ICIC Express Letters
Volume **12**, Number **11**, November **2018**

ICIC International ©2018 ISSN 1881-803X
pp. 1163–1169

THE COMPARISON OF SEMI-PARAMETRIC TRENDING PANEL DATA MODELS AND NON-PARAMETRIC TIME-VARYING COEFFICIENT PANEL DATA MODELS

Liang Yin[1] and Weiguo Wang[2]

[1]School of Science
Dalian Nationalities University
No. 18, Liaohe West Road, Jinzhou New District, Dalian 116600, P. R. China
lyindufe@163.com

[2]School of Economics
Dongbei University of Finance and Economics
No. 217, Jianshan Street, Shahekou District, Dalian 116025, P. R. China
wwguodufe@163.com

Received April 2018; accepted July 2018

Abstract. *Semi-parametric trending panel data models can separate the nonlinear trend of variables to reflect the nonlinear characteristics of time variable and non-parametric time-varying coefficient panel data models can reflect the model's parameters variation with the time through the varying coefficient function. This paper analyzes the two models through the model specification, estimation methods and Monte Carlo simulation. The results show that it is necessary to choose suitable models to analyze the economic problems and summarize the economic regulation combining with the feature of data and economic theory in the complex economic phenomenon.*
Keywords: Semi-parametric trending panel data models, Time-varying coefficient models, Monte Carlo simulation

1. **Introduction.** Semi-parametric models are the most important statistical models in statistics which have the interpretability of the parameter part and the flexibility of the non-parametric part. Since they were introduced, they have been widely concerned in the field of statistics [1-4]. In many empirical problems, the semi-parametric regression model can be much closer to the real situation and can make full use of the information contained in the data. Chen et al. [5] first propose semi-parametric trending panel data models with cross-sectional dependence which can reflect the variation of variables by separating the nonlinear time trend of the model. There is another model proposed by Li et al. [6] which can reflect the nonlinear characteristics of the model and the parameter of the model is a varying coefficient function. Varying coefficient model is very popular in statistics. Silvapulle et al. [7] use local linear method to estimate time-varying trend or time-varying coefficient function. Zhou et al. [8] use the vector autoregressive model and the least square support vector machine to estimate the parameters of the linear time-varying structure system.

There is a rich literature on analyzing the realistic economic problems using the two models [9,10]. Two models have some similarities: not only the models but also explanatory variables contain nonlinear time trend. They also have some differences: parameters of the semi-parametric trending panel model are fixed. However, in addition to the nonlinear time trend, coefficients of the parameter partial of the non-parametric time varying coefficient model are varying with time. Both models have their own characteristics and are widely used. If we use inappropriate models to analyze the data of the problems, we

DOI: 10.24507/icicel.12.11.1163

will get wrong conclusions in the complex economic problems. Therefore, it is necessary to compare and analyze the two models. This paper compares the similarities and differences between the two models by model specification, estimation methods and Monte Carlo simulation. By comparison, we can find some differences between the two models under the same data. We need to use appropriate models to analyze the different data generation processes. There are some similarities and differences between the two models. However, as far as we know very few research has been done in the comparative analysis. In the complexity of real data, the innovation of this paper is comparing the two models from three aspects and providing a theoretical foundation for the application of the models.

This paper is organized as follows. In Section 2 and Section 3 the model specification of semi-parametric trending panel data models and non-parametric time-varying panel data models are discussed. In Section 4 the simulation results of the different data generating progress are presented. Finally, Section 5 concludes the paper.

2. Semi-Parametric Trending Panel Data Models and Its Estimation Methods.

The model we consider in this paper which is proposed in [5] is semi-parametric trending panel data model of the form:

$$Y_{it} = X_{it}\beta + f_t + \alpha_i + e_{it} \tag{1}$$

$$X_{it} = g_t + x_i + v_{it} \tag{2}$$

where $X_{it} = (X_{1t}, X_{2t}, \ldots, X_{pt})$ are P-dimension vector of explanatory variables, and $\beta = (\beta_1, \beta_2, \ldots, \beta_p)^T$ are P-dimension vector of unknown parameters. $f_t = f(t/T)$, $g_t = g(t/T)$ are both time trend functions with $f(\cdot)$, $g(\cdot)$ being unknown, and both $\{e_{it}\}$, $\{v_{it}\}$ are independent and identically distributed (i.i.d) across time but correlated among individuals. $\{\alpha_i\}$ is allowed to be correlated with $\{X_{it}\}$ through some unknown structure, while $\{e_{it}\}$ is assumed to be independent of $\{v_{it}\}$. In the above models, we impose the following restrictions on the fixed effects $\{\alpha_i\}$ and the individual effects $\{x_i\}$.

$$\sum_{i=1}^{N} \alpha_i = 0 \tag{3}$$

$$\sum_{i=1}^{N} x_i = 0_p \tag{4}$$

where 0_p is the P-dimension null vector. In model (1), Y_{it} is non-stationary variable which contains the nonlinear time trend, but the coefficients of the model are constants. Models (1) and (2) cover and extend some existing models. When $\beta = 0$, model (1) reduces to the non-parametric model discussed in [11]. When $N = 1$, models (1) and (2) reduce to the models discussed in [12].

The estimation method of semi-parametric trending panel data models is introduced in [5]. This paper developed a pooled semi-parametric profile likelihood dummy variable approach based on the first-stage local linear fitting to estimate both the parameter vector and the non-parametric time trend function. The authors defined a loss function by using the kernel function as a weighted function to get the parameter estimator and time trend through minimizing the loss function.

3. Non-Parametric Trending Time-Varying Coefficient Panel Data Model and Its Estimation Methods.

The other model we consider in this paper which is proposed in [6] is non-parametric trending time-varying coefficient panel data model of the form:

$$Y_{it} = X_{it}\beta_t + f_t + \alpha_i + e_{it}, \quad i = 1, \ldots, N,\ t = 1, \ldots, T \tag{5}$$

$$X_{it} = g_t + x_i + v_{it} \tag{6}$$

where $X_{it} = (X_{it,1}, X_{it,2}, \ldots, X_{it,p})$, $\beta_t = (\beta_{t,1}, \beta_{t,2}, \ldots, \beta_{t,p})$, all β_t and f_t are unknown functions, α_i reflects unobserved individual effect and $\{e_{it}\}$ is stationary and weakly dependent for each i and independent of $\{X_{it}\}$ and α_i. T is the time series length and N is the cross section size. In model (5), Y_{it} is non-stationary variable which contains the nonlinear time trend and the coefficients of the model are varying with the time. This is the difference with the model (1). Model (5) is called a fixed effects model when α_i is allowed to be correlated with $\{X_{it}\}$ through some unknown structure. And it is called random-effects model when α_i is uncorrelated with $\{X_{it}\}$. For the purpose of identification, we assume that α_i satisfies $\sum_{i=1}^{N} \alpha_i = 0$.

The estimation method of non-parametric trending time-varying coefficient panel data model is introduced in [6]. This paper develops two methods to estimate the trend function and the coefficient function without taking the first difference to eliminate the fixed effects. The first one eliminates the fixed effects by taking cross-sectional averages, and then uses a non-parametric local linear approach to estimate the trend function and the coefficient function. The second one is pooled local linear dummy variable approach. This is motivated by a least squares dummy variable method proposed in parametric panel data analysis. This method removes the fixed effects by deducting a smoothed version of cross-time average from each individual.

There are some similarities and differences of the estimations methods between the two models. Because the second model has two estimation methods, we will describe them respectively among the two models. For the similar part, firstly, all methods of the models use the kernel function as weighted function. Secondly, we get the parameter estimator and trend function directly through minimizing corresponding function for the method of the first model and the first method of the second model.

For the different part, the first method of the second model eliminates the fixed effects by taking cross-sectional averages and the second method of the second model removes the fixed effects by deducting a smoothed version of cross-time average from each individual. Moreover, the second method of the second model uses a non-parametric dummy variable technique for the model and then applies the two-step algorithm for the local linear dummy variable method.

4. **Monte Carlo Simulation.** In this part, we verify the different estimation results of the two models through the Monte Carlo simulation under the different data generation process. Firstly, we use the two models to analyze the data which is generated by semi-parametric trending panel data models. Secondly, we use the two models to study the data which is generated by non-parametric time-varying coefficient panel data models.

4.1. **Data generating process.** Consider a panel data model of the form:

$$Y_{it} = X_{it}\beta_t + f_t + \alpha_i + e_{it},\ 1 \le i \le N,\ 1 \le t \le T \tag{7}$$

Assume that, X_{it} is one-dimension vector, $f(u) = 2u^3 + u$, $\alpha_i = \frac{1}{T}\sum_{i=1}^{T} X_{it}$, $i = 1, \ldots, N-1$, $\alpha_N = -\sum_{i=1}^{N-1} \alpha_i$. The error terms e_{it} are generated as follows: for each $1 \le t \le T$, let $\tilde{e}_{.t} = (e_{1t}, e_{2t}, \ldots, e_{Nt})$ which is an N-dimensional vector and $e_{.t}$ $(1 \le t \le T)$ is generated as an N-dimensional vector of independent Gaussian variables with zero mean and covariance matrix $c_{ij} = 0.8^{|j-i|}$, $1 \le i, j \le N$. From the way e_{it} are generated. It is easy to see that

$$E(e_{it}, e_{js}) = 0, \quad 1 \le i, j \le N,\ t \ne s,$$

$$E(e_{it}, e_{jt}) = 0.8^{|j-i|}, \quad 1 \le i, j \le N,\ 1 \le t \le T$$

The above equations imply that $\{e_{it}\}$ is cross-sectional dependent and time independent. The explanatory variables X_{it} are generated by $X_{it} = g(t/T) + x_i + v_{it}$, $1 \le i \le N$, $1 \le t \le T$. $x_i : U(-0.2, 0.2)$, $x_N = -\sum_{i=1}^{N-1} x_i$, $\{v_{it}\}$ is independent of $\{e_{it}\}$ and is generated in the same way as $\{e_{it}\}$ but with a different covariance matrix $(d_{ij})_{N\times N}$, where $d_{ij} = 0.5^{|j-i|}$ for $1 \leq i, j \leq N$.

In order to compare the two models, we set β with two forms in our simulation. One is constant which is generated in the form of semi-parametric trending panel data models and the other varies with the time that is generated in the form of non-parametric time-varying panel data models. In different forms, we use two models mentioned above to study the differences of the estimators of the two models.

4.2. **Monte Carlo simulation results.** To evaluate the effectiveness of the two models, the simulation studies are carried out in the different data generation progress. Firstly, we analyze the data which is generated as the form: $Y_{it} = X_{it}\beta + f_t + \alpha_i + e_{it}$, where β is constant in the data generation process in this model. We next compare two estimation methods of the two models. The results are given in Table 1 and Table 2.

Secondly, we analyze the data which is generated as the form: $Y_{it} = X_{it}\beta_t + f_t + \alpha_i + e_{it}$, where $\beta_t = 0.5 * (t/T)$ is time function, $T = 5$, that is, $\beta = (0.1, 0.2, 0.3, 0.4, 0.5)$. We next compare two classes of estimation methods. The results are given in Table 3 and Table 4.

Table 1. Semi-parametric trending panel data models estimation (β is constant)

True value	$N\backslash T$	5	10	20	30
$\beta = 0.1$	10	0.1007 (0.3672)	0.0972 (0.2057)	0.0988 (0.1300)	0.1037 (0.0998)
	20	0.1011 (0.2047)	0.1020 (0.1344)	0.0998 (0.0855)	0.1003 (0.0619)
	30	0.0966 (0.2006)	0.1012 (0.1087)	0.1008 (0.0684)	0.1006 (0.0540)
	40	0.1027 (0.1421)	0.0996 (0.1089)	0.0999 (0.0570)	0.10053 (0.0505)
	50	0.0961 (0.1338)	0.1009 (0.0877)	0.1010 (0.0315)	0.0999 (0.0430)
$\beta = 1.0$	10	1.0070 (0.1869)	1.0018 (0.0922)	0.9924 (0.0645)	0.9991 (0.0514)
	20	1.0075 (0.1429)	1.0010 (0.0663)	0.9997 (0.0449)	1.0008 (0.0357)
	30	0.9962 (0.0960)	0.9980 (0.0551)	0.9994 (0.0366)	0.9995 (0.0289)
	40	0.9980 (0.0910)	0.9952 (0.0535)	0.9989 (0.0325)	0.9996 (0.0243)
	50	0.9981 (0.0816)	0.9969 (0.0425)	0.9959 (0.0298)	1.0018 (0.0207)
$\beta = 2.0$	10	2.0043 (0.3778)	1.9992 (0.2262)	1.9953 (0.1394)	2.0015 (0.1078)
	20	2.0104 (0.23697)	1.9957 (0.1472)	2.0010 (0.0936)	2.0005 (0.0726)
	30	2.0034 (0.1566)	2.0003 (0.1184)	1.9950 (0.0746)	1.9979 (0.0588)
	40	2.001 (0.2350)	2.0055 (0.1096)	1.9999 (0.0554)	2.0016 (0.0490)
	50	1.9963 (0.1771)	2.0009 (0.0732)	1.9986 (0.0542)	2.0008 (0.0460)

Note: Means and SDs of estimators

Table 2. Non-parametric time-varying coefficient model estimation ($T = 5$)

True value	$N\backslash\beta$	β_1	β_2	β_3	β_4	β_5
$\beta = 0.1$	10	0.0686 (3.4785)	0.0642 (1.9163)	0.0597 (0.8610)	0.0552 (0.3125)	0.0508 (0.2708)
	20	0.0424 (1.6033)	0.0648 (0.8607)	0.0785 (0.3723)	0.0965 (0.1379)	0.1146 (0.1576)
	30	0.0578 (1.4289)	0.0753 (0.7820)	0.0928 (0.3504)	0.1104 (0.1341)	0.1279 (0.1331)
	40	0.1240 (1.0098)	0.1246 (0.5625)	0.1252 (0.2593)	0.1258 (0.1002)	0.1264 (0.0851)
	50	0.0708 (0.8235)	0.0839 (0.4521)	0.0970 (0.2043)	0.1101 (0.0803)	0.1232 (0.0799)
$\beta = 1.0$	10	1.0753 (3.4789)	1.0438 (1.9909)	1.0123 (0.9657)	0.9809 (0.4032)	0.9494 (0.3036)
	20	0.9935 (2.1615)	0.9961 (1.1522)	0.9987 (0.4892)	1.0013 (0.1726)	1.0039 (0.2024)
	30	0.7625 (1.3208)	0.8072 (0.7261)	0.8518 (0.3331)	0.8965 (0.1418)	0.9411 (0.1523)
	40	0.8286 (0.8377)	0.8727 (0.4818)	0.9169 (0.2348)	0.9610 (0.0968)	1.0052 (0.0679)
	50	1.2425 (0.7728)	1.1957 (0.3861)	1.1489 (0.1444)	1.1021 (0.0477)	1.0554 (0.0961)
$\beta = 2.0$	10	2.0531 (3.5169)	2.0366 (1.9478)	2.0201 (0.8878)	2.0036 (0.3369)	1.9871 (0.2953)
	20	1.9661 (1.9549)	1.9704 (1.0796)	1.9746 (0.4911)	1.9789 (0.1893)	1.9831 (0.1744)
	30	1.9901 (1.2099)	1.9868 (0.6612)	1.9836 (0.2958)	1.9803 (0.1135)	1.9771 (0.1144)
	40	2.0328 (1.0569)	2.0261 (0.5749)	2.0193 (0.2551)	2.0126 (0.0974)	2.0059 (0.1020)
	50	2.0583 (0.8585)	2.0410 (0.4739)	2.0238 (0.2159)	2.0066 (0.0845)	1.9893 (0.0797)

Note: Means and SDs of estimators

Table 3. Semi-parametric trending panel data models estimation (β is time-varying)

$N\backslash T$	$T = 5$	$T = 10$	$T = 20$	$T = 30$	$T = 40$
$N = 10$	0.3376 (0.2328)	0.3004 (0.2013)	0.2718 (0.1058)	0.2580 (0.0960)	0.2595 (0.0891)
$N = 20$	0.3073 (0.1723)	0.2860 (0.1180)	0.2739 (0.0763)	0.2624 (0.0654)	0.2603 (0.0529)
$N = 30$	0.3321 (0.1454)	0.2830 (0.0980)	0.2696 (0.0455)	0.2620 (0.0390)	0.2600 (0.0408)
$N = 40$	0.3110 (0.1720)	0.2843 (0.0917)	0.2710 (0.0476)	0.2634 (0.0343)	0.2608 (0.0371)
$N = 50$	0.3228 (0.1331)	0.2909 (0.0680)	0.2700 (0.0491)	0.2602 (0.0355)	0.2589 (0.0330)

Note: Means and SDs of estimators

TABLE 4. Non-parametric time-varying coefficient model estimation ($T = 5$)

T	$N\backslash\beta$	$\beta_1 = 0.1$	$\beta_2 = 0.2$	$\beta_3 = 0.3$	$\beta_4 = 0.4$	$\beta_5 = 0.5$
$T = 5$	10	0.1167 (3.5208)	0.2229 (1.9168)	0.3290 (0.8481)	0.4351 (0.3145)	0.5412 (0.3162)
	20	0.1274 (1.7534)	0.2259 (0.9598)	0.3245 (0.4293)	0.4230 (0.1618)	0.5216 (0.1572)
	30	0.0985 (1.3339)	0.1947 (0.7216)	0.2909 (0.3163)	0.3872 (0.1178)	0.4835 (0.1263)
	40	0.1114 (1.0528)	0.2023 (0.5792)	0.2931 (0.2622)	0.3840 (0.1018)	0.4749 (0.0981)
	50	0.1087 (0.8599)	0.2078 (0.5050)	0.3069 (0.4729)	0.4059 (0.4729)	0.5050 (0.0782)

Note: Means and SDs of estimators

From the results of Tables 1-4, we can find that both models can reflect similar characteristics under different data generation process. Firstly, parameter estimator varies stably with the sample quantity varying. Secondly, for the fixed period, increasing the quantity of the individual can reduce the standard deviation of the parameter estimator. Thirdly, for the fixed individual, increasing the quantity of period can reduce the standard deviation of the parameter estimator.

However, both models have the different characteristics from each other. Firstly, semi-parametric trending panel data models can reflect the nonlinear characteristics by separating time trend to improve the effect of parameter estimation of the model and the parameters in the model are fixed in this progress. Secondly, non-parametric time-varying coefficient models reflect the nonlinear characteristics through variation of the parameters and the parameters of the model vary with the time. Thirdly, the estimators of the time-varying coefficient models are not good as semi-parametric trending panel data model under the data generated by semi-parametric trending panel data model. In a similar way, the estimators of the semi-parametric trending panel data model are not good as time-varying coefficient models under the data generated by non-parametric time-varying coefficient.

5. **Conclusions.** We have compared two models through model specification, estimation methods and simulation. From the simulation results we can find that two models have both similarities and differences. For the similar part, increasing the quantity of the time or the individual can reduce the standard deviation of the parameter estimator. For the different part, semi-parametric trending panel data models can reflect the nonlinear characteristics by separating time trend to improve the effect of parameter estimation of the model and the parameters of the model are fixed in this progress. In contrast, non-parametric time-varying coefficient models reflect the nonlinear characteristics through variation of the parameters and time trend, and the parameters of the model vary with the time. In the empirical analysis, we need to choose suitable models to analyze economic phenomena and summarize the economic law combining with the feature of data and economic theory.

Acknowledgement. This work is partially supported by the Natural Science Funds (71171305), (71201019) and (13YJC790023).

REFERENCES

[1] P. M. Robinson, Root-N-consistent semi-parametric regression, *Econometrica: Journal of the Econometric Society*, vol.56, no.4, pp.931-954, 1988.

[2] J. Q. Fan and T. Huang, Profile likelihood inferences on semi-parametric varying coefficient partially linear models, *Bernoulli*, vol.11, no.6, pp.1031-1057, 2005.
[3] J. Hahn and G. Ridder, Asymptotic variance of semi-parametric estimators with generated regressors, *Econometrica*, vol.81, no.1, pp.315-340, 2013.
[4] X. Hu, Semi-parametric inference for semi-varying coefficient panel data model with individual effects, *Journal of Multivariate Analysis*, vol.154, pp.262-281, 2017.
[5] J. Chen, J. Gao and D. Li, Semi-parametric trending panel data models with cross-sectional dependence, *Journal of Econometrics*, vol.171, no.1, pp.71-85, 2012.
[6] D. Li, J. Chen and J. Gao, Non-parametric time-varying coefficient panel data models with fixed effects, *Econometrics Journal*, vol.14, pp.387-408, 2011.
[7] P. Silvapulle, R. Smyth, X. Zhang, P. Fenech and S. Russell, Nonparametric panel data model for crude oil and stock market prices in net oil importing countries, *Energy Economics*, vol.67, pp.256-267, 2017.
[8] S. Zhou, Y. Ma, L. Liu, J. Kang, Z. Ma and L. Yu, Output-only modal parameter estimator of linear time-varying structural systems based on vector TAR model and least squares support vector machine, *Mechanical Systems and Signal Processing*, vol.98, pp.722-755, 2018.
[9] W. G. Wang and L. Yin, Parameters estimation of semi-parametric trending threshold panel data model, *Quantitative & Technical Economics*, vol.9, pp.124-137, 2014.
[10] A. Afonso, M. G. Arghyrou, D. M. Gadea and A. Kontonikas, "Whatever it takes" to resolve the European sovereign debt crisis? Bond pricing regime switches and monetary policy effect, *Working Paper, No.6691*, CESifo, 2017.
[11] P. M. Robinson, Non-parametric trending regression with cross-sectional dependence, *Journal of Econometrics*, vol.169, no.1, pp.4-14, 2012.
[12] J. T. Gao and K. Hawthorne, Semi-parametric estimation and testing of the trend of temperature series, *Econometrics Journal*, vol.9, pp.332-355, 2006.