The generalize maximum Tsallis entropy estimator in kink regression model

Payap Tarkhamtham¹,², Woraphon Yamaka¹,³ and Songsak Sriboonchitta¹

¹Center of Excellence in Econometrics, Faculty of Economics, Chiang Mai University, Chiang Mai, Thailand
²International College of Digital Innovation, Chiang Mai University, Chiang Mai, Thailand
³Email: woraphon.econ@gmail.com

Abstract. Under the limited information situation, underdetermined or ill-posed problem in statistical inference is likely to arise. To solve these problems the generalized maximum entropy (GME) was proposed. In this study, we apply a generalized maximum Tsallis entropy (Tsallis GME) to estimate the kink regression using Monte Carlo Simulation and find that Tsallis GME performs better than the Least squares and Maximum likelihood estimators when the error is generated from unknown distribution. In addition, we can claim that the GME is a robust estimator and suggest that Tsallis GME can be used as an alternative estimator for kink regression model.

1. Introduction
In macroeconomic studies, the limitation of data is one of the most common issues we could find in the use of historical data and thereby estimating a realistic macroeconomic model is almost beyond possibility. Although the data collection system has been developed rapidly in recent years, many countries were found to lack data collection innovation, in particular in the less- and under-developed countries. [1] and [2] suggested in their works that the limited data will lead to underdetermined, or ill-posed problem. To solve these problems, the Entropy approach is conducted.

After [3] introduced the Shannon’s entropy measure [4], the concept was further developed in early 1990’s and was subsequently extended by [5] as the generalized maximum entropy (GME) estimator for estimating linear regression models. The estimated parameters of the model are reparameterized as the expected values of discrete probability distribution defined on bounded supports. The computation of the GME is straightforward, we can maximize Shannon’s entropy in the objective function subject to the constraints imposed by the model structure, data and other additive constraints. In the last decade, this estimation approach has been found in various applications, for example political science, communications and information, engineering, physics, biology, chemistry, finance, and economics. Although the GME estimation for a linear regression model has been applied successfully in many fields, there is no ability to explain the non-linear relationship between the dependent and independent variables. There are many previous evidences suggesting that the structure of the economic data might exhibit non-linear behaviour during the last few decades, especially after the economic globalization. Thus, the linear regression fails to capture this non-linear relationship. To deal with this problem, [6] proposed the application of GME estimator to non-linear kink regression model of [7]. They derived the GME problem where the Shannon’s entropy is set as the objective function and kink regression
structure is set as the constraint. Also, they have conducted a simulation study to compare the GME estimates with those from maximum likelihood (MLE) and least-square (LS) estimations. They found that GME estimator performs slightly better compared to those traditional estimators especially when the data is small and normal assumption does not hold.

In previous literature, [5] and [8] carried out several Monte Carlo experiments using artificial data to compare the performance of GME estimator to the ML and LS estimators. They found that GME is superior to the other estimators. In addition, [9] made an experiment study comparing a GME with ML estimator in the Tobit regression framework. They revealed that GME estimator was superior to ML estimator when observations are small sample and the non-normal distribution assumption is given. Moreover, [10] also made the similar experiment in the regression framework and the results showed the superior performance of GME under various scenarios. A review of GME refers to [11].

As stated earlier, GME estimator is introduced and used as an efficient method alternative to ML and LS estimators. However, in this study, a new estimator has been introduced based on the Tsallis entropy of [12]. This entropy measure is non-logarithmic and is obtained through the joint generalization of the averaging procedures and the concept of information gain [13]. Applications of Tsallis entropy estimation can be found in [14]. Their aim to compare the performance of GME and Ordinary least square (OLS) and found that GME estimator has better than OLS. In term of GME, Tsallis entropy estimates are not differing much from Shannon estimates. Therefore, in this study, we replace the Shannon entropy measure with Tsallis entropy. As far as we know, the estimation of Tsallis GME in kink regression has not been proposed in any literatures, and thus this study is the first with an aim to apply the Tsallis GME estimator in kink regression model. As a consequence, the main objective here is to improve the estimation method of kink regression when the error has non-normal distribution and the data is limited. In continuation, we also investigate this method by comparing with the Tsallis GME and Shannon GME.

The reminder of this study is organized as follows: Tsallis GME approach to the kink regression is presented in Section 2. In Section 3, the Monte Carlo experiments and results are introduced to evaluate and illustrate this method. Finally, Section 4 closes the study.

2. Methodology

2.1. Kink regression model

In this study two regime kink regression model is considered and it relies on

\[ Y_t = \beta_1' (x_{1,t} - \gamma_1) + \beta_2' (x_{1,t} - \gamma_1)_+ + \ldots + \beta_K' (x_{K,t} - \gamma_K) + \beta_K' (x_{K,t} - \gamma_K)_+ + \alpha Z_t + \varepsilon_t, \]

where \( Y_t \) is \([T \times 1]\) sequence of response variable at time \( t \), \( x_{i,t} \) is a matrix of \((T \times K)\) predictor variables at time \( t \). \( Z_t \) is the regime independent exogenous variable. The relationship between \( Y_t \) and \( x_{i,t} \) is non-linear while there is a linear relationship between \( Y_t \) and \( Z_t \). Therefore, the relationship of \( x_{i,t} \) with \( Y_t \) changes at the unknown location called threshold or kink point \( \gamma_k \), thus \( \beta \) is a matrix of \((T \times K \times 2)\) coefficients of unknown parameters. In other words, the kink regression function of this model is continuous in the variables \( x_{i,t} \) and \( Z_t \), but the slope with respect to \( x_{i,t} \) is discontinuous at the kink points. The parameter \((\beta_1^+, \ldots, \beta_K^+)\) and \((\beta_1^-, \ldots, \beta_K^-)\) are the coefficients with respect to variable \( x_{i,t} \) for value of \( x_{i,t} > \gamma_k \) and with respect to variable \( x_{i,t} \) for value of \( x_{i,t} \leq \gamma_k \), respectively. \( Z \) is a vector of covariates whose relationship with \( Y_t \) is linear. Following [7], the response variables are subject to regime-change at unknown threshold parameter or kink point \((\gamma_1, \ldots, \gamma_K)\) so the model can separate the data into two or more regimes. Note that the threshold or kink parameter \( \gamma_k \) is strictly
compact in the interior of the support of $x_{k,i}$. In this study, the assumption of the error term $\varepsilon_t$ is relaxed from the normal or any other distributions, and this study only assumes $E(\varepsilon_t) = 0$.

2.2. Tsallis entropy measure
In this study, we modify the GME estimator [5] by using Tsallis entropy measure to estimate an unknown parameter in kink regression model of [7]. Recall that [4] measure is

$$H(p) = -\sum_{k=1}^{K} p_k \log p_k,$$

where $p_k$ is the probability of the entropy function and $\sum_{k=1}^{K} p_k = 1$. As we mentioned earlier, this Shannon measure is replaced by Tsallis measure [15]

$$H^T(p) = \frac{1}{1-c} \left( \sum_{k=1}^{K} p^c_k - 1 \right),$$

where the value of $c$ is a positive constant and depends on the particular units used. These two entropy measures are indexed by a single parameter $c$, which we restrict to be strictly positive: $c > 0$. In particularly, $c = 2$ (see [14] and [15]). If $c = 1$, they become Shannon entropy. We apply the Tsallis GME with $c = 2$ to solve the inverse problem or ill-posed problem to the kink regression model. The parameters $(\beta^+_1, \ldots, \beta^+_K)$ and $(\beta^-_1, \ldots, \beta^-_K)$ can be computed by

$$\beta^+_k = \sum_m p^+_{km} z_{km}^{-}, x_{k,i} \leq \gamma_k, \quad \beta^-_k = \sum_m p^-_{km} z_{km}^{+}, x_{k,i} > \gamma_k,$$ (4)

where $p^+_{km}$ and $p^-_{km}$ are $M$ dimensional estimated probability distribution defined on the vector of support $z_{km}^-$ and $z_{km}^+$. Note that we can use our prior to specify the number of support and range of support of each parameter estimate. For the threshold or kink point, it is also computed by

$$\gamma_k = \sum_m h_{km} q_{km},$$ (5)

where $q_k = [q_{k1}, \ldots, q_{km}]$ is the vector of $q_{k1}$ and $q_{km}$ which are the lower and upper support bounds, respectively, of $\gamma_k$. Likewise, the error term $\varepsilon_t$ is also constructed by

$$\varepsilon_t = \sum_m w_{mt} v_{mt},$$ (6)

where $v_t = [v_{t1}, \ldots, v_{tm}]$ is the support value and $w_{mt}$ is an $M$ dimensional proper probability weights defined on the set $v_t$. For simplicity, we consider kink regression without $Z_t$ variable. Therefore, the study can reparameterize all unknowns $\beta^+_k, \beta^-_k, \gamma_k$, and $\varepsilon_t$, thereby rewriting equation (1) as

$$Y_t = \sum_m p^-_{im} z_{im}^{+} (x_{k,i} - \sum_m h_{im} q_{im}), + \sum_m p^+_{im} z_{im}^{-} (x_{k,i} > \sum_m h_{im} q_{im}), + \sum_m p^+_{im} z_{im}^{+} (x_{k,i} > \sum_m h_{im} q_{im}), + \sum_m w_{mt} v_{mt}.$$ (7)

Then, the optimization problem of this model is

$$H(p, h, w) = \arg\max_{p, h, w} \left\{ H(p) + H(h) + H(w) \right\} = -\frac{1}{1-c} \left( \sum_m \sum_k (p_{km})' - 1 \right) - \frac{1}{1-c} \left( \sum_m \sum_k (h_{km})' - 1 \right) - \frac{1}{1-q} \left( \sum_m \sum_k (w_{mt})' - 1 \right).$$ (8)
subject to
\[ Y_t = \sum_{m} p_{lm}^{-} z_{lm} (x_{1,t} - \sum h_{lm} q_{lm}) + \sum_{m} p_{lm}^{+} z_{km} (x_{1,t} - \sum h_{lm} q_{lm}) + \sum_{m} w_{lm} v_{lm}, \]
\[ \sum_{m} p_{lm}^{+} = 1, \sum_{m} p_{lm}^{-} = 1, \sum_{m} h_{km} = 1, \sum_{m} w_{lm} = 1, \]  \( \lambda_2 \)

where \( \tilde{p}_{lm}^{-} , \tilde{p}_{lm}^{+} , \tilde{h}_{lm} \) and \( \tilde{w}_{lm} \) are the probabilities on the interval \([0, 1]\), which require that the probabilities sum to one. For example, let consider one regressor \((k = 1)\), the constrained optimization problem can be solved by
\[ L = H(p, h, w) + \lambda_1^i (Y_t - \sum_{m} p_{lm}^{-} z_{lm} (x_{1,t} - \sum h_{lm} q_{lm}) - \sum_{m} p_{lm}^{+} z_{km} (x_{1,t} - \sum h_{lm} q_{lm}) + \sum_{m} w_{lm} v_{lm} + \lambda_2 (1 - \sum_{m} p_{lm}^{-} ) + \lambda_3 (1 - \sum_{m} p_{lm}^{+} ) + \lambda_4 (1 - \sum_{m} h_{km} ) + \lambda_5 (1 - \sum_{m} w_{lm} ), \]

where \( \lambda_i^i, i = 1,...,5 \) 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. Thus, this optimization yields
\[ \tilde{p}_{lm}^{-} = \left( \frac{1-c}{c} \right) \left[ \sum_{m} \tilde{\lambda}_{lm} z_{lm} (x_{1,t} - \sum h_{lm} q_{lm} ) + \tilde{\lambda}_2 \right] \]
\[ \tilde{p}_{lm}^{+} = \left( \frac{1-c}{c} \right) \left[ \sum_{m} \tilde{\lambda}_{lm} z_{km} (x_{1,t} - \sum h_{lm} q_{lm} ) + \tilde{\lambda}_3 \right] \]
\[ \tilde{w}_{lm} = \left( \frac{1-c}{c} \right) \left[ \sum_{m} \tilde{\lambda}_{lm} v_{lm} + \tilde{\lambda}_5 \right] \]
\[ \tilde{h}_{lm} = \left( \frac{1-c}{c} \right) \left[ \sum_{m} \tilde{\lambda}_{lm} p_{lm}^{-} z_{lm} (x_{1,t} - \sum q_{lm} ) - \sum_{m} \tilde{\lambda}_{lm} p_{lm}^{+} z_{km} (x_{1,t} - \sum q_{km} ) - \tilde{\lambda}_4 \right]. \]

Summing up equation (8)-15, the objective entropy function in equation (8), subject to the kink regression in equation (9) and additive restrictions in equation (10), is solved by the Lagrangian method for the first-order conditions to get the optimal unique solutions \( \tilde{p}_{lm}^{-} , \tilde{p}_{lm}^{+} , \tilde{h}_{lm} , \tilde{w}_{lm} \) and \( \tilde{\lambda} \). In addition, by taking the second derivative of the Lagrangian with respect to \( \tilde{p}_{lm}^{-} , \tilde{p}_{lm}^{+} , \tilde{h}_{lm} \) and \( \tilde{w}_{lm} \). The hessian for the GME linear regression problem is negative definite, and this ensures that the entropy maximization problem can reach a unique global solution.

3. Experiment study
The simulation and experiment study are conducted for assessing the finite sample performance of the GME estimator. In this section, we followed the simulation scheme given in [6] by dividing this experiment into two parts. In the first experiment, we varied the GME parameter and error support matrices and examined the sensitivity of the GME estimates to the prior information imposed. We consider the following three different settings for the number of support \((M)\): \( M = 3, 5, 7 \) thus \( z_{lm} = [-z, 0, z], [-z, -z/2, 0, z/2], [-z, -z/2, -z/2, z/3, 0, z/3, z/2, z] \), respectively. Note that the data simulation is carried out in the same manner with the only change being the number and interval of support. The support space of the \( \beta_1 \), \( y_1 \), and \( \varepsilon \) is determined to be symmetric around zero. For the interval of support, several support bounds for all estimated parameters are also investigated. The support space of \( \beta_1 \) is given to be in the range \([-z, 0, z]\) when \( z = 3, 6, \) and \( 9 \) while the range
$[-v, 0, v]$, for $v = 1, 2,$ and $3$, are given for model error $\varepsilon_i$. For the kink parameter, the support space of $\gamma_1$ is given to be in the range $[-h, 0, h]$ when $h = 5, 10,$ and $15$.

In the second part, the performance of GME with the least ML and LS estimators are compared. To do this, the study evaluates these estimators in terms of the bias and Mean Squared Error (MSE) of parameter estimates. To this end, the study considers a Kink regression model as follows:

$$Y_i = \alpha + \beta_i^- (x_{i1} - \gamma_i) + \beta_i^+ (x_{i1} - \gamma_i) + \varepsilon_i.$$  \hfill (16)

In this experiment, the study firstly draws a random sample from the equation (16) and uses the data to create an estimation sample. To do this, the study simulates $x_{i1} \sim Unif[-2, 2]$ and kink or threshold parameter value is $\gamma_i = 3$. The true values for coefficient parameters are $\alpha = 1, \beta_1^- = 2,$ and $\beta_1^+ = -1$. To make a fair comparison, the error terms are generated from various distributions namely $\chi^2(2), N(0,1)$, and $Unif(-2, 2)$. Then, we generate new data during each Monte Carlo iteration using the true parameter values as specified above. Sample size $n = 20$ and $n = 40$ are used each iteration. Then, the performance of estimator is evaluated in terms of the bias of each parameter.

### 3.1. Experiment results

For the first study, table 1 summarizes the estimated parameters under various GME specifications. We propose various number and range of parameter support to investigate the effect of the choice of $M$ and value of support $Z, h$ and $v$ on the parameter estimates. According to the suggestion of [2], it is difficult to specify the range and number of support points since they are complicated and composite. Therefore, various choice of number and range of support is taken into account in this experiment. The study investigates the performance of the GME estimator by providing various specifications of support space and the number of support points. To make a fair comparison, the error is generated $N(0,1)$ for all cases. According to table 1, where we can see that with the higher number of support, the lower MSE and bias of the estimated parameters are obtained. Anyway, when the support space is widened, bias and MSE do not increase in some cases. Hence, these results bring this study to conclude GME estimation is sensitive to the specified number and range of support. It is found that the GME estimation does have a lower bias and MSE as we widen the range of support and increase the number of support. In other words, GME estimation will perform better by increasing the number of support points and widening the range of the support.

### 3.2. Estimation comparison

Then, the study turns the focus on the second experiment to compare the performance of GME estimator with the other estimators. Specifically, the performance of Tsallis GME estimator is compare with Shannon GME, ML and LS estimation. In the case of Shannon and Tsallis GME estimators, this study estimates the model using three different parameter support matrices: $z = [-6, -3, 1.5, 0, 1.5, 3.6]$ as the support for the intercept, and coefficient terms, $v = [-2, -1, -0.5, 0, 0.5, 1.2]$ as the support for the error term, $h = [-10, -5, -2.5, 0, 2.5, 5, 10]$ as the support for the kink or threshold parameter. Note that all parameters are shrunk toward zero under this specification. In this experiment, we also specify the parameter support to be symmetric about the “true” parameters. Thus, the estimates will be closest to the true parameters yet compatible with the observed data. Moreover, the error of each model is generated from $\chi^2(2)$, $N(0, 1)$ and $Unif(2, 2)$; and the simulation results are illustrated in tables 2, 3, and 4, respectively.

For table 2-4, we present a Monte Carlo results which examine the risk measure of LS, MLE, Shannon GME, and Tsallis GME estimators. Table 2 provides the results for experiment using normal distribution as the error assumption, table 3 provides the results for experiment using chi-square
distribution as the error assumption and table 4 provides the results for experiment using uniform distribution as the error assumption.

\textbf{Table 1.} Choice of number of supports and support space.

| Support space \((n=20)\) | Bias | MSE |
|--------------------------|------|-----|
| \(z\) | \(h\) | \(v\) | \(M\) | \(\alpha\) | \(\beta_{-}\) | \(\beta_{+}\) | \(\gamma_{1}\) | \(\alpha\) | \(\beta_{-}\) | \(\beta_{+}\) | \(\gamma_{1}\) |
| [-3.3] | [-5.5] | [-1.1] | 3 | 0.0606 | 0.0491 | 0.0277 | 0.0687 | 0.1100 | 0.0485 | 0.0184 | 0.0952 |
| [-3.3] | [-5.5] | [-1.1] | 5 | 0.0600 | 0.0501 | 0.0273 | 0.0731 | 0.1100 | 0.0487 | 0.0184 | 0.0967 |
| [-3.3] | [-5.5] | [-1.1] | 7 | 0.0600 | 0.0493 | 0.0276 | 0.0697 | 0.1100 | 0.0485 | 0.0184 | 0.0952 |
| [-6.6] | [-10.10] | [-2.2] | 3 | 0.0279 | 0.0052 | 0.0079 | 0.0290 | 0.0429 | 0.0019 | 0.0024 | 0.0684 |
| [-6.6] | [-10.10] | [-2.2] | 5 | 0.0200 | 0.0048 | 0.0062 | 0.0172 | 0.0356 | 0.0019 | 0.0020 | 0.0530 |
| [-6.6] | [-10.10] | [-2.2] | 7 | 0.0199 | 0.0048 | 0.0062 | 0.0172 | 0.0356 | 0.0019 | 0.0020 | 0.0530 |
| [-9.9] | [-15.15] | [-3.3] | 3 | 0.0585 | 0.0137 | 0.0137 | 0.1098 | 0.1156 | 0.0049 | 0.0040 | 0.2488 |
| [-9.9] | [-15.15] | [-3.3] | 5 | 0.0585 | 0.0137 | 0.0137 | 0.1098 | 0.1156 | 0.0049 | 0.0040 | 0.2488 |
| [-9.9] | [-15.15] | [-3.3] | 7 | 0.0601 | 0.0159 | 0.0146 | 0.1199 | 0.1161 | 0.0052 | 0.0042 | 0.2603 |

The similar patterns are obtained from different cases. It is observed that the coefficient \(\alpha\) estimates are close to the true values, since the bias and MSE of coefficient \(\alpha\) are the lowest when compared with the other two estimators. Furthermore, it can be seen that when the sample size are \(n = 20\) and \(n = 40\), the bias and MSE are not much different, which suggests that the performance of GME seems to be affected little when a sample increases. In contrast to GME, we can observe the MLE and LS provide a lower bias and MSE when using a larger sample size. This indicates that the larger sample size cannot reduce uncertainty and its performance seems to be a little affected. Next, we turn our attention to threshold parameter \(\gamma_{1}\), the bias and MSE for GME estimated parameters increase when the error is generated from chi-square.

Comparing the Shannon GME and Tsallis GME, we note that when the error is generated from normal distributions, the MSE and bias of the Shannon and Tsallis GMEs are not much different. When \(n = 20\), the bias and MSE of Tsallis GME are smaller than those of the Shannon GME, see \(\alpha\), \(\beta_{-}\), and \(\gamma_{1}\). However, in the case of \(\beta_{+}\), bias and MSE of Tsallis GME become larger. Considering \(n = 40\) the situation seems to be different as the parameter bias and MSE of Tsallis GME are mostly larger than those of the Shannon GME. Next, we investigate the performance of the proposed estimator when the errors are generated from some non-normal distributions. Under an errors \(\chi^{2}(2)\) error distribution, we observe that the bias and MSE of parameters estimated by Tsallis and Shannon GMEs are similar. The performance is quite similar between \(n = 20\) and \(n = 40\). Last but not least, when the error is assumed to be uniform distribution, we observe that the Tsallis GME completely...
outperform the Shannon GME. The value of bias and MSE of Tsallis GME are smaller than those of the Shannon GME.

Table 2. Kink regressions with $N(0,1)$ errors.

|       | Tsallis | Shannon | MLE    | LS     |
|-------|---------|---------|--------|--------|
| $n = 20$ | Bias | MSE | Bias | MSE | Bias | MSE | Bias | MSE |
| $\alpha$ | 0.0062 | 0.0067 | 0.0127 | 0.0094 | 0.0156 | 0.0533 | 0.0321 | 0.0613 |
| $\beta_1$ | 0.0014 | 0.0002 | 0.0024 | 0.0012 | 0.0020 | 0.0009 | 0.0033 | 0.0004 |
| $\beta_2$ | 0.0010 | 0.0001 | 0.0004 | 0.0002 | 0.0090 | 0.0012 | 0.0070 | 0.0011 |
| $\gamma_1$ | 0.0008 | 0.0027 | 0.0118 | 0.0463 | 0.0221 | 0.0233 | 0.0354 | 0.0199 |
| $n = 40$ | Bias | MSE | Bias | MSE | Bias | MSE | Bias | MSE |
| $\alpha$ | 0.0292 | 0.0713 | 0.0038 | 0.0049 | 0.0138 | 0.0177 | 0.0136 | 0.0137 |
| $\beta_1$ | 0.0140 | 0.0059 | 0.0132 | 0.0076 | 0.0015 | 0.0005 | 0.0005 | 0.0003 |
| $\beta_2$ | 0.0096 | 0.0037 | 0.0042 | 0.0014 | 0.0006 | 0.0004 | 0.0102 | 0.0004 |
| $\gamma_1$ | 0.0351 | 0.0240 | 0.0444 | 0.0799 | 0.0026 | 0.0082 | 0.0132 | 0.0077 |

Table 3. Kink regressions with $\chi^2(2)$ errors.

|       | Tsallis | Shannon | MLE    | LS     |
|-------|---------|---------|--------|--------|
| $n = 20$ | Bias | MSE | Bias | MSE | Bias | MSE | Bias | MSE |
| $\alpha$ | 0.1000 | 0.2760 | 0.0835 | 0.1868 | 1.9497 | 4.5538 | 1.6332 | 3.1732 |
| $\beta_1$ | 0.0412 | 0.0305 | 0.0445 | 0.0352 | 0.0021 | 0.0228 | 0.0119 | 0.0160 |
| $\beta_2$ | 0.0317 | 0.0147 | 0.0293 | 0.0158 | 0.0074 | 0.0135 | 0.0413 | 0.0124 |
| $\gamma_1$ | 0.1149 | 0.5476 | 0.1504 | 0.7247 | 0.0296 | 0.4588 | 0.0849 | 0.4639 |
| $n = 40$ | Bias | MSE | Bias | MSE | Bias | MSE | Bias | MSE |
| $\alpha$ | 0.0002 | 0.0190 | 0.0196 | 0.0415 | 1.9081 | 4.0513 | 1.9206 | 4.0011 |
| $\beta_1$ | 0.1051 | 0.1167 | 0.1038 | 0.1138 | 0.0211 | 0.0097 | 0.0152 | 0.0076 |
| $\beta_2$ | 0.0692 | 0.0496 | 0.0639 | 0.0425 | 0.0064 | 0.0069 | 0.0163 | 0.0063 |
| $\gamma_1$ | 0.4526 | 2.5881 | 0.4193 | 2.2657 | 0.0279 | 0.2443 | 0.1174 | 0.3710 |

Table 4. Kink regressions with $Unif(2,2)$ errors.

|       | Tsallis | Shannon | MLE    | LS     |
|-------|---------|---------|--------|--------|
| $n = 20$ | Bias | MSE | Bias | MSE | Bias | MSE | Bias | MSE |
| $\alpha$ | 0.0118 | 0.0237 | 0.0003 | 0.0396 | 0.0141 | 0.2775 | 0.1913 | 0.1731 |
| $\beta_1$ | 0.0026 | 0.0012 | 0.0061 | 0.0063 | 0.0124 | 0.0105 | 0.0116 | 0.0038 |
| $\beta_2$ | 0.0014 | 0.0008 | 0.0041 | 0.0016 | 0.0024 | 0.0054 | 0.0003 | 0.0019 |
| $\gamma_1$ | 0.0198 | 0.0481 | 0.0289 | 0.0908 | 0.0280 | 0.1926 | 0.0308 | 0.1048 |
| $n = 40$ | Bias | MSE | Bias | MSE | Bias | MSE | Bias | MSE |
| $\alpha$ | 0.0409 | 0.0603 | 0.0512 | 0.0906 | 0.0198 | 0.1083 | 0.1863 | 0.2054 |
| $\beta_1$ | 0.0365 | 0.0269 | 0.0394 | 0.0303 | 0.0030 | 0.0033 | 0.0143 | 0.0064 |
| $\beta_2$ | 0.0157 | 0.0068 | 0.0174 | 0.0076 | 0.0040 | 0.0016 | 0.0092 | 0.0005 |
| $\gamma_1$ | 0.1585 | 0.5602 | 0.1652 | 0.5205 | 0.0200 | 0.0459 | 0.0186 | 0.0539 |

In summary, we can conclude that GME estimator is not completely superior to ML and LS estimators. Since, the GME estimator is outperform when the errors are generated from unknown
distribution. However, in terms of $\chi^2(2)$ and $N(0,1)$ the bias and MSE of GME are not much different from those of MLE and LS. In addition, we also find that Tsallis GME is likely the perform well in various cases, especially when the error is generated form unknown distribution.

4. Conclusions
In this study, we modified the GME estimator for the estimation of kink regression model by replacing the Shannon entropy with Tsallis entropy. By using this estimator, we can solve the ill-posed problem, especially when the data is limited and the error has unknown distribution. Simulation results validate that our GME estimator can provide accurate estimates for all unknown parameters. Moreover, the study carries out several Monte Carlo experiments comparing the Tsallis GME estimator to the Shannon GME, ML and LS estimators. In case of $n = 20$, the result shows that the estimation bias and MSE are lower for GME than for MLE and LS when the error is generated from unknown distribution ($Unif(2,2)$). Furthermore, it is apparent that bias and MSE values are not quite stable under various support bounds and sample sizes. Therefore, this study can conclude that the Tsallis GME can be used as an alternative estimator for the kink regression model estimation. Investigation on the selection of the single parameter $c$, and automatic specification of this parameter to achieve adaptiveness and further improvement shall be of interest for future studies.

References
[1] Golan A and Dose V 2001 A generalized information theoretical approach to tomographic reconstruction Journal of Physics A: Mathematical and General 34(7) 1271
[2] Heckelei T, Mittelhammer R and Jansson T 2008 A Bayesian alternative to generalized cross entropy solutions for underdetermined econometric models Food and Resource Economics Discussion Paper 2
[3] Jaynes E T 1957 Information theory and statistical mechanics Phys. Rev. 106 620
[4] Shannon C A 1948 Mathematical theory of communication Bell System Tech. 27 379
[5] Golan A, Judge G and Miller D 1996 Maximum entropy econometrics: robust estimation with limited data (John Wiley & Sons: Chichester, U.K.)
[6] Sriboonchitta S, Yamaka W, Manejuk P and Pastpipatkul P 2017 A generalized information theoretical approach to non-linear time series model Robustness in Econometrics (Springer Cham) 333
[7] Hansen B E 2017 Regression kink with an unknown threshold Journal of Business & Economic Statistics 35(2) 228
[8] Golan A 1998 Maximum Entropy, Likelihood and Uncertainty: A Comparison, Maximum Entropy and Bayesian Methods Fundamental Theories of Physics 98 35
[9] Golan A, Judge G and Perloff J 1997 Recovering Information from Censored and Ordered Multinomial Response Data Journal of Econometrics 79 23
[10] Wu X 2009 A weighted generalized maximum entropy estimator with a data-driven weight Entropy 11(4) 917
[11] Golan A 2002 Information and entropy econometrics-editor’s view Journal of Econometrics 107(1-2) 1-15
[12] Tsallis C 1988 Possible generalizations of Boltzmann-Gibbs statistics Journal of statistical Physics 52 479
[13] Tabass M S, Borzadaran G M and Amini M 2016 Renyi entropy in continuous case is not limit of discrete case Mathematical Sciences and Applications e-Notes 4(1) 113
[14] Tabass M S and 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
[15] Tsallis C 2002 Entropic nonextensivity: a possible measure of complexity Chaos, Solitons and Fractals 13 371-391