Thai Journal of Mathematics : (2017) 133–145
Special Issue on Entropy in Econometrics

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

Generalized Information Theoretical Approach to Panel Regression Kink Model

Phachongchit Tibprasorn[†], **Paravee Maneejuk**[‡,§,1], **and Songsak Sriboochitta**[‡,§]

[†]Faculty of Business Administration and Liberal Arts
Rajamangala University of Technology Lanna, Lampang, Thailand
e-mail : hally_cute@hotmail.com (P. Tibprasorn)

[‡]Faculty of Economics, Chiang Mai University
Chiang Mai 50200, Thailand

[§]Center of Excellence in Econometrics, Chiang Mai University,
Chiang Mai 50200, Thailand
e-mail : mparavee@gmail.com (P. Maneejuk)
e-mail : songsakecon@gmail.com (S. Sriboochitta)

Abstract : This study aims at exploring the use of generalized maximum entropy for estimating the unknown parameters in macroeconomic panel data models with two special concerns. First, the macroeconomic empirical panel studies usually face to a problem of data limitations, where the exiting estimation methods are often hard to get a significant results. Hence, we consider the GME estimator as one of the effective solutions for this problem. Second, there is often a discontinuity found in a relationship between explanatory and response variables in the macroeconomic studies. Therefore, in this study, the panel regression kink design based on the GME estimator is proposed for examining a discontinuous slope of the relationship between variables especially when the data is limited. The performance of the proposed method is evaluated through simulation study and later in the empirical analysis of the foreign direct investment. Both experiments showed a satisfiable performance of the model as well as the considered estimation technique.

[1]Corresponding author.

Keywords : maximum entropy; panel regression discontinuity; FDI analysis.
2010 Mathematics Subject Classification : 62P20; 91B84.

1 Introduction

Panel data set for macroeconomic study possess several major advantages over the conventional units like cross-sectional data or time series. It derives theoretical ability to isolate the effects of specific actions or more general policies. But one of the problems that bothers researchers when conducting macroeconomic studies using state-level panel data is data limitations. This can be the consequence of many reasons, such as uneven development in different countries. This inequality can cause different levels of data quality, data collection as well as source, and make equal data-lengths across countries become difficult to access. The limited observations can bring about under-determined or ill-posed problems to the analysis, and hard to get meaningful results. For instance, the sample size generally corresponds to the shape of the sampling distribution in such a way that increase in the sample size will make the sampling distribution more bell-shaped (according to the central limit theorem). However, when the sample size is limited, the shape of the corresponding distributions will become more difficult to obtain, which in turn make the optimal solution cannot be obtained through traditional estimation techniques. [Song and Cheon [1], 2006; and Lee and Cheon [2], 2014]

Apart from that problem in the estimation, this study also realizes the existence of discontinuity in the relationship between explanatory and response variables, which often occurs in the macroeconomic studies. For instance, there is a discontinuous adjustment in the economic growth through the different stages of economic development or the adjustment of cost in the dynamic of investment. Consequently, this study is focusing on two important issues; first, the estimation technique to handle the limited information in panel data model; and second, the discontinuous relationship between macroeconomic variables. To overcome these concerns, we introduce a panel regression kink design -extended from the regression kink of Hansen [3] (2017)- and suggest to use the Generalized Maximum Entropy (GME) approach to estimate the unknown parameters. This method allows for the examination of discontinuous slope of the relationship between variables when there is also limited information.

For empirical examination, this study applies the panel regression kink model based on GME estimator to examine the structure of Foreign Direct Investment (FDI) flows into European region in which the impacts of some macroeconomic variables on inward FDI are considerably discontinuous, especially due to the UK withdrawal from EU, or the so-called Brexit. The effect of this change in the context of foreign investment will be explored in this study in Section 4 application. Prior to the application, the general idea of panel regression kink model will be described in Section 2. Then, Section 3 contains the explanation and some mathematical properties of generalized maximum entropy estimator. In Section 5, we conduct a simulation study to evaluate the performance of GME estimator for the proposed model under finite samples. And Section 6 contains conclusion.

2 A General Idea of Panel Regression Kink Model

Consider a form of the panel regression kink model shown in Eq. (2.1), there are two error components namely an unobserved individual fixed effect u_i and a zero-mean idiosyncratic random disturbance ε_{it}, where $i = 1, \cdots, N$ correspond to individuals and $t = 1, \cdots, T$ correspond to time.

$$\begin{aligned} y_{it} =& \beta_1^-(X'_{1,it} - \gamma_1)_- + \beta_1^+(X'_{1,it} - \gamma_1)_+ \\ &+ \cdots + \beta_K^-(X'_{K,it} - \gamma_K)_- + \beta_1^+(X'_{K,it} - \gamma_K)_+ \\ &+ \theta D_{it} + u_i + \varepsilon_{it} \end{aligned} \tag{2.1}$$

The dependent variable y_{it} is $NT \times 1$ scalar, the regressor X_{it} is $NT \times K$ matrix explanatory variables. γ is a kink or a threshold parameter. β^- and β^+ are slope associated with different regimes. In this model, we use $(A)_-$ and $\min[A, 0]$ to represent respectively negative $(A)_+$ and $\max[A, 0]$ positive parts of a real number in A in order to divide the regressor X_{it} into two regimes. θ denotes $G \times 1$ vector of other dependent variables D_{it}, which are linearly related to y_{it}. The assumption about this error term is that it usually has no correlation with X_{it} and the individual-specific effect error component u_i. We also assume that u_i varies across individuals but is constant over time.

3 Estimation: Generalized Maximum Entropy Estimation

In this study, we propose the use of maximum entropy estimator to estimate the unknown parameters in the panel data kink regression. The maximum entropy concept is proposed in Jaynes (1957) to estimate the unknown probabilities of discrete probability distribution. Under this maximum entropy principle, one chooses the distribution for which the information is just sufficient to determine the probability assignment. The term of entropy refers to the uncertainty, regarding predicting the future outcome, represented by discrete probability distribution. Shannon [4] (1948) defined the entropy of the distribution of probabilities $p = (p_1, \cdots, p_K)$ as the measure$H(p) = -\sum_{k=1}^{K} p_k \log p_k$, where $0 \log 0$ tend to zero. This entropy measure reaches a maximum when $p_1 = p_2 = \cdots = p_K = \dfrac{1}{K}$.

To apply the concept of GME to be an estimator for our model, we follow the work of Sriboochitta, Yamaka, Maneejuk, and Pastpipatkul [5] (2017) but extend to the inverse problem of the panel data framework. Here, the point estimates $(\beta_1^-, \cdots, \beta_K^-)$, $(\beta_1^+, \cdots, \beta_K^+)$ and θ_g can be viewed as expectations of random variables with M support value for each estimated parameter value (k), $Z = [z_1, \cdots, z_K]$ where $z_k = [z_{k1}, \cdots, z_{km}] \forall k = 1, \cdots, K$ and $2 \leq M < \infty$. Note that z_{k1} and z_{km} are the lowest and highest value, respectively, of β. Thus we can express parameters β_k^- and β_k^+ as

$$\begin{aligned} \beta_k^- &= \sum_m p_{km}^- z_{km}^-, \ X_{k,it} \leq \gamma_k \\ \beta_k^+ &= \sum_m p_{km}^+ z_{km}^+, \ X_{k,it} > \gamma_k \end{aligned} \tag{3.1}$$

$$\theta_g = \sum_m r_{gm} a_{gm} Z_{g,it}, \tag{3.2}$$

where p_{km}^- and p_{km}^+ are the M dimensional estimated probability distribution defined on the set z_{km}^- and z_{km}^+, respectively. Likewise, r_m is M dimensional estimated probability distribution and a_m is support value. For the threshold γ_k, we also view that (k) element of γ as a discrete random variable with M support value $q_k = [q_{k1}, \cdots, q_{km}]$, where q_{k1} and q_{km} are the lowest and highest value, respectively, of γ_k

$$\gamma_k = \sum_m h_{km} q_{km} \tag{3.3}$$

Then, as we do with the estimated parameters, ε_t is also assumed to be bounded a priori. Golan, Judge, and Miller [6] (1996) assume that the errors can be bounded in $v_t = [v_{t1}, \cdots, v_{tm}]$, where v_{t1} and v_{tm} are the lowest and highest value, respectively, of ε_t. Then for each random error, ε_t, there exist $w \in [0, 1]$ such that

$$\varepsilon_t = \sum_m w_{tm} v_{tm} \tag{3.4}$$

However, it is difficult to specify the appropriate bounds in the panel data kink regression model. To get around this difficulty, Pukelsheim [7] (1994), Golan, Judge, and Miller [6] (1996) suggested setting the error bounds using the 3σ rule. That is, $v_{t1} = -3\sigma$ and $v_{tm} = 3\sigma$, where σ is the standard deviation of ε_t. In practice, we can get σ from either Ordinary Least Squares estimator or standard deviation of dependent variable y_{it}. Finally, the individual-specific effect error component α_i can be expressed as

$$\alpha_i = \sum_m f_{im} g_{im}, \tag{3.5}$$

where f_{im} is the a vector of positive probability that sum to one and this error can be bounded in $g_t = [g_{t1}, \cdots, g_{tM}]$, where g_{t1} and g_{tM} are the lowest and highest value, respectively, of α_i. Therefore, by using the re-parametrized unknowns $\beta_k^-, \beta_k^+, \theta_g, \gamma_g, \alpha_i$ and ε_t, we can rewrite Eq.(2.1) as

$$y = \mathrm{p}^- \mathrm{Z}^- (\mathrm{X} - \mathrm{hQ})_- + \mathrm{p}^+ \mathrm{Z}^+ (\mathrm{Z} - \mathrm{hQ})_+ + \mathrm{rAD} + \mathrm{fG} + \mathrm{wV}, \tag{3.6}$$

where $\mathbf{Z}^{-}$is $K\times KM$ matrix, $\mathbf{p}^{-}$ is $KM\times 1$, $\mathbf{Z}^{+}$ is $K\times KM$ matrix, $\mathbf{p}^{+}$ is $KM\times 1$, $\mathbf{h}$ is $KM\times 1$, $\mathbf{r}$ is $KM\times 1$, $\mathbf{A}$ is $G\times GM$ matrix, $\mathbf{f}$ is $NM\times 1$ matrix, $\mathbf{G}$ is $N\times NR$, w is $NTM\times 1$ matrix, and $\mathbf{V}$ is $NT\times NTM$. We assume that the unknown probabilities on the parameter and the weight on the error are independent and estimate them jointly by solving the constrained optimization problem

$$\begin{aligned}\mathbf{H}(\mathbf{p}^{-},\mathbf{p}^{+},\mathbf{h},\mathbf{r},\mathbf{f},\mathbf{w}) = \arg\max_{p^{-},p^{+},h,r,f,w} & \{\mathbf{H}(\mathbf{p}^{-})+\mathbf{H}(\mathbf{p}^{+})+\mathbf{H}(\mathbf{h})\\ & +\mathbf{H}(\mathbf{r})+\mathbf{H}(\mathbf{f})+\mathbf{H}(\mathbf{w})\}\\ & = -\mathbf{p}^{-}\log\mathbf{p}^{-}-\mathbf{p}^{+}\log\mathbf{p}^{+}-\mathbf{h}'\log\mathbf{h}\\ & -\mathbf{r}'\log\mathbf{r}-\mathbf{f}'\log\mathbf{f}-\mathbf{w}'\log\mathbf{w}, \end{aligned}\tag{3.7}$$

subject to

$$\mathbf{y}= \mathbf{p}^{-}\mathbf{Z}^{-}(\mathbf{X}-\mathbf{hQ})_{-}+\mathbf{p}^{+}\mathbf{Z}^{+}(\mathbf{X}-\mathbf{hQ})_{+}+\mathbf{rAD}+\mathbf{fG}+\mathbf{wV},\tag{3.8}$$

$$i_K = (\mathbf{I}_K\otimes i'_M)\mathbf{p}^{-},\tag{3.9}$$
$$i_K = (\mathbf{I}_K\otimes i'_M)\mathbf{p}^{+},\tag{3.10}$$
$$i_K = (\mathbf{I}_K\otimes i'_M)\mathbf{h},\tag{3.11}$$
$$i_G = (\mathbf{I}_K\otimes i'_M)\mathbf{r},\tag{3.12}$$
$$i_N = (\mathbf{I}_N\otimes i'_M)\mathbf{f},\tag{3.13}$$
$$\mathbf{i}_{NT} = (\mathbf{I}_{NT}\otimes i'_M)\mathbf{w}.\tag{3.14}$$

Then, the Lagrangian function is

$$\begin{aligned}\mathbf{L} = & -\mathbf{p}^{-}\log\mathbf{p}^{-}-\mathbf{p}^{+}\log\mathbf{p}^{+}-\mathbf{h}'\log\mathbf{h}-\mathbf{r}'\log\mathbf{r}-\mathbf{f}'\log\mathbf{f}-\mathbf{w}'\log\mathbf{w}\\ & +\lambda'[\mathbf{y}-\mathbf{p}^{-}\mathbf{Z}^{-}(\mathbf{X}-\mathbf{hQ})_{-}+\mathbf{p}^{+}\mathbf{Z}^{+}(\mathbf{X}-\mathbf{hQ})_{+}-\mathbf{rAD}-\mathbf{fG}-\mathbf{wV}]\\ & +\theta[i_K-\mathbf{I}_K\otimes i'_M)\mathbf{p}^{-}]+\Phi[i_K-(\mathbf{I}_K\otimes i'_M)\mathbf{p}^{+}]\\ & +\phi[i_K-(\mathbf{I}_K\otimes i'_M)\mathbf{h}]+\vartheta[i_G-(\mathbf{I}_G\otimes i'_M)\mathbf{r}]\\ & +\varphi[i_N-(\mathbf{I}_N\otimes i'_M)\mathbf{f}]+\xi[i_{NT}-(\mathbf{I}_{NT}\otimes i'_M)\mathbf{w}]. \end{aligned}\tag{3.15}$$

By taking the gradient of $\mathbf{L}$ to derive the first-order conditions, we obtain

$$\frac{\partial\mathbf{L}}{\partial\mathbf{p}^{-}} = -log\mathbf{p}^{-}-i_{KM}-\mathbf{Z}'(\mathbf{X}'-\mathbf{hQ})_{-}\lambda-(\mathbf{I}_K\otimes i_M)\theta=0,\tag{3.16}$$
$$\frac{\partial\mathbf{L}}{\partial\mathbf{p}^{+}} = -log\mathbf{p}^{+}-i_{KM}-\mathbf{Z}'(\mathbf{X}'-\mathbf{hQ})_{-}\lambda-(\mathbf{I}_K\otimes i_M)\Phi=0,\tag{3.17}$$
$$\frac{\partial\mathbf{L}}{\partial\mathbf{r}} = -log\mathbf{r}-i_{GM}-\mathbf{A}'\mathbf{D}'\lambda-(\mathbf{I}_K\otimes i_M)\Phi=0,\tag{3.18}$$
$$\frac{\partial\mathbf{L}}{\partial\mathbf{h}} = -log\mathbf{h}-i_K-\mathbf{Z}'(\mathbf{Q})_{-}\lambda-\mathbf{Z}'(\mathbf{Q})_{+}\lambda-(\mathbf{I}_K\otimes i_M)\phi=0,\tag{3.19}$$
$$\frac{\partial\mathbf{L}}{\partial\mathbf{f}} = -log\mathbf{h}-i_{NM}-\mathbf{G}'\lambda-(\mathbf{I}_N\otimes i'_M)\varphi=0,\tag{3.20}$$

$$\frac{\partial \mathbf{L}}{\partial \mathbf{w}} = -log\mathbf{w} - i_{NTM} - \mathbf{V}'\lambda - (\mathbf{I}_{NT} \otimes i'_M)\xi = 0, \quad (3.21)$$

$$\frac{\partial \mathbf{L}}{\partial \lambda} = \mathbf{y} - \mathbf{p}^{-}\mathbf{Z}(\mathbf{X} - \mathbf{hQ})_{-} - \mathbf{p}^{+}\mathbf{Z}^{+}(\mathbf{X} - \mathbf{hQ})_{+} - \mathbf{rAD} - \mathbf{fG} - \mathbf{wV}, \quad (3.22)$$

$$\frac{\partial \mathbf{L}}{\partial \theta} = i_K - (\mathbf{I}_K \otimes i'_M)\mathbf{p}^{-}, \quad (3.23)$$

$$\frac{\partial \mathbf{L}}{\partial \mathbf{\Phi}} = i_K - (\mathbf{I}_K \otimes i'_M)\mathbf{p}^{+}, \quad (3.24)$$

$$\frac{\partial \mathbf{L}}{\partial \phi} = i_K - (\mathbf{I}_K \otimes i'_M)\mathbf{h}, \quad (3.25)$$

$$\frac{\partial \mathbf{L}}{\partial \vartheta} = i_G - (\mathbf{I}_G \otimes i'_M)\mathbf{r}, \quad (3.26)$$

$$\frac{\partial \mathbf{L}}{\partial \varphi} = i_N - (\mathbf{I}_N \otimes i'_M)\mathbf{f}, \quad (3.27)$$

$$\frac{\partial \mathbf{L}}{\partial \xi} = i_{NT} - (\mathbf{I}_{NT} \otimes i'_M)\mathbf{w}. \quad (3.28)$$

After some algebra, we obtain

$$\mathbf{p}^{-} = \exp\{-\mathbf{Z}'(\mathbf{X}' - \mathbf{hQ})_{-}\lambda\}\exp\{i_{KM} - (i_{KM} - (\mathbf{I}_K \otimes i_M)\theta)\}, \quad (3.29)$$

$$\mathbf{p}^{+} = \exp\{-\mathbf{Z}'(\mathbf{X}' - \mathbf{hQ})_{+}\lambda\}\exp\{-i_{KM} - (i_{KM} - (\mathbf{I}_K \otimes i_M)\Phi)\}, \quad (3.30)$$

$$\mathbf{r} = \exp\{-\mathbf{A}'\mathbf{D}'\lambda\}\exp\{-i_{GM} - (\mathbf{I}_G \otimes i_M)\vartheta\}, \quad (3.31)$$

$$\mathbf{h} = \exp\{-\mathbf{Z}'(\mathbf{Q})_{-}\lambda\} - \exp]\{-\mathbf{Z}'(\mathbf{Q})_{+}\lambda\}\exp\{-i_K - (\mathbf{I}_K \otimes i'_M)\phi\}, \quad (3.32)$$

$$\mathbf{f} = \exp\{-\mathbf{G}'\lambda\exp\{-i_{KM} - (\mathbf{I}_N \otimes i'_M)\varphi\} = 0, \quad (3.33)$$

$$-\mathbf{w} = \exp\{-\mathbf{V}'\lambda\exp\{-i_{NTM} - (\mathbf{I}_{NT} \otimes i'_M)\xi\} \quad (3.34)$$

where $\otimes$ denotes the Hadamard product. Since the additive constraints Eqs.(3.9)-(3.14) and $\exp\{i_{KM} - (i_{KM} - (\mathbf{I}_K \otimes i_M)\theta)\}$, $\exp\{-i_{KM} - (i_{KM} - (\mathbf{I}_K \otimes i_M)\Phi)\}$, $\exp\{-i_{GM} - (\mathbf{I}_G \otimes i_M)\vartheta\}$, $\exp\{-i_K - (\mathbf{I}_K \otimes i'_M)\phi\}$, $\exp\{-i_{KM} - (\mathbf{I}_N \otimes i'_M)\varphi\}$ and $\exp\{-i_{NTM} - (\mathbf{I}_{NT} \otimes i'_M)\xi\}$ are constant, this optimization yields

$$\mathbf{p}^{-} = \frac{\exp\{-\mathbf{Z}'(\mathbf{X}' - \mathbf{hQ})_{-}\lambda\}}{(\mathbf{I}_K \otimes i_M)\exp\{-\mathbf{Z}'(\mathbf{X}' - \mathbf{hQ})_{-}\lambda\}}, \quad (3.35)$$

$$\mathbf{p}^{+} = \frac{\exp\{-\mathbf{Z}'(\mathbf{X}' - \mathbf{hQ})_{+}\lambda\}}{(\mathbf{I}_K \otimes i_M)\exp\{-\mathbf{Z}'(\mathbf{X}' - \mathbf{hQ})_{+}\lambda\}}, \quad (3.36)$$

$$\mathbf{r} = \frac{\exp\{a\mathbf{A}'\mathbf{D}'\lambda\}}{(\mathbf{I}_G \otimes i_M)\exp\{a\mathbf{A}'\mathbf{D}'\lambda\}}, \quad (3.37)$$

$$\mathbf{h} = \frac{\exp\{-\mathbf{Z}'(\mathbf{Q})_{-}\lambda\} - \exp\{-\mathbf{Z}(\mathbf{Q})_{+}\lambda\}}{(\mathbf{I}_K \otimes i'_M)(\exp\{-\mathbf{Z}'(\mathbf{Q})_{-}\lambda\} - \exp\{-\mathbf{Z}(\mathbf{Q})_{+}\lambda\})}, \quad (3.38)$$

$$\mathbf{f} = \frac{\exp\{-\mathbf{G}'\lambda\}}{(\mathbf{I}_N \otimes i'_M)\exp\{-\mathbf{G}'\lambda\}}, \quad (3.39)$$

$$\mathbf{w} = \frac{\exp\{-\mathbf{V}'\lambda\}}{(\mathbf{I}_{NT} \otimes i'_M)\exp\{-\mathbf{V}'\lambda\}}, \quad (3.40)$$

4 Simulation Study

In this section, we conduct a simulation and experiment study to investigate the finite sample performance of the GME estimator for the proposed model, panel kink regression. We compare the performance of the estimation when the number of support varies in its value, for example, three-point support $z_{1m} = [-z, 0, z]$ and five-point support $z_{1m} = [-z, -z/2, 0, 0, z/2, z]$. The support space of the $\beta_1^-, \beta_1^+, \gamma_1, u_i$ and ε_t are chosen to be uniformly symmetric around zero. In this study, several choices of the support for each of GME unknown parameters and errors are used. The support space of parameter and two error components are provided in Table 1. Our model is based on

$$y_{it} = \beta_1^-(X'_{1,it} - \gamma_1)_- + \beta_1^+(X'_{1,it} - \gamma_i) + \varepsilon_{it}, \tag{4.1}$$

where $X'_{1,it}$ is generated by $X'_{1,it} = 1 + 1x_{1,it} + \omega_{it}$, where ω_{it} is generated from $N(0, 1)$ and $x_{1,it}$ is generated from $N(0, \gamma_1)$. In the error term, u_i and ε_{it} are generated independently from $N(0, \sigma_u^2$ and $N(0, \gamma_1)$, respectively, where $\sigma_u^2 = \rho\sigma^2$ and $\sigma_\varepsilon^2 = (1-\rho)\sigma^2$ with $\sigma^2 = \sigma_\varepsilon^2 + \sigma_u^2$. We set $\sigma^2 = 1$ and $\rho = 1$ in all simulations. The threshold value is $\gamma = 0.5$. The true value for parameter β_1^-, and β_1^+ are set to be 1 and 2 respectively. In this Monte Carlo simulation, we consider sample size $N = 5,\ T = 10$ and $N = 10,\ T = 5$.

In the simulation results shown in Table 1, the choice of number of support points (M) and value of support space z, h, g, and v are considered in this simulations. Heckelei, Mittelhammer, and Jansson [8] (2008) suggested that the prior information on the support points and the number of support points are complicated, composite, and difficult to be specified. Thus, we examine the performance of the GME estimator by varying its support space and the number of support points. The estimated results for this simulation are shown in Table 1. We can see that the estimated parameters are close to their true values, nonetheless, except when the number of support points is large (i.e. $M = 7$), the threshold parameter, γ_1 seems to be affected by the change in the number of support points since it deviates a bit from its true value. When we consider the range of the support, the performance is quite similar and stable in all cases. In summary, this simulation result shows that the GME can perform well for our model.

5 An Application

This section presents an application of the panel kink regression model to the analysis of Foreign Direct Investment (FDI) flows into European countries. The study aims to examine how the inward FDI in EU countries changes due to the UK withdrawal from EU, or the so-called Brexit. In this experiment, macroeconomic factors are considered in order to explain the changes in inward FDI. However, we pay attention only to 10 countries in EU, namely Belgium, Czech Republic, Germany, Ireland, Spain, France, Italy, the Netherlands, Portugal, and Finland.

Table 1: Choices of number of supports and values of support space

Value of support space					**TRUE**		
N=10/T=5					1	2	0.5
z	h	g	v	M	β_1^-	β_1^+	γ_1
[-5,5]	[-1,1]	[-5,5]	[-5,5]	3	1.0797	2.0515	0.5124
[-10,10]	[-1,1]	[-6,6]	[-6,6]	3	1.1536	1.9033	0.5131
[-15,15]	[-1,1]	[-7,7]	[-7,7]	3	0.9647	2.0085	0.5115
[-5,5]	[-1,1]	[-5,5]	[-5,5]	5	0.9368	1.9275	0.4857
[-10,10]	[-1,1]	[-6,6]	[-6,6]	5	1.0323	2.0187	0.4877
[-15,15]	[-1,1]	[-7,7]	[-7,7]	5	0.9215	1.8907	0.4875
[-5,5]	[-1,1]	[-5,5]	[-5,5]	7	0.9904	2.1117	0.3588
[-10,10]	[-1,1]	[-6,6]	[-6,6]	7	1.0136	1.9207	0.3544
[-15,15]	[-1,1]	[-7,7]	[-7,7]	7	0.9189	2.0562	0.3548
N=20/T=10							
z	h	g	v	M	β_1^-	β_1^+	γ_1
[-5,5]	[-1,1]	[-5,5]	[-5,5]	3	1.0572	2.0409	0.5021
[-10,10]	[-1,1]	[-6,6]	[-6,6]	3	1.0493	2.0139	0.5011
[-15,15]	[-1,1]	[-7,7]	[-7,7]	3	1.0074	2.0001	0.5017
[-5,5]	[-1,1]	[-5,5]	[-5,5]	5	1.0273	2.0559	0.4947
[-10,10]	[-1,1]	[-6,6]	[-6,6]	5	1.1078	1.9728	0.4994
[-15,15]	[-1,1]	[-7,7]	[-7,7]	5	1.0247	2.0034	0.4874
[-5,5]	[-1,1]	[-5,5]	[-5,5]	7	1.0481	2.0865	0.3544
[-10,10]	[-1,1]	[-6,6]	[-6,6]	7	1.0138	2.0654	0.3551
[-15,15]	[-1,1]	[-7,7]	[-7,7]	7	1.0146	2.0578	0.3554

Source: Calculation.

5.1 Data

The macroeconomic factors considered as independent variable consist of gross domestic product (GDP), the harmonized index of consumer prices (HICP) -as an indicator of prices stability or inflation-, gross fixed capital formation (GFCF), and unit labor cost (ULC). Again, Foreign Direct Investment (FDI) is considered a dependent variable. All the data is collected annually from 2007 to 2016 in which the descriptive statistics of the variables are presented in Table 2.

Figure 1 presents a scatterplot matrix of the considered variables. The variables are plotted against each other in a diagonal line from top left to bottom right. For example, the first square in the second row is an individual scatterplot of FDI and GDP, where FDI is as the X-axis and GDP is as the Y-axis. But the first square in the third row shows an individual scatterplot of FDI and GFCF. In essence, the boxes on the upper right hand side of the whole scatterplot are mirror images of the plots on the lower left hand. This scatterplot can help determine rough linear correlation between variables. However, from this scatterplot, less of correlation among these variables can be found.

Table 2: Data description

Variable	FDI	GDP	GFCF	ULC	HICP
Mean	54,201.2	927,764.7	193,724.9	100.6	1.48
Median	26,497.3	517,445.5	107,013.5	100.6	1.45
Standard deviation	93,577.3	887,540.8	181,936.6	6.1	1.39
Maximum	734,010.3	3,144,050.0	630,034.0	114.9	6.30
Minimum	−28,375.2	138,302.9	25,122.0	78.1	−1.70

Source: Calculation.

Note: All the data is collected from Eurostat, except for the data of FDI which is retrieved from World Bank. FDI is measured in million USD; GDP and GFCF are in million euros; HICP is measured in annual percentage change; and ULC is index (2010=100).

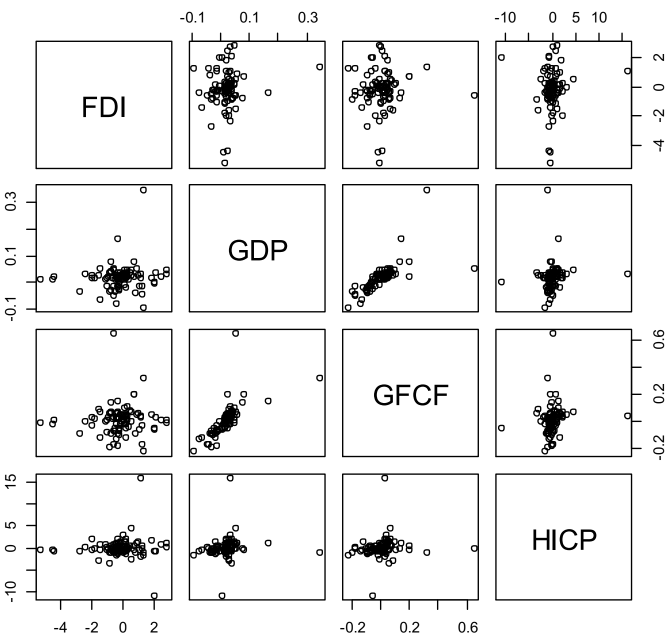

Figure 1: Scatterplot matrix

5.2 Testing Kink Effect

Since the nonlinear structure of the model is proposed in this study, we use the entropy ratio test, which is introduced by Golan and Dose [9] (2001), to check whether the kink regression model is significantly preferable to the linear regression model. It corresponds to the likelihood ratio test which measures the entropy discrepancy between the constrained and unconstrained models. Consider the constrained simple linear repression model with one regressor,

$$Y_t = \beta_1(x'_{1,t}) + \varepsilon_t. \tag{5.1}$$

The null hypothesis of entropy test is presented in Eq. (3.27) with the restriction of $\beta_1 = \beta_1^- = \beta_1^+$. Under this hypothesis, the kink variable γ does not exist while the alternative unconstrained hypothesis can be presented by

$$Y_t = \beta_1^-(x'_{1,t} \leq \gamma_1)_- + \beta_1^+(x'_{1,t} > \gamma_1)_+ + \varepsilon_t. \tag{5.2}$$

Hence, the entropy ratio statistic can be defined as

$$ER = 2\left|H_U(\beta_1^- \neq \beta_1^+) - H_R(\beta_1 = \beta_1^- = \beta_1^+)\right|, \tag{5.3}$$

where H_U is the unrestricted hypothesis (linear regression) and H_R is the restricted hypothesis (kink regression). Under certain regularity assumptions (see, Lee and Cheon [2], 2014; and Sriboonchitta et al. [5], 2017), ER converges in distribution to χ^2_F as $T \to \infty$ when the restriction is true, and F is degree of freedom. The approximate confidence level (α) for this test is computed by setting $ER(\cdot) < C_\alpha$, where C_α is the critical value of the χ^2_F at a significance level at α, and it is chosen; therefore $Pr(\chi^2_F < C_\alpha) = \alpha$.

Table 3: Entropy ratio test

	H_{linear}	H_{kink}	Entropy Ratio	Interpretation
FDI vs. GDP	−94.45	−100.22	−11.52	No kink effect
FDI vs. GFCF	−96.07	−86.03	20.07***	Kink effect exists.
FDI vs. ULC	−92.59	−91.39	2.38	No kink effect
FDI vs. HICP	−96.75	−94.25	5.001*	Kink effect exists.

Source: Calculation.

Note: "***" is significant at 1% level.

Table 3 shows results regarding the entropy ratio test. We can see that there is a significant kink effect occurred in a relationship between FDI and GFCF in addition to FDI and HICP.

5.3 Estimated Results

The estimated impacts on FDI flows into European countries are illustrated in Table 4. Even though the data set used in this experiment has a problem of data limitation, incomplete ranges, overall results can still reflect the satisfiable performance of the proposed method.

It shows that all explanatory variables create significant impacts on the inward FDI. Firstly, We find a positive relationship between GDP and FDI in which the coefficient is equal to 1.0726. Typically, GDP is considered as one of the main factors affecting inward FDI as an increase in production of goods and services leads to a rise in GDP. This, in turn, makes the economy more attractive to the foreign investors (Anaya and Indra, 2016). On the other hand, a negative effect of unit labor cost on inward FDI is found with a significant coefficient -0.6557, meaning that 1% increase of unit labor cost can lead to 0.66% decrease in the inward FDI. These results conform with the literature. (see, for example, Parcon [10], 2008; Bellak et al. [11], 2009; and Hatzius et al. [12], 1996)

Moreover, specific capability of our method allows for the nonlinear impacts of gross fixed capital formation (GFCF) as well as harmonized index of consumer prices (HICP, representing inflation) on inward FDI. The impacts are separated into two sections according to a kink point. As shown in Table 4, there is a large negative slope for any values of GFCF below the kink point around 0.1498 (0.15%) and the slope switches to a positive one for the values of GFCF greater than that kink value. It indicates that if the growth rate of GFCF is less than 0.15%, it will create a negative impact on inward FDI with coefficient -11.3034. On the other hand, the growth rate of GFCF greater than 0.15% can encourage inward FDI with coefficient 4.8466.

Table 4: Kink regression coefficients and standard errors (in parentheses)

FDI	**GDP**	**GFCF**	**ULC**	**HICP**
β	1.0726		-0.6557	
	(4.0549)		(5.2753)	
β^-		-11.3034		0.1104
(regime 1)		(0.0923)		(0.0923)
β^+		4.8466		0.2002
(regime 1)		(4.7981)		(0.0816)
γ		0.1498		0.3534
(Kink parameter)		(0.0110)		(0.0785)

Source: Calculation.

Note: All of the estimated parameters are significant at 1% level.

In the case of HICP, the result shows a small positive slope for any inflation rates under a significant kink value around 0.3534 (0.35%), which switches to a negative slope for the inflation rates beyond that kink value. This result indicates that changes in HICP lead to a change in inward FDI in the same direction, but the impact is discontinuous due to the kink effect. According to previous studies by Huybens and Smith [13] (1999), Boyd et al. [14] (2001) and Andinuur [15] (2013), low inflation rate is taken as a sign of internal economic stability in the host country and this would, in turn, increase the return on inward FDI. When inflation rate is low, nominal interest rate declines, therefore a low cost of capital. Moreover, the availability of capital at cheap lending rate would enable foreign investors not only to locate better partners in the host countries with sufficient domestic investment to supplement, but also to maximize the return on their investment.

6 Conclusion

So far, this study has dealt with two important issues: the data limitations, which can often be found in panel data models, and the discontinuous relationship between economic variables. To overcome these concerns when they happen simultaneously, the panel regression kink design based on the Generalized Maximum Entropy (GME) estimator is suggested in this study. The performance of this proposed method is evaluated through simulation study and later in the empirical analysis of foreign direct investment. Both experiments show a satisfiable performance of the model as well as the considered estimation technique, even though there is incomplete data for some countries.

In addition, the estimated results obtained from this proposed method just emphasize the distinct and discontinuous impacts of macroeconomic variables on the FDI. We conducted the experiment using the data of 10 European countries and found that all the considered variables, including gross domestic product (GDP), the harmonized index of consumer prices (HICP), gross fixed capital formation (GFCF), and unit labor cost (ULC), significantly affected the inward FDI. However, the effects of GFCF and HICP were discontinuous due to the kink effect. For the case of GFCF, the coefficients appeared to be negative in the first regime but positive in the second regime, whereas the effects of HICP were found to be positive in both regimes. Our finding suggests the appropriate level of inflation (in terms of HICP) can encourage inward FDI in EU countries.

References

[1] S. H. Song, S. Y. Cheon, A study of generalized maximum entropy estimator for the panel regression model, The Korean Journal of Applied Statistics 19 (2006) 521-534.

[2] J. J. Lee, S. Y. Cheon, Dual Generalized Maximum Entropy Estimation for Panel Data Regression Models, CSAM (Communications for Statistical Applications and Methods) 21 (5) (2014) 395-409.

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

[4] C. E. Shannon, A mathematical theory of communication, Bell System Technical Journal 27 (1948) 379-423.

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

[6] A. Golan, G. Judge, D. Miller, Maximum Entropy Econometrics: Robust Estimation with Limited Data, Chichester UK, John Wiley, 1996.

[7] F. Pukelsheim, The three-sigma rule, The American Statistician 48 (2) (1994) 88-91.

[8] T. Heckelei, R. Mittelhammer, T. Jansson, A Bayesian alternative to generalized cross entropy solutions for underdetermined econometric models, Food and Resource Economics Discussion Paper 2 (2008).

[9] A. Golan, V. Dose, A generalized information theoretical approach to tomographic reconstruction, Journal of Physics A: Mathematical and General 34 (7) (2001) 1271-1283.

[10] H. Parcon, Labor market flexibility as a determinant of FDI inflows, University of Hawaii at Manoa, Department of Economics, Working Papers, (2008).

[11] C. Bellak, M. Leibrecht, Do low corporate income tax rates attract FDI? Evidence from Central-and East European countries, Applied Economics 41 (21) (2009) 2691-2703.

[12] K. Hatzius, Y. Li, M. Werner, B. J. Jungnickel, Phase diagram and crystallization of polycarbonate/polycaprolactone blends, Macromolecular Materials and Engineering 243 (1) (1996) 177-187.

[13] E. Huybens, B. D. Smith, Inflation, financial markets and long-run real activity, Journal of Monetary Economics 43 (2) (1999) 283-315.

[14] J. H. Boyd, R. Levine, B. D. Smith, The impact of inflation on financial sector performance, Journal of monetary Economics 47 (2) (2001) 221-248.

[15] J. A. M. E. S. Andinuur, Inflation, Foreign Direct Investment and Economic Growth in Ghana, Doctoral dissertation, University of Ghana, 2013.

(Received 31 August 2017)
(Accepted 30 October 2017)