(0)}\right)\boldsymbol{m}'\left(x_i;s_i,\boldsymbol{\gamma}^{(0)}\right)\right)$$

and its inverse which is $\hat{\boldsymbol{S}}^{-1}$ is a consistent estimate of $\boldsymbol{S}^{-1}$, often a numerical algorithm is used to minimize $Q(\boldsymbol{\gamma}) = \boldsymbol{g}'(\boldsymbol{\gamma})\hat{\boldsymbol{S}}^{-1}\boldsymbol{g}(\boldsymbol{\gamma})$, the vector of the GMM estimators is obtained after the convergence of a numerical iterative process, at each iteration we might want to readjust $\boldsymbol{\gamma}^{(0)}$ in a similar way as when performing an iteratively feasible nonlinear weighted least-squares procedures where the weights are re-estimated at each step of the iterations.

The vector of the GMM estimators $\hat{\boldsymbol{\gamma}}$ is consistent in general, this follows from general theory of GMM procedures, *i.e.*, we have $\hat{\boldsymbol{\gamma}}\xrightarrow{\ p\ }\boldsymbol{\gamma}_0$, with $\xrightarrow{\ p\ }$ denotes convergence in probability. In addition, under suitable differentiability imposed on the vector ($\boldsymbol{\gamma}$), the vector of GMM estimators $\hat{\boldsymbol{\gamma}}$ has an asymptotic multivariate normal distribution, *i.e.*,

$$\sqrt{n}\left(\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0\right)\xrightarrow{\ L\ }N\left(0,V\right),\ \ \boldsymbol{V} = \left(\boldsymbol{D}'\boldsymbol{S}^{-1}\boldsymbol{D}\right)^{-1}$$

$\xrightarrow{\ L\ }$ denotes convergence in law with

$$\boldsymbol{D} = \lim_{n\to\infty} E_{\boldsymbol{\gamma}_0}\left(\frac{\partial\boldsymbol{g}\left(\boldsymbol{\gamma}_0\right)}{\partial\boldsymbol{\gamma}'}\right)$$

which is a r by p matrix, r is the number of sample moments used or equivalently the number of elements of the vector $\boldsymbol{g}(\boldsymbol{\gamma})$. $\boldsymbol{D}$ can be estimated by

$$\hat{\boldsymbol{D}} = \frac{\partial\boldsymbol{g}\left(\hat{\boldsymbol{\gamma}}\right)}{\partial\boldsymbol{\gamma}'}$$

and V can be estimated by

$$\hat{\boldsymbol{V}} = \left(\hat{\boldsymbol{D}}'\hat{\boldsymbol{S}}^{-1}\hat{\boldsymbol{D}}\right)^{-1}$$

and consequently, the asymptotic variance of $\hat{\boldsymbol{\gamma}}$ can be estimated.

2.2. Testing Moment Restrictions

One of the advantages of GMM procedures is that it can lead to distribution free chi-square test. The asymptotic null distribution of the statistic no longer de-

pends on the parameters using statistic based on the same objective function to obtain the vector of GMM estimators $\hat{\boldsymbol{\gamma}}$. Howeuver, the test for moment restrictions when used as goodness-of-fit test, it might not be consistent as it might fail to detect departures from the specified model even as the sample size $n \to \infty$. The inclusion of some more relevant sample moments so that the number of sample moments used r is still manageable and as results of this will make the test more powerful yet without creating too many numerical difficulties on inverting the matrix $\hat{\boldsymbol{S}}$ in the same vein as for chi-square test for count models using probability generating function as given by Luong [17] is a topic for further studies.

For testing the moment restrictions

$$H_0 : E_{\boldsymbol{\gamma}}\left(m\left(X_i; s_i, \boldsymbol{\gamma}\right)\right) = 0.$$

Assuming we have already minimized $Q(\boldsymbol{\gamma}) = \boldsymbol{g}'(\boldsymbol{\gamma})\hat{\boldsymbol{S}}^{-1}\boldsymbol{g}(\boldsymbol{\gamma})$ to obtain $\hat{\boldsymbol{\gamma}}$, it follows from standard results of GMM theory that the following statistic can be used and the statistic has an asymptotic chi-square distribution with $r - p$ degrees of freedom, assuming $r > p$, *i.e.*,

$$nQ(\hat{\boldsymbol{\gamma}}) \xrightarrow{\ L\ } \chi^2_{r-p}. \tag{8}$$

3. Numerical Illustrations and Simulations

For illustrations of the newly developed methods, we shall examine the symmetric GNL distribution and compare the efficiencies of GMM estimators vs the efficiencies of method of moment estimators (MOM) as given by Reed [1] (p 47). The characteristic function of the symmetric GNL distribution only has 4 parameters as $\alpha = \beta$, it is easy to see that its characteristic function is reduced to

$$\phi_{\boldsymbol{\gamma}}(s) = \mathrm{e}^{-\rho\sigma^2 \frac{s^2}{2}}\left(1 + \frac{s^2}{\alpha^2}\right)^{-\rho} \mathrm{e}^{i\theta s}, \quad \boldsymbol{\gamma} = \left(\theta, \alpha, \sigma^2, \rho\right)'.$$

The location parameter instead of being $\rho\mu$ we can use $\theta = \rho\mu$ and Reed's MOM estimator for θ can be obtained independently of ρ.

It is not difficult to see that

$$Re\phi_{\boldsymbol{\gamma}}(s) = \mathrm{e}^{-\rho\sigma^2 \frac{s^2}{2}}\left(1 + \frac{s^2}{\alpha^2}\right)^{-\rho}\cos(\theta s), \tag{9}$$

$$Im\phi_{\boldsymbol{\gamma}}(s) = \mathrm{e}^{-\rho\sigma^2 \frac{s^2}{2}}\left(1 + \frac{s^2}{\alpha^2}\right)^{-\rho}\sin(\theta s), \tag{10}$$

The vector of parameters of the symmetric GNL distribution is $\boldsymbol{\gamma} = \left(\theta, \alpha, \sigma^2, \rho\right)'$. Since the model has a location parameter it is expected that the modulus of the characteristic function only has three parameters, it is easy to see that indeed this is the case with

$$\left|\phi_{\boldsymbol{\gamma}}(s)\right| = \mathrm{e}^{-\rho\sigma^2 \frac{s^2}{2}}\left(1 + \frac{s^2}{\alpha^2}\right)^{-\rho},$$

$$\left|\phi_{\pmb{\gamma}}\left(s\right)\right|^{2} = \mathrm{e}^{-\rho\sigma^{2}s^{2}}\left(1+\frac{s^{2}}{\alpha^{2}}\right)^{-2\rho}$$

which gives the following derivatives

$$\left(-\frac{1}{2}\frac{\partial\left|\phi_{\pmb{\gamma}}\left(s\right)\right|^{2}}{\partial\alpha}\right) = 2\left(\frac{s^{2}}{\alpha^{3}}\left(1+\frac{s^{2}}{\alpha^{2}}\right)^{-1}\right)\left|\phi_{\pmb{\gamma}}\left(s\right)\right|^{2},$$

$$\left(-\frac{1}{2}\frac{\partial\left|\phi_{\pmb{\gamma}}\left(s\right)\right|^{2}}{\partial\sigma^{2}}\right) = \left(\rho\frac{s^{2}}{2}\right)\left|\phi_{\pmb{\gamma}}\left(s\right)\right|^{2}$$

$$\left(-\frac{1}{2}\frac{\partial\left|\phi_{\pmb{\gamma}}\left(s\right)\right|^{2}}{\partial\rho}\right) = \frac{1}{2}\left|\phi_{\pmb{\gamma}}\left(s\right)\right|^{2}\left(\sigma^{2}s^{2}+2\ln\left(1+\frac{s^{2}}{\alpha^{2}}\right)\right)$$

and the variance of the basic estimating equation $v_{\pmb{\gamma}}\left(h\left(x_{i};s_{i},\pmb{\gamma}\right)\right)$ can be obtained using expression (1), the number of sample moments will be $r = 5$, *i.e.* there will be 5 elements for the vector $\pmb{g}\left(\pmb{\gamma}\right)$.

The first two cumulants which give the mean and the variance of the distribution are $c_{1} = \theta$, $c_{2} = \rho\left(\sigma^{2}+\frac{2}{\alpha^{2}}\right)$ and the method of moment estimators (MOM) estimators can be obtained in closed form. Let k_{i} be the i-th sample cumulant, the MOM estimators have been given by Reed [1] (p 477) with $\hat{\theta}_{\mathrm{MOM}} = k_{1}$,

$$\hat{\alpha}_{\mathrm{MOM}} = \left(20\frac{k_{4}}{k_{6}}\right)^{1/2}, \quad \hat{\rho}_{\mathrm{MOM}} = \frac{100}{3}\frac{k_{4}^{3}}{k_{6}^{2}} \quad \text{and} \quad \hat{\sigma}_{\mathrm{MOM}}^{2} = \frac{k_{2}}{\hat{\rho}_{\mathrm{MOM}}} - \frac{2}{\hat{\rho}_{\mathrm{MOM}}^{2}}.$$

Reed [1] (p 477) also noted that the MOM estimators often give a negative sign value for positive parameters like ρ or σ^{2}. It might be due to the lack of efficiency for MOM estimators for these parameters and therefore it is natural to consider alternatives like the GMM estimators. For financial applications with log returns data recorded as percentages being small in magnitudes, the parameters θ, ρ, σ^{2} tend to have small values in magnitude only α is around 18. GMM procedures with a number of sample moments larger than 20 might lead to the inverse of the matrix $\hat{\pmb{S}}$ difficult to obtain numerically.

By fixing $\alpha = 18, \rho = 0.1$ we consider the ranges $\theta = -0.04, -0.03, \cdots, 0, \cdots, 0.03, 0.04$ and $\sigma^{2} = 0.10, 0.11, \cdots, 0.15$. We simulate $M = 50$ samples of size $n = 100$ each and compute the MOM estimators and the GMM estimators. The GMM estimator for the location parameter θ is identical to the MOM estimator and only with this parameter that MOM estimator performs well. For other parameters, GMM estimators perform much better than MOM estimators.

Using simulated samples, we can estimate the individual mean square error for each parameter for each estimator. The vector of GMM estimators is denoted by $\hat{\pmb{\gamma}} = \left(\hat{\theta}, \hat{\alpha}, \hat{\rho}, \hat{\sigma}^{2}\right)'$. For GMM estimators, the following mean square errors are estimated using simulated samples, $\mathrm{MSE}\left(\hat{\theta}\right)$, $\mathrm{MSE}\left(\hat{\alpha}\right)$, $\mathrm{MSE}\left(\hat{\rho}\right)$, $\mathrm{MSE}\left(\hat{\sigma}^{2}\right)$. Similarly, we also estimate

$$\mathrm{MSE}\left(\hat{\theta}_{\mathrm{MOM}}\right), \ \mathrm{MSE}\left(\hat{\alpha}_{\mathrm{MOM}}\right), \ \mathrm{MSE}\left(\hat{\rho}_{\mathrm{MOM}}\right), \ \mathrm{MSE}\left(\hat{\theta}_{\mathrm{MOM}}\right), \ \mathrm{MSE}\left(\hat{\sigma}^2_{\mathrm{MOM}}\right).$$

with these quantities estimated, we can estimate the asymptotic overall relative of efficiency of these two methods as

$$\mathrm{ARE} = \frac{\mathrm{MSE}\left(\hat{\theta}\right) + \mathrm{MSE}\left(\hat{\alpha}\right) + \mathrm{MSE}\left(\hat{\rho}\right) + \mathrm{MSE}\left(\hat{\sigma}^2\right)}{\mathrm{MSE}\left(\hat{\theta}_{\mathrm{MOM}}\right) + \mathrm{MSE}\left(\hat{\alpha}_{\mathrm{MOM}}\right) + \mathrm{MSE}\left(\hat{\rho}_{\mathrm{MOM}}\right) + \mathrm{MSE}\left(\hat{\theta}_{\mathrm{MOM}}\right) + \mathrm{MSE}\left(\hat{\sigma}^2_{\mathrm{MOM}}\right)}.$$

Since GMM procedures perform better than MOM procedures for estimating α, σ^2, ρ and consequently overall GMM procedures perform much better based on the limited simulation study. More numerical studies should be conducted but it appears that the GMM procedures have the potentials to outperform MOM procedures for many models especially with models with more than 3 parameters as sample moments of order greater than 3 often are not stable in general. The results are displayed in table A in the **Appendix**.

In the second limited study, we consider the normal distribution which is also used for modelling log-returns data. The normal distribution is $N\left(\mu, \sigma^2\right)$, the ML estimators for μ and σ^2 are $\hat{\mu}_{\mathrm{ML}}$ and $\hat{\sigma}^2_{\mathrm{ML}}$. They are given respectively by the sample mean and the sample variance. The characteristic function for the normal distribution is

$$\phi_{\gamma}(s) = \mathrm{e}^{i\mu s - \frac{\sigma^2 s^2}{2}}, \ \ \gamma = \left(\mu, \sigma^2\right)',$$

$$Re\phi_{\gamma}(s) = \mathrm{e}^{-\frac{\sigma^2 s^2}{2}}\cos(\mu s), \ \ Im\phi_{\gamma}(s) = \mathrm{e}^{-\frac{\sigma^2 s^2}{2}}\sin(\mu s), \tag{11}$$

$$\left|\phi_{\gamma}(s)\right| = \mathrm{e}^{-\frac{\sigma^2 s^2}{2}} \ \ \text{and} \ \ \left|\phi_{\gamma}(s)\right|^2 = \mathrm{e}^{-\sigma^2 s^2} \ \ \text{and} \ \ -\frac{1}{2}\frac{\partial\left|\phi_{\gamma}(s)\right|^2}{\partial\sigma^2} = \frac{s^2}{2}\mathrm{e}^{-\sigma^2 s^2} \tag{12}$$

We only have 3 sample moments for GMM estimation for the normal model and we also use $M = 50$ simulated samples each with size $n = 1000$ to obtains GMM estimators and ML estimators for μ and σ^2 for the range of parameters with $\mu = -0.04, -0.03, \cdots, 0, 0.01, \cdots, 0.04$ and $\sigma^2 = 0.10, \cdots, 0.15$. The ranges of the parameters are often encountered for financial data and we estimate the individual mean square errors for the GMM estimators $\hat{\mu}, \hat{\sigma}^2$ and ML estimators $\hat{\mu}_{\mathrm{ML}}, \hat{\sigma}^2_{\mathrm{ML}}$. We found GMM estimator for μ is slightly less efficient than ML estimator but overall, GMM estimators are nearly as efficient as ML estimators based on the simulation results obtained and we estimate the overall relative efficiency

$$\mathrm{ARE} = \frac{\mathrm{MSE}\left(\hat{\mu}\right) + \mathrm{MSE}\left(\hat{\sigma}^2\right)}{\mathrm{MSE}\left(\hat{\mu}_{\mathrm{ML}}\right) + \mathrm{MSE}\left(\hat{\sigma}^2_{\mathrm{ML}}\right)},$$

the estimate ARE is close to 1 for the parameters being considered in Table A1 & Table A2.

4. Conclusion

Based on theoretical results and numerical results, it appears that:

1) The new GMM procedures are relatively simple to implement. The number of sample moments can be kept below the number ten yet the methods appear to have good efficiencies and offer good alternatives to MOM procedures which in general are not efficient for models with more than three parameters.

2) The proposed procedures are simpler to implement than GMM procedures based on a continuum of moment conditions and consequently might be of interests for practitioner who want to use these methods to analyze date where the model characteristic function is simple and have closed form but the density function is complicated, these situations often occur in practice.

3) The methods are less simulation oriented and consequently faster in computing time for implementations.

4) The estimators obtained have good efficiencies for some models being considered but more numerical and simulation works are needed to confirm the efficiencies using different parametric models and larger scale of simulations. In addition, further studies are needed for the topic on adding sample moments to make the chi-square goodness-of-fit test consistent without creating extensive numerical difficulties when it comes to obtaining the efficient matrix which is used for the quadratic form of the GMM objective function.

Acknowledgements

Conflicts of Interest

The author declares no conflict of interest regarding the publication of the paper.

References

[1] Reed, W.J. (2007) Brownian-Laplace Motion and Its Use in Financial Modelling. *Communications in Statistics, Theory and Methods*, **36**, 473-484. https://doi.org/10.1080/03610920601001766

[2] Kotz, S., Kozubowski, T.J. and Podgorski, K. (2001) The Laplace Distribution and Generalizations. Birkhauser, Boston. https://doi.org/10.1007/978-1-4612-0173-1

[3] Luong, A. and Bilodeau, C. (2017) Simulated Minimum Hellinger Distance Estimation for Some Financial and Actuarial Models. *Open Journal of Statistics*, **7**, 743-759. https://doi.org/10.4236/ojs.2017.74052

[4] Luong, A. (2017) Likelihood and Quadratic Distance Methods for the Generalized Asymmetric Laplace Distribution for Financial Data. *Open Journal of Statistics*, **7**, 347-368. https://doi.org/10.4236/ojs.2017.72025

[5] Luong, A. (2020) Generalized Method Moments and Generalized Estimating Functions Based on Generating Function for Count Models. *Open Journal of Statistics*, **10**, 516-539. https://doi.org/10.4236/ojs.2020.103031

[6] Feuerverger, A. and McDunnough, P. (1981) On the Efficiency of Empirical Cha-

racteristic Function. *Journal of the Royal Statistical Society, Series B*, **43**, 20-27.
https://doi.org/10.1111/j.2517-6161.1981.tb01143.x

[7] Groparu-Cojocaru, I. and Doray, L.G. (2013) Inference for the Generalized Normal Laplace Distribution. *Communications in Statistics—Simulation and Computation*, **42**, 1989-1997. https://doi.org/10.1080/03610918.2012.687064

[8] Koutrouvelis, I.A. (1980) Regression Type Estimation of the Parameters of Stable Laws. *Journal of the American Statistical Association*, **75**, 918-928.
https://doi.org/10.1080/01621459.1980.10477573

[9] Tran, K.C. (1998) Estimating Mixtures of Normal Distributions via Empirical Characteristic Function. *Econometric Reviews*, **17**, 167-183.
https://doi.org/10.1080/07474939808800410

[10] Carrasco, M. and Florens, J.P. (2000) Generalization of GMM to a Continuum of Moment Conditions. *Econometric Theory*, **16**, 797-834.
https://doi.org/10.1017/S0266466600166010

[11] Carrasco, M. and Kotchoni, R. (2017) Efficient Estimation Using the Characteristic Function. *Econometric Theory*, **33**, 479-526.
https://doi.org/10.1017/S0266466616000025

[12] Martin, V., Hurn, S. and Harris, D. (2013) Econometric Modelling with Time Series: Specification, Estimation and Testing. Cambridge University Press, Cambridge.

[13] Hayashi, F. (2000) Econometrics. Princeton University Press, Princeton.

[14] Hamilton, J.D. (1994) Time Series Analysis. Princeton University Press, Princeton.

[15] Godambe, V.P. and Thompson, M.E. (1989) An Extension of Quasi-Likelihood Estimation. *Journal of Statistical Planning and Inference*, **22**, 137-152.
https://doi.org/10.1016/0378-3758(89)90106-7

[16] Morton, R. (1981) Efficiency of Estimating Equations and the Use of Pivots. *Biometrika*, **68**, 227-233. https://doi.org/10.1093/biomet/68.1.227

[17] Yuan, K.-H. and Jennrich, R.I. (1998) Asymptotic of Estimating Equations under Natural Conditions. *Journal of Multivariate Analysis*, **65**, 245-260.
https://doi.org/10.1006/jmva.1997.1731

Appendix

Table A1. Overall estimate asymptotic relative efficiencies of the GMM estimators vs MOM estimators.

σ^2 \ θ	−0.04	−0.03	−0.02	−0.01	0.00	0.01	0.02	0.03	0.04
0.10	0.3709	0.0084	0.0104	0.2435	0.0419	0.5732	0.5725	0.0006	0.5008
0.11	0.0000	0.0834	0.0000	0.7649	0.0002	0.4282	0.0025	0.2395	0.0859
0.12	0.1362	0.2370	0.2177	0.0021	0.3731	0.5520	0.1105	0.0000	0.6634
0.13	0.0034	0.0880	0.0000	0.1197	0.0002	0.0670	0.2390	0.0188	0.1897
0.14	0.0530	0.0000	0.0004	0.4699	0.0702	0.0220	0.0001	0.0725	0.2501

Simulation studies for symmetric GNL distributions with $\alpha = 18, \rho = 0.1$.

Table A2. Overall estimate asymptotic relative efficiencies of the GMM estimators vs ML estimators.

σ^2 \ μ	−0.04	−0.03	−0.02	−0.01	0.00	0.01	0.02	0.03	0.04
0.10	1.0005	1.0009	1.0008	1.0005	1.0012	1.0014	1.0001	1.0011	1.0004
0.11	1.0005	1.0020	1.0036	1.0021	1.0026	1.0035	1.0020	1.0022	1.0038
0.12	1.0009	1.0020	1.0015	1.0018	1.0008	1.0032	1.0018	1.0013	1.0031
0.13	1.0022	1.0021	1.0040	1.0020	1.0040	1.0012	1.0023	1.0020	1.0012
0.14	1.0017	1.0022	1.0023	1.0025	1.0050	1.0157	1.0008	1.0019	1.0036

Simulation studies for normal distributions $N(\mu, \sigma^2)$.