Thai **J**ournal of **M**athematics : (2019) 185–200
Special Issue: Structural Change Modeling
and Optimization in Econometrics 2018

http://thaijmath.in.cmu.ac.th
ISSN 1686-0209

High-Order Generalized Maximum Entropy Estimator in Kink Regression Model

Payap Tarkhamtham and Woraphon Yamaka[1]

Centre of Excellence in Econometrics, Faculty of Economics
Chiang Mai University, Chiang Mai 50200, Thailand
e-mail : `woraphon.econ@gmail.com (W. Yamaka)`

Abstract : Investigation was made on the performance of the high-order Generalized Maximum Entropy (GME) estimators, namely Rényi and Tsallis GME, in the nonlinear kink regression context with an aim to replace the Shannon entropy measure. Used for performance comparison was the Monte Carlo Simulation to generate the sample size $n = 20$ and $n = 50$ with various error distributions. Then, the obtained model was applied to the real data. The results demonstrate that the high-order GME estimators are not much different from the Shannon GME estimator and are not completely superior to the Shannon GME in the simulation study. Nevertheless, according to the MAE criteria, Rényi and Tsallis GME perform better than the Shannon GME. Thus, it can be concluded that high-order GME estimator can be used as alternative tool in the nonlinear econometric framework.

Keywords : Shannon; Rényi; Tsallis; generalized maximum entropy; Kink regression.

1 Introduction

The concept of entropy was originally proposed by Shannon [1] under the information theory. Over the last several decades, entropy theory has played a major role in different areas where it has been successfully exploited and the related literature has grown dramatically. In this study, we focus on the econometric area,

[1]`Corresponding author.`

particularly the model estimation. After that, Janyes [2] developed the Shannon entropy measure and proposed a maximum entropy to uncover the unknown probability distribution of undetermined problems. Later, there have been several researchers who developed the entropy measure such as Rényi [3] measure which is the generalization of Shannon entropy and Tsallis [4] measure which is obtained through the joint generalization of the averaging procedures and the concept of information gain([5]).

Then, Golan et al. [6] generalized the maximum entropy of Jaynes [2] and introduced a generalized maximum entropy (GME) based on Shannons entropy-information measure to recover the unknown probability distribution of the regression model problem and for solving linear inverse problems. So, their suggested that Golan et al. [6] have considered applications of GME when the model is ill-conditioned and have linear inverse problems and suggested that GME is a feasible alternative method of estimation. The estimated parameters of the model are reparametrized as the expected values of discrete probability distribution defined on bounded supports. The computation of the GME is straightforward as the objective entropy measure is maximized subject to the constraints imposed by the model structure, data and other additive constraints. With its several advantages for the data limitation and unknown distribution of the error, GME has been widely used in various fields, political science, communications and information, engineering, physics, finance, and economics. (see [7] , [5], and [8])

Later, the study of Golan and Perloff [9] proposed higher-order entropy, namely Rnyi and Tsallis measures, as the new objective function of the GME estimator. They found that these two high-order entropy perform better than the Shannon's entropy, as these two estimations showed the lower mean squared error when compared to the Shannon's entropy for some values of order α , where $\alpha = \{1, 2, ...\}$.

In the previous literature, we found that GME are widely employed in the linear regression. Regression analysis can be considered one of the most widely used data analysis techniques in engineering, social sciences, biology, data mining, pattern recognition, etc. Generally speaking, this model is used to investigate the effect of the set of independent variables on the dependent variable. Nevertheless, the application of GME in non-linear model is limited. There are a few works such as Zheng and Gohin [10] which proposes a generalized maximum entropy (GME) approach to estimate nonlinear dynamic stochastic decision models. They found that the GME approach provides a similar accuracy level but much higher computational effciency for nonlinear models and shows favorable properties for small sample size data. Futhermore, the work of Sriboonchitta et al. [11] that applied the GME to estimate the unknown parameter in the nonlinear kink regression model [12]. We note that the regression function of this model is continuous, but the slope changes suddently at a kink point (threshold).

Inspired by the work of Golan and Perloff [9] that showed the high-order GME in linear model could outperform the conventional GME in terms of lower mean squared error, we in this paper apply the high order GME in the kink regression by replacing Shannon entropy measure with Rényi entropy and Tsallis entropy. Each of these entropy measures are indexed by a single parameter.

In this study, the kink regression model and GME are presented in Section 2. Section 3 provides an experiment study of GME kink regression. Section 4 presents an application study for our model by using the real data of the United States and Thailand. Finally, conclusion is provided in Section 5.

2 Methodology

2.1 Kink Regression Model

Our two regimes Kink regression model takes the form

$$\begin{aligned} Y_t = {} & \beta_1^-(x_{1,t} - \gamma_1) + \beta_1^+(x_{1,t} - \gamma_1) +, \ldots, \\ & + \beta_K^-(x_{K,t} - \gamma_K) + \beta_K^+(x_{K,t} - \gamma_K) + \beta_0 Z_t + \varepsilon_t \end{aligned} \tag{2.1}$$

where Y_t is $[T \times 1]$ continuous dependent variable at time t, $x_{k,t}$ is a matrix of $(T \times K)$ continuous independent variables at time t, and Z_t is the regime independent exogenous variable. The linear relationship appears among Y_t and Z_t, while the relationship between Y_t and $x_{k,t}$ is non-linear as their relationship changes at the unknown location called threshold or kink point γ_K. However, the kink regression function model is continuous in the variables $x_{k,t}$ and Z_t, but the slope with respect to $x_{k,t}$ is discontinuous at the threshold or kink point γ_K. Then, β is a matrix of $(T \times K \times 2)$ coefficients of unknown parameters and consist of $({\beta_1}^-, ..., {\beta_K}^-)$ and $({\beta_1}^+, ..., {\beta_K}^+)$ with respect to variable ${x'}_{k,t}$ for value of ${x'}_{k,t} \leq \gamma_k$, respectively and the coefficients with respect to variable ${x'}_{k,t}$ for value of ${x'}_{k,t} > \gamma_k$. According to Hansen [12], the regressor variables are subject to regime-change at kink point $(\gamma_1, ..., \gamma_K)$ thus these regressors can be separated into two regimes. Furthermore, the assumption of the error term ε_t is the distribution can be not normal $E(\varepsilon_t) = 0$.

2.2 Generalized Maximum Entropy Estimator

For a random vector x with K discrete values x_k, each with a probability $p_k = P(x_k)$ and $p = \{p_1, \ldots, p_k\}$ where p_k is a proper probability mass function, the Shannon measure is

$$H(x) = -\sum_k p_k \log p_k, \tag{2.2}$$

where $0 \log 0 = 0$ and $\sum\limits_k p_k = 1$. The two families of information measure are indexed by an order α. The Renyi measure is

$$H_\alpha^R(x) = \frac{1}{1-\alpha} \log \sum_k p_k^\alpha, \tag{2.3}$$

and the Tsallis measure is

$$H_\alpha^T(x) = c \frac{\sum_k p_k^\alpha - 1}{1-\alpha}, \tag{2.4}$$

where the value of c is a positive constant and depends on the particular units used. For simplicity, we set $c = 1$. (see [9]). Both Renyi and Tsallis measures become the Shannon measure as a special case when $\alpha \to 1$. Rather than searching for the point estimates $(\beta_1^-), \cdot, \beta_k^-$ and $(\beta_1^+), \cdot, \beta_k^+$,we can view these unknown parameters as expectations of random variables with M support value for (k), $Z = [z_1, \cdot, z_K]$ where $z_k = [\underline{z}_{k1}, \cdot, \overline{z}_{km}]$ for all $k = 1, \cdot, K$. Note that $\underline{z}_{k1}$ and $\overline{z}_{km}$ denote the lower and upper bound. In this study, we apply these three measures as the objective function in the GME estimator to estimate the parameters β_k^- and β_k^+ which can be reparametrized as

$$\begin{aligned} \beta_k^- &= \sum_m p_{km}^- z_{km}^+, \quad x_{k,t} \leq \gamma_k \\ \beta_k^+ &= \sum_m p_{km}^- z_{km}^+, \quad x_{k,t} > \gamma_k \end{aligned} \tag{2.5}$$

where p_{km}^- and p_{km}^+ are M dimensional estimated probability distributions defined on the set of support z_{km}^- and z_{km}^+. For the threshold or kink point, it can be computed by

$$\gamma_k = \sum_m h_{km} q_{km}, \tag{2.6}$$

where q_{km} is a vector of $q_k = [\underline{q}_{k1}, ..., \bar{q}_{km}]$, while $\underline{q}_{k1}$ and $\bar{q}_{km}$ are lower and upper bound of supports. Likewise, the error term ε_t is also computed by

$$\varepsilon_t = \sum_m w_{tm} v_{tm} \tag{2.7}$$

ε_t is also constructed as the mean value of random variable v where $v_t = [v_{t1}, ..., v_{tM}]$ is the support value and w_t is an M dimensional proper probability weight defined on the set v_t. From the reparametrized unknown variables β_k^- ,β_k^+, γ_k, and ε_t, we can rewrite Eq(2.1) as

$$\begin{aligned} Y_t = &\sum_m p_{1m}^- z_{1m}^- \left(x_{1,t} \leq \sum_m h_{1m} q_{1m} \right)_- \\ &+ \sum_m p_{1m}^+ q_{1m}^+ \left(x_{1,t} > \sum_m h_{1m} q_{1m} \right)_+ \\ &+, \ldots, + \sum_m p_{Km}^- z_{Km}^- \left(x_{1,t} \leq \sum_m h_{Km} q_{Km} \right)_- \\ &+ \sum_m p_{Km}^+ q_{Km}^+ \left(x_{1,t} > \sum_m h_{Km} q_{Km} \right)_+ + \sum_m w_{tm} v_{tm} \end{aligned} \tag{2.8}$$

Note that: For simplicity, we consider kink regression without Z_t variable. Then, we can construct our GME model as in the followings

$$\begin{aligned} H(p,h,w) &= \arg\max\{H(p)+H(h)+H(w)\} \equiv -\sum_k\sum_m p_{km}^- \log p_{km}^- \\ &- \sum_k\sum_m p_{km}^+ \log p_{km}^+ - \sum_k\sum_m h_{km}\log h_{km} - \sum_t\sum_m w_{tm}\log w_{tm} \end{aligned} \tag{2.9}$$

$$\begin{aligned} H(p,h,w) &= \arg\max\{H(p)+H(h)+H(w)\} \equiv \frac{1}{1-\alpha}\log\sum_k\sum_m p_{km}^{\alpha,-} \\ &+ \frac{1}{1-\alpha}\log\sum_k\sum_m p_{km}^{\alpha,+} + \frac{1}{1-\alpha}\log\sum_k\sum_m h_{km}^{\alpha} + \frac{1}{1-\alpha}\sum_k\sum_m w_{tm}^{\alpha} \end{aligned} \tag{2.10}$$

$$\begin{aligned} H(p,h,w) &= \arg\max\{H(p)+H(h)+H(w)\} \\ &\equiv \frac{1}{1-\alpha}\left(\sum_k\sum_m p_k^{\alpha,-} - 1\right) + \frac{1}{1-\alpha}\left(\sum_k\sum_m p_k^{\alpha,+} - 1\right) \\ &+ \frac{1}{1-\alpha}\left(\sum_k\sum_m h_{km}^{\alpha} - 1\right) + \frac{1}{1-\alpha}\left(\sum_t\sum_m w_{tm}^{\alpha} - 1\right). \end{aligned} \tag{2.11}$$

These objective functions are subject to the following constraints

$$\begin{aligned} Y_t &= \sum_m p_{1m}^- z_{1m}^- \left(x_{1,t} - \sum_m h_{1m}q_{1m}\right)_- + \sum_m p_{1m}^+ z_{1m}^+ \left(x_{1,t} - \sum_m h_{1m}q_{1m}\right)_+ \\ &+,\ldots,+ \sum_m p_{im}^- z_{im}^- \left(x_{i,t} - \sum_m h_{im}q_{im}\right)_- + \sum_m p_{im}^+ z_{im}^+ \left(x_{i,t} - \sum_m h_{im}q_{im}\right)_+ \\ &+ \sum_m w_{tm}v_{tm} \end{aligned} \tag{2.12}$$

$$\sum_m p_{km}^- = 1,\ \sum_m p_{km}^+ = 1,\ \sum_m h_{km} = 1,\ \sum_m w_{tm} = 1 \tag{2.13}$$

where p, h, and w are the probability on the interval [0,1]. If we consider one regressor ($k = 1$), this optimization problem can be solved by

$$\begin{aligned} L &= H(p,h,w) + \lambda_1(\theta) + \lambda_2\left(1 - \sum_m p_{km}^-\right) \\ &+ \lambda_3\left(1 - \sum_m p_{km}^+\right) + \lambda_4\left(1 - \sum_m h_{km}\right) \\ &+ \lambda_5\left(1 - \sum_m w_{tm}\right) \end{aligned} \tag{2.14}$$

where λ_i, $i = 1, ..., 5$ are the vectors of Lagrange multipliers for the data constraint as well as the additive constraint on the unknown probabilities and weights for the parameter and error, respectively. Summing up the objective entropy function subject to the kink regression in problem in Eq. 2.12 and additional restrictions in Eq. 2.13, we can solve by the Lagrangian method and then the first-order condition to get the optimal unique solution $\widehat{p}_{1m}^{-}$, $\widehat{p}_{1m}^{+}$, $\widehat{h}_{1m}$ and $\widehat{w}_{tm}$. Thus, the solution of the Shannon GME, Renyi GME, and Tsallis GME model are the following:

1. Shannon GME

$$\widehat{p}_{1m}^{-} = \frac{\exp(-z_{1m}^{-}\sum_t \widehat{\lambda}_{1t}(x'_{1,t} - \sum_m h_{1m}q_{1m})_{-})}{\sum_m \exp(-z_{1m}^{-}\sum_t \widehat{\lambda}_{1t}(x'_{1,t} - \sum_m h_{1m}q_{1m})_{-})}, \tag{2.15}$$

$$\widehat{p}_{1m}^{+} = \frac{\exp(-z_{1m}^{+}\sum_t \widehat{\lambda}_{1t}(x'_{1,t} - \sum_m h_{1m}q_{1m})_{+})}{\sum_m \exp(-z_{1m}^{+}\sum_t \widehat{\lambda}_{1t}(x'_{1,t} - \sum_m h_{1m}q_{1m})_{+})}, \tag{2.16}$$

$$\widehat{w}_{tm} = \frac{\exp(-\widehat{\lambda}_{1t}v_{1m})}{\sum_m \exp(-\widehat{\lambda}_{1t}v_{1m})}, \tag{2.17}$$

$$\widehat{h}_{1m} = \frac{\exp\left(-\left(\begin{array}{c}\sum_t \widehat{\lambda}_{1t}p_{1m}^{-}z_{1m}^{-}(x'_{1,t} - \sum_m q_{1m})_{-} \\ -\sum_t \widehat{\lambda}_{1t}p_{1m}^{+}z_{1m}^{+}(x'_{1,t} - \sum_m q_{1m})_{+}\end{array}\right)\right)}{\sum_m \exp\left(-\left(\begin{array}{c}\sum_t \widehat{\lambda}_{1t}p_{1m}^{-}z_{1m}^{-}(x'_{1,t} - \sum_m q_{1m})_{-} \\ -\sum_t \widehat{\lambda}_{1t}p_{1m}^{+}z_{1m}^{+}(x'_{1,t} - \sum_m q_{1m})_{+}\end{array}\right)\right)} \tag{2.18}$$

2. Renyi GME

$$\widehat{p}_{1m}^{-} = \left\{\left(\frac{1-\alpha}{\alpha}\right)\frac{\exp(-z_{1m}^{-}\sum_t \widehat{\lambda}_{1t}(x'_{1,t} - \sum_m h_{1m}q_{1m})_{-})}{\sum_m \exp(-z_{1m}^{-}\sum_t \widehat{\lambda}_{1t}(x'_{1,t} - \sum_m h_{1m}q_{1m})_{-})}\right\}^{1/(\alpha-1)}, \tag{2.19}$$

$$\widehat{p}_{1m}^{+} = \left\{\left(\frac{1-\alpha}{\alpha}\right)\frac{\exp(-z_{1m}^{+}\sum_t \widehat{\lambda}_{1t}(x'_{1,t} - \sum_m h_{1m}q_{1m})_{+})}{\sum_m \exp(-z_{1m}^{+}\sum_t \widehat{\lambda}_{1t}(x'_{1,t} - \sum_m h_{1m}q_{1m})_{+})}\right\}^{1/(\alpha-1)}, \tag{2.20}$$

$$\widehat{w}_{tm} = \left\{ \left(\frac{1-\alpha}{\alpha}\right) \frac{\exp(-\widehat{\lambda}_{1t} v_{1m})}{\sum\limits_m \exp(-\widehat{\lambda}_{1t} v_{1m})} \right\}^{1/(\alpha-1)}, \tag{2.21}$$

$$\widehat{h}_{1m} = \left\{ \left(\frac{1-\alpha}{\alpha}\right) \frac{\exp\left(-\left(\begin{array}{l} \sum\limits_t \widehat{\lambda}_{1t} p_{1m}^- z_{1m}^- (x'_{1,t} - \sum\limits_m q_{1m})_- \\ -\sum\limits_t \widehat{\lambda}_{1t} p_{1m}^+ z_{1m}^+ (x'_{1,t} - \sum\limits_m q_{1m})_+ \end{array}\right)\right)}{\sum\limits_m \exp\left(-\left(\begin{array}{l} \sum\limits_t \widehat{\lambda}_{1t} p_{1m}^- z_{1m}^- (x'_{1,t} - \sum\limits_m q_{1m})_- \\ -\sum\limits_t \widehat{\lambda}_{1t} p_{1m}^+ z_{1m}^+ (x'_{1,t} - \sum\limits_m q_{1m})_+ \end{array}\right)\right)} \right\}^{1/(1-\alpha)} \tag{2.22}$$

3. Tsallis GME

$$\widehat{p}_{1m}^- = \left\{ \left(\frac{1-\alpha}{\alpha}\right) \left[\sum_m \widehat{\lambda}_{1m} z_{1m}^- (x'_{1,t} - \sum_m h_{1m} q_{1m})_- + \lambda_{2k} \right] \right\}^{1/(\alpha-1)} \tag{2.23}$$

$$\widehat{p}_{1m}^+ = \left\{ \left(\frac{1-\alpha}{\alpha}\right) \left[\sum_m \widehat{\lambda}_{1m} z_{1m}^+ (x'_{1,t} - \sum_m h_{1m} q_{1m})_+ + \lambda_{3k} \right] \right\}^{1/(\alpha-1)} \tag{2.24}$$

$$\widehat{w}_{1m} = \left\{ \left(\frac{1-\alpha}{\alpha}\right) \left[\sum_m \widehat{\lambda}_{1m} v_{1m} + \lambda_{5t} \right] \right\}^{1/(\alpha-1)} \tag{2.25}$$

$$\widehat{h}_{1m} = \left\{ \left(\frac{1-\alpha}{\alpha}\right) \left[\begin{array}{l} \sum\limits_m \lambda_{1m} p_{1m}^- z_{1m}^- (x'_{1,t} - \sum\limits_m q_{1m})_- \\ -\sum\limits_m \lambda_{1m} p_{1m}^+ z_{1m}^+ (x'_{1,t} - \sum\limits_m q_{1m})_+ - \lambda_{4k} \end{array} \right] \right\}^{1/(\alpha-1)}. \tag{2.26}$$

In addition,by taking the second derivative of the Lagrangian with respect to$\widehat{p}_{1m}^-$, $\widehat{p}_{1m}^+$, $\widehat{h}_{1m}$ and $\widehat{w}_{tm}$.The hessian matrix for the GME linear regression problem must be negative definite, and this ensures that the entropy maximization problem can reach a unique global solution. Finally, these obtained probabilities are then used to compute the parameter estimates in the kink regression model, threshold parameter and error term as provided respectively in Eq.2.15 - Eq.2.26.

3 Experiment Study

In this section, we use a Monte Carlo simulation to examine the finite sample performance of alpha-order GME, as known as high-order GME. Then, the performance of two high-order GMEs are compared with Shannon GME in terms of the Bias and Mean Squared Error (MSE) of parameters. To this end, we consider the following model

$$Y_t = \beta_0 + \beta_1^-(x_{1,t} - \gamma_1)_- + \beta_1^+(x_{1,t} - \gamma_1)_+ + \varepsilon_t \tag{3.1}$$

In the data simulation, we firstly simulate a random sample (x) from normal distribution with mean equal to one and variance equal to one. The true values for coefficient parameters β_0, β^-,β^+, and γ are set to be 1, 2, -1, and 3 respectively. However, to make a fair comparison, we generated the error term from normal and non-normal distributions which consist of $N(0,1)$, $t(0,1,4)$, and $Unif(-2,2)$. Then, we generated a new sample during each Monte Carlo iteration by using the true values of coefficient parameters as specified above with the sample size 25 and 50 observations. Moreover, we determine the number of support points equal to 3 and we also specify the parameter support to be symmetric about the true parameters. The three support points consist of$z = [-5,0,5]$ as the support of intercept term, $v = [-3,0,3]$ as the support of error term, and $h = [-10,0,10]$ as the support of threshold. Finally, the performance of various model is evaluated in terms of Bias and MSE for each parameter which are given by

$$Bias = R^{-1}\sum_{r=1}^{R}\left(\tilde{\phi}_r - \phi_r\right),$$

and

$$MSE = R^{-1}\sum_{r=1}^{R}\left(\tilde{\phi}_r - \phi_r\right)^2.$$

where $\tilde{\phi}_r$ and ϕ_r are the estimated value and true value, respectively. R is the number of replications.

According to Tables 1-6, we found that Shannon GME seems to outperform the high α-order GME (both Renyi and Tsallis) in terms of lower MSE and Bias, especially, β^- and β^+ provide a strong accuracy when compared with Renyi and Tsallis GMEs. We can see that whether the sample size is $n = 20$ or $n = 40$, the bias and MSE are not much different, which suggests that the performance of GME seems to be affected little when a sample size is increased. The Bias and MSE are not much different with the normal and non-normal error distributions. In terms of threshold, Shannon GME estimator seems to be better in capturing threshold, except Tsallis GME estimator when the number of observations is 50 and error distribution is uniform. Then, when the high- αorder GME estimator are higher, the Bias and MSE seem to be higher in almost all cases.

Table 1: Experiment result of Kink regressions 25 observations with $N(0,1)$ errors

GME	Mean Square Error (MSE)				Bias			
	β_0	β^-	β^+	γ	β_0	β^-	β^+	γ
Shannon	0.8842	0.0152	0.0455	1.3169	0.6223	0.0772	0.1287	0.6480
Renyi								
$\alpha = 2$	1.0700	0.0457	0.0437	2.8457	0.8793	0.1447	0.1798	1.3984
$\alpha = 3$	0.9454	0.0616	0.0595	3.4481	0.8292	0.2348	0.2166	1.7204
$\alpha = 4$	0.8434	0.0818	0.0526	3.4346	0.7931	0.2388	0.1905	1.6082
$\alpha = 5$	0.9740	0.0758	0.0506	3.3555	0.8534	0.2013	0.2033	1.6293
Tsallis								
$\alpha = 2$	1.0055	0.0436	0.0514	2.3027	0.7579	0.1376	0.1770	1.1174
$\alpha = 3$	0.9680	0.0906	0.0573	3.3604	0.8900	0.2186	0.2237	1.6432
$\alpha = 4$	1.4248	0.0669	0.0626	3.7045	1.0148	0.2058	0.2339	1.8147
$\alpha = 5$	1.5283	0.1064	0.0600	3.2126	1.1166	0.2597	0.2121	1.5149

Table 2: Experiment result of Kink regressions 25 observations with $t(0,1,4)$ errors

GME	Mean Square Error (MSE)				Bias			
	β_0	β^-	β^+	γ	β_0	β^-	β^+	γ
Shannon	0.3254	0.0231	0.0266	1.5047	0.3311	0.0970	0.1051	0.7250
Renyi								
$\alpha = 2$	0.4080	0.0505	0.0272	2.2311	0.5018	0.1784	0.1293	1.1890
$\alpha = 3$	0.8248	0.0484	0.0442	2.9423	0.7463	0.1962	0.1916	1.6092
$\alpha = 4$	0.7648	0.0451	0.0383	2.4944	0.8052	0.1713	0.1773	1.3557
$\alpha = 5$	0.8030	0.0588	0.0517	2.9639	0.8138	0.1942	0.2121	1.5892
Tsallis								
$\alpha = 2$	0.5555	0.0428	0.0284	1.9820	0.4859	0.1402	0.1115	0.9116
$\alpha = 3$	0.8329	0.0545	0.0376	2.3977	0.7863	0.1661	0.1795	1.3107
$\alpha = 4$	0.9599	0.0325	0.0450	2.3312	0.8402	0.1262	0.1857	1.2163
$\alpha = 5$	0.8306	0.0415	0.0559	2.8382	0.7386	0.1767	0.2121	1.6004

Table 3: Experiment result of Kink regressions 25 observations with $Unif(-2,2)$ errors

GME	Mean Square Error (MSE)				Bias			
	β_0	β^-	β^+	γ	β_0	β^-	β^+	γ
Shannon	0.4234	0.0095	0.0092	0.3328	0.3985	0.069	0.0525	0.3865
Renyi								
$\alpha=2$	0.366	0.0369	0.0298	1.2198	0.457	0.1246	0.1312	0.9411
$\alpha=3$	0.7108	0.0558	0.0477	2.2841	0.7108	0.1676	0.1988	1.4458
$\alpha=4$	0.4894	0.0476	0.0274	1.477	0.5976	0.1548	0.1409	1.0884
$\alpha=5$	0.6679	0.0658	0.0253	1.4995	0.6779	0.2044	0.1305	1.0906
Tsallis								
$\alpha=2$	1.0299	0.0437	0.0255	1.0335	0.7811	0.1487	0.1212	0.7405
$\alpha=3$	0.6054	0.0485	0.0251	1.32	0.5837	0.1437	0.1169	0.9197
$\alpha=4$	1.1335	0.045	0.0401	1.2801	0.8421	0.1168	0.161	0.9347
$\alpha=5$	2.0212	0.0616	0.0499	1.6711	1.2508	0.173	0.1977	1.1395

Table 4: Experiment result of Kink regressions 50 observations with $N(0,1)$ errors

GME	Mean Square Error (MSE)				Bias			
	β_0	β^-	β^+	γ	β_0	β^-	β^+	γ
Shannon	0.0464	0.0136	0.0035	0.6246	0.1349	0.0504	0.0325	0.3857
Renyi								
$\alpha=2$	0.9444	0.0224	0.0425	2.2115	0.8332	0.1030	0.1760	1.2244
$\alpha=3$	0.6038	0.0407	0.0373	2.4301	0.7017	0.1710	0.1670	1.4405
$\alpha=4$	0.6640	0.0348	0.0351	2.1293	0.7189	0.1517	0.1693	1.3122
$\alpha=5$	0.8635	0.0388	0.0378	2.4502	0.7941	0.1604	0.1709	1.3425
Tsallis								
$\alpha=2$	0.4750	0.0216	0.0261	1.5307	0.4495	0.0917	0.1117	0.8353
$\alpha=3$	1.0947	0.0295	0.0369	2.1336	0.9115	0.1335	0.1803	1.3904
$\alpha=4$	1.1288	0.0186	0.0421	2.1534	0.9751	0.0976	0.1874	1.4133
$\alpha=5$	1.1736	0.0238	0.0434	2.5972	1.0141	0.1324	0.1910	1.4837

Table 5: Experiment result of Kink regressions 50 observations with $t(0, 1, 4)$ errors

GME	Mean Square Error (MSE)				Bias			
	β_0	β^-	β^+	γ	β_0	β^-	β^+	γ
Shannon	0.7198	0.0091	0.0332	1.3994	0.5435	0.0610	0.1051	0.6946
Renyi								
$\alpha = 2$	1.4313	0.0333	0.0543	2.3292	1.0227	0.1447	0.2040	1.2723
$\alpha = 3$	0.9469	0.0250	0.0574	2.4758	0.8308	0.1141	0.2100	1.3536
$\alpha = 4$	1.2527	0.0364	0.0622	3.0052	0.9834	0.1466	0.2074	1.5506
$\alpha = 5$	0.9933	0.0380	0.0504	2.9124	0.9191	0.1655	0.2082	1.6543
Tsallis								
$\alpha = 2$	0.6048	0.0317	0.0397	1.4375	0.5243	0.1354	0.1303	0.8350
$\alpha = 3$	1.1429	0.0327	0.0547	2.5068	0.9475	0.1493	0.1935	1.3588
$\alpha = 4$	1.1399	0.0447	0.0449	2.9902	0.9843	0.1800	0.1940	1.5935
$\alpha = 5$	1.4863	0.0445	0.0618	3.1913	1.1516	0.1665	0.2276	1.7271

Table 6: Experiment result of Kink regressions 50 observations with $Unif(-2, 2)$ errors

GME	Mean Square Error (MSE)				Bias			
	β_0	β^-	β^+	γ	β_0	β^-	β^+	γ
Shannon	0.5330	0.0391	0.0559	0.9139	0.5348	0.1036	0.1641	0.6575
Renyi								
$\alpha = 2$	0.3884	0.0039	0.0181	0.5000	0.5592	0.0507	0.1148	0.5692
$\alpha = 3$	0.4382	0.0140	0.0352	0.7110	0.4926	0.0846	0.1450	0.7710
$\alpha = 4$	0.4528	0.0034	0.0396	0.9204	0.6105	0.0495	0.1673	0.8722
$\alpha = 5$	0.5727	0.0136	0.0204	0.6902	0.6374	0.0873	0.1199	0.7625
Tsallis								
$\alpha = 2$	0.2694	0.0018	0.0063	0.1351	0.3916	0.0312	0.0632	0.2874
$\alpha = 3$	0.7942	0.0136	0.0330	0.3749	0.7998	0.0811	0.1530	0.5268
$\alpha = 4$	0.8093	0.0351	0.0430	0.5190	0.7731	0.1444	0.1854	0.6315
$\alpha = 5$	0.6848	0.0157	0.0253	0.2778	0.6868	0.0722	0.1257	0.4457

4 Case Study

In this section, the high α-order GMEs proposed in this paper are applied to the real data and compared to the Shannon GME estimator. We consider the growth/debt problem of Reinhart and Rogoff [13], and follow the scheme of Hansen [12] who suggested that the growth of economy tends to slow down when the government debt relative to GDP exceeds a threshold. Thus, we applied our model in US and Thai economy. Our data set are collected from Reinhart and Rogoff [13] consisting of yearly data of GDP growth and Debt/GDP ratio for US and Thailand. The data cover the period 1996 2010 with 15 observations for Thailand and the period 1961 2010 with 50 observations for US. In these two datasets, we choose the economic growth as the dependent variable, while Debt/GDP ratio is defined as independent variable. The GDP growth and Debt/GDP ratio for US and Thailand are illustrated in Figures 1 and 2, respectively.

Figure 1: GDP growth and Debt/GDP ratio of USA 1961 2010

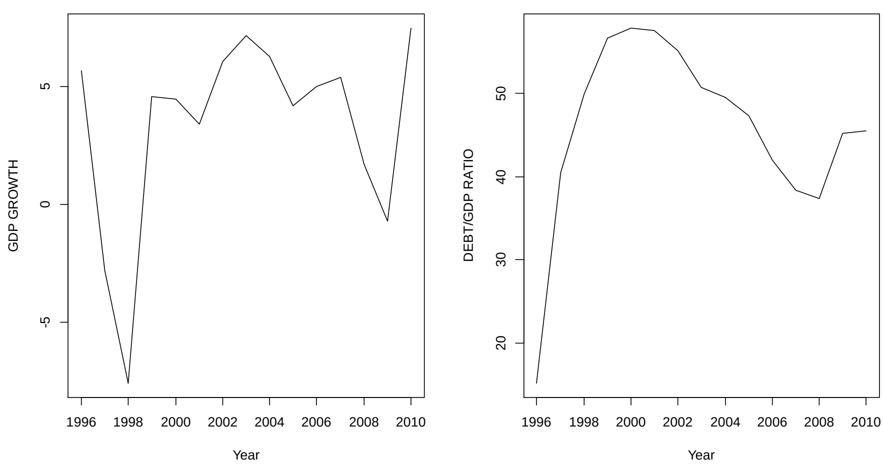

Figure 2: GDP growth and Debt/GDP ratio of Thailand 1996 2010

In these two cases, as we do not have any information about the possible range of the support vectors, we fix them to be as large as possible. In particular, all the support vectors for the regression parameters (β_0, β^-, andβ^+) are fixed to [30,0,30] whereas the support vectors for the residual terms are computed according to the three-sigma-rule $[3\sigma, 0, 3\sigma]$ where σ is the standard deviation of GDP growth. In the kink parameter γ, the information of this parameter is very important part for our model, thus, we fix the support vectors for this parameter around the mean of Debt/GDP ratio, say $\bar{x} - 0.2\bar{x}, \bar{x}, \bar{x} + 0.2\bar{x}$. That is, we set the lower and upper bounds to be 20% far from the mean. We note that the researcher needs to make sure that the real kink value is located within this range, otherwise we may obtain unreliable parameter estimate.

In the first US dataset, the results show that the coefficient (β^-, β^+) have a positive effect during economic recession (regime 1) and its had a negative effect during economic rising, it means that Debt/GDP ratio has a positive effect during economic recession and vice versa for economic boom which seems to correspond to the Keynesian economics. These results reveal that USA has a strong economic structure like other advanced countries.

Moreover, we can see that there exist two patterns of results. First, Shannon measure has a steep positive slope for low Debt/GDP ratio then switches to a negative slope at kink point around 40%. Meanwhile, Reni measure and Tsallis measures have a low positive slope for low Debt/GDP ratio and then switches to a negative slope at kink point around 40%.

Table 7: Coefficients (standard errors) from Kink regression application of US

GME	β_0	β^-	β^+	γ	MAE
Shannon	5.1158(0.5761)	0.3931(0.1113)	-0.1033(0.0280)	40.6128	1.5169
Renyi					
$\alpha = 2$	3.7740(0.6185)	0.1944(0.1196)	-0.0533(0.0301)	40.6077	1.6454
$\alpha = 3$	3.6903(0.6258)	0.1839(0.1210)	-0.0562(0.0304)	40.6074	1.6901
$\alpha = 4$	3.7110(0.6290)	0.1909(0.1216)	-0.0629(0.0306)	40.609	1.7162
$\alpha = 5$	3.6861(0.6354)	0.1963(0.1228)	-0.0666(0.0309)	40.6099	1.7475
Tsallis					
$\alpha = 2$	5.1204(0.5761)	0.3950(0.1113)	-0.1016(0.0280)	40.6128	1.5122
$\alpha = 3$	5.0807(0.5781)	0.3893(0.1108)	-0.1040(0.0282)	40.6713	1.526
$\alpha = 4$	4.9539(0.5825)	0.3775(0.1111)	-0.1040(0.0284)	40.7086	1.5481
$\alpha = 5$	4.8592(0.5918)	0.3624(0.1088)	-0.1042(0.0290)	40.9892	1.5756

Table 8: Coefficients (standard errors) from Kink regression application of Thailand

GME	$\beta_0 a$	β^-	β^+	γ	MAE
Shannon	-0.5150(1.9939)	-0.2205(0.1608)	1.0610(0.4486)	49.6376	3.8529
Renyi					
$\alpha = 2$	-0.5041(1.9750)	-0.2199(0.1585)	1.0610(0.4541)	49.8151	3.8246
$\alpha = 3$	-0.5082(1.9806)	-0.2199(0.1592)	1.0610(0.4524)	49.7627	3.8325
$\alpha = 4$	-0.5164(1.9974)	-0.2782(0.1606)	1.0610(0.4554)	49.747	3.9692
$\alpha = 5$	-0.5163(1.9969)	-0.2782(0.1605)	1.0610(0.4556)	49.752	3.9684
Tsallis					
$\alpha = 2$	-0.5190(2.0073)	-0.2782(0.1618)	1.0610(0.4522)	49.6481	3.9848
$\alpha = 3$	-0.5179(2.0033)	-0.2782(0.1613)	1.0610(0.4535)	49.6882	3.9785
$\alpha = 4$	-0.5190(2.0074)	-0.2782(0.1618)	1.0610(0.4522)	49.6475	3.9848
$\alpha = 5$	-0.5234(1.9997)	-0.2781(0.1627)	0.9575(0.4279)	49.2245	3.9084

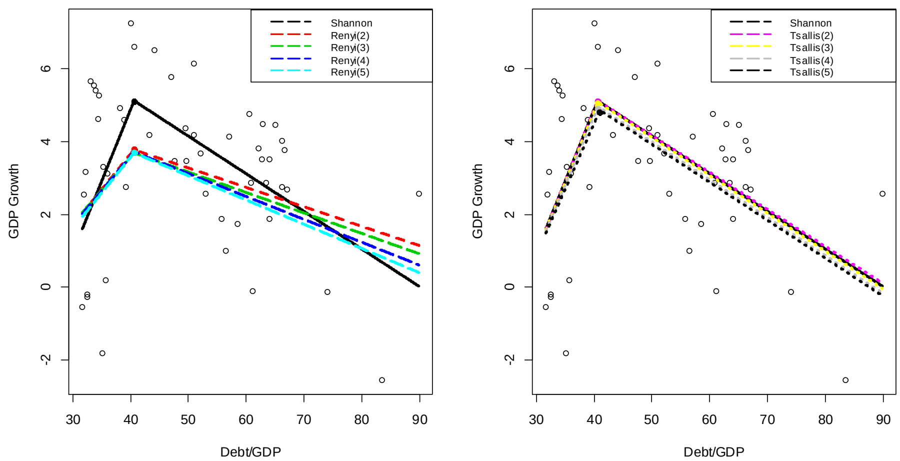

Figure 3: Kink plot of US data

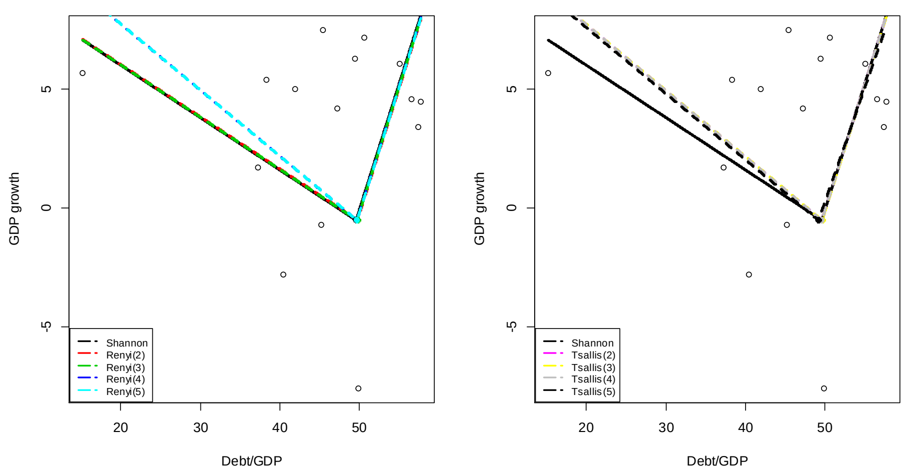

Figure 4: Kink plot of Thai data

The results from Table 8 seem contradictory to Keynesian economic approach which suggests that Debt/GDP ratio should have a positive effect on economic growth during economic recession (regime 1). In Thailands economy, the Debt/GDP ratio had a negative effect during economic recession. We suspect that the Thai economic structure may not change in our sample period as we have only 15 observations over the 15 yearss period and Thai economic structure is different from advanced countries. However, according to Figure 4, we may say that the higher Debt/GDP ratio may not contribute to a high growth for the Thai economy as the economic performance may not strong enough for returning the debt efficiently. But, when the Debt/GDP ratio is greater than kink point (around 49%), the effect of Debt/GDP ratio switches to a positive slope.

Finally, it is important to compare the estimators in this real application study by considering mean absolute error (MAE). MAE of an estimator measures the average of the difference between the real value and what is estimated. Small MAE value is needed in statistics because it is closer to actual data and leads to a better estimator. In this study, best estimator would be chosen based on the smallest value of MAE. Tables 7-8 show MAE in the last column for each estimator and it is found that the smallest MAE is obtained from the Tsallis GME with order 2 and Renyi GME with order 2, respectively, for US and Thai datasets.

5 Conclusion

In this paper, we apply the GME estimator which is able to solve the ill-posed problem or the limited data problems. We are following the work of Sriboonchitta et al. [11] and then we aim to replace Shannon GME estimator by high-order GME estimators consisting of Renyi and Tsallis. We then demonstrate the performance of high-order GME estimator to compare with that of Shannon GME estimator with various error distributions. Simulation results validate that the high-order GME estimators can provide accurate estimates for all unknown parameters. However, the Shannon GME estimator can capture the threshold or kink point better than high-order GME estimators, except the case of error distribution is uniform and the observation number is 50. Furthermore, we can see that the number of observations seems to be affected a little, the Bias and MSE are not much different. Likewise, the Bias and MSE values are not much different under the wide range of the error distributions. When the alpha-order gets higher, the Bias and MSE seem to be higher in almost all cases. The application results show that both Renyi and Tsallis GME estimators can provide accurate estimates for all unknown parameters of both countries and the results of both estimators are not much different from the Shannon GME. However, these two estimators seem to show a better performance than the conventional GME in terms of MAE. Thus, we can conclude that high-order GME estimator can be used as alternative tool.

Acknowledgements : We are grateful for financial support from Centre of Excellence in Econometrics, Faculty of Economics, Chiang Mai University. We thanks Nguyen, H. T. for valuable comment to improve this paper.

References

[1] Shannon, C. E. (1948). A mathematical theory of communication. Bell system technical journal, 27(3), 379-423.

[2] Jaynes, E. T. (1957). Information theory and statistical mechanics. Physical review, 106(4), 620.

[3] Rényi, A. (1961). On measures of entropy and information. HUNGARIAN ACADEMY OF SCIENCES Budapest Hungary.

[4] Tsallis, C. (1988). Possible generalization of Boltzmann-Gibbs statistics. Journal of statistical physics, 52(1-2), 479-487.

[5] Tabass, M. S., & Borzadaran, G. M. (2017). The generalized maximum Tsallis entropy estimators and applications to the Portland cement dataset. Communications in Statistics-Simulation and Computation, 46(4), 3284-3293.

[6] Golan, A., Judge, G. G., & Miller, D. (1996). Maximum entropy econometrics: Robust estimation with limited data (pp. 6-38). New York: Wiley.

[7] Pipitpojanakarn, V., Maneejuk, P., Yamaka, W., & Sriboonchitta, S. (2017). How does economic growth affect the well-being in Asia?.

[8] Tanak, A. K., Borzadaran, G. M., & Ahmadi, J. (2017). Maximum Tsallis entropy with generalized Gini and Gini mean difference indices constraints. Physica A: Statistical Mechanics and its Applications, 471, 554-560.

[9] Golan, A., & Perloff, J. M. (2002). Comparison of maximum entropy and higher-order entropy estimators. Journal of Econometrics, 107(1-2), 195-211.

[10] Zheng, Y., & Gohin, A. (2018). Estimating dynamic stochastic decision models: explore the generalized maximum entropy alternative (No. 2058-2018-5309).

[11] Sriboochitta, S., Yamaka, W., Maneejuk, P., & Pastpipatkul, P. (2017). A generalized information theoretical approach to non-linear time series model. In Robustness in Econometrics (pp. 333-348). Springer, Cham.

[12] Hansen, B. E. (2017). Regression kink with an unknown threshold. Journal of Business & Economic Statistics, 35(2), 228-240.

[13] Reinhart, C. M., & Rogoff, K. S. (2010). Growth in a time of debt. American Economic Review, 100(2), 573-78.

(Received 29 November 2018)
(Accepted 4 February 2019)