Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 902602, 12 pages
http://dx.doi.org/10.1155/2015/902602

Research Article

Detection of Outliers in Panel Data of Intervention Effects Model Based on Variance of Remainder Disturbance

Yanfang Lyu

School of Mathematics and Quantitative Economics, Dongbei University of Finance and Economics, Dalian 116025, China

Correspondence should be addressed to Yanfang Lyu; yanfang902@dufe.edu.cn

Received 18 July 2014; Revised 19 September 2014; Accepted 24 September 2014

Academic Editor: Erol Egrioglu

The presence of outliers can result in seriously biased parameter estimates. In order to detect outliers in panel data models, this paper presents a modeling method to assess the intervention effects based on the variance of remainder disturbance using an arbitrary strictly positive twice continuously differentiable function. This paper also provides a Lagrange Multiplier (LM) approach to detect and identify a general type of outlier. Furthermore, fixed effects models and random effects models are discussed to identify outliers and the corresponding LM test statistics are given. The LM test statistics for an individual-based model to detect outliers are given as a particular case. Finally, this paper performs an application using panel data and explains the advantages of the proposed method.

1. Introduction

Outliers are observations in the dataset that appear to be unusual and discordant. If the sample contains outliers, the inappropriate estimation from contaminated observations may be strongly distorted and leads to unreliable results. Intervention effects may cause serious bias in estimating parameters as explained in the work of Fox [1], Martin and Yohai [2], Chang et al. [3], and Verardi and Croux [4]. Therefore, it is very important to identify these outliers in large datasets for both natural science and social science disciplines such as engineering, biology, education, medicine, economy, and sociology.

Detection of potential outliers plays a very important role in obtaining valuable and accurate information particularly in the field of engineering. An outlier of engineering observations may be due to an error in data transmission or human error in measurement. Some experts in the field of engineering noted the importance of identification of outliers. Febrero et al. [5] highlighted its importance: the analysis of outliers is an important aspect of any statistical analysis of data and especially is important to identify days or periods in which the NO_x levels are significantly large. Outlier identification is important to analyze the traffic volume data collected and used for a variety of purposes in intelligent transportation system [6]. Outlier detection is also important in the field of ecological engineering, as identification of atypical observations is an important concern in water quality monitoring [7].

Initial researches in outlier detection focused on time series-based outliers. Some influential studies have made contributions to the detection and identification of outliers [1, 3, 8–13]. Fox [1] proposed a fundamental approach for detecting and identifying outliers in a time series model. For linear regression models, it is a general approach to detect outliers to employ the mean-shift outlier model and some test statistics equivalent to the external studentized residuals [8, 9]. Chang et al. [3] provided an iterative procedure to detect and identify the outliers. Two distinct kinds of outliers are considered, namely, additive outlier (AO) and innovation outlier (IO), and the outliers in time series are regarded as being generated by dynamic intervention models at unknown time points. In recent years, researchers begin to be concerned about the outliers' detection of more complicated models. Shi and Chen [14] developed outlier detection for multilevel models and the proposed test can be used to detect outliers at any level in multilevel models and for any combinations of units. Chen [15] provided an approach to estimate the panel data model with a mixed fractional ARIMA remainder process, in which the data may contain different types of outliers. By the modified inverse Fourier transform, the outliers for the spectral Whittle

approach can be quickly detected and identified. Willems et al. [16] diagnosed multivariate outliers. Riani et al. [17] found an unknown number of multivariate outliers. Cerioli [18] developed multivariate outlier tests based on the high-breakdown Minimum Covariance Determinant estimator. Yan [19] proposed a novel method integrating self-organizing map (SOM) with adaptive nonlinear map (ANLM) for multivariate outlier detection. Yuen and Mu [20] proposed a novel probabilistic method for robust parametric identification and outlier detection in linear regression problems. The crux of this method is to calculate the probability of outlier, which quantifies how probable a data point is an outlier. Rapallo [21] made use of log-linear models and exact goodness-of-fit tests to specify the notions of outlier and pattern of outliers. Kuhnt et al. [22] introduced a new technique for the detection of outliers in contingency tables, where outliers are unusual cell counts with respect to classical log-linear Poisson models.

It is very difficult to identify outliers directly by eye, especially when faced by a panel data model with unfamiliar features and uncertain large datasets. As pointed out in Bramati and Croux [23], outliers are not always detectable by looking at residuals from a least squares fit, and diagnostic measures like the Cook distance suffer from the masking effect, as soon as multiple outliers are present. Not much effort has been given to the diagnostics and influence assessment of outliers in panel data models. Although a few researchers tend to be aware of this, there is little literature on the detection of outlier in a panel data model as an important issue.

This paper focuses on the detection and identification of outliers in panel data models with individual effects. Because of the presence of individual effects in panel data models, the traditional mean-shift model cannot differentiate individual effects from the mean disturbance. Thus, the mean-shift model cannot be applied in detecting outliers for panel data models with individual effects. A panel data model with outliers is likely to contaminate the residuals. This means that the variance of error term probably has a large deviation in the outlier model. Therefore, in this paper, a variance intervention effects model is proposed to study the detection of outlier. This paper is concerned with outlier detections through a method of variance intervention effects on the remainder disturbance using arbitrary strictly positive twice continuously differentiable function. Even if the error term u was observable in Baltagi [24], the equations of maximum likelihood (ML) estimator would still be highly nonlinear and difficult to solve explicitly. The test statistics based on Lagrange Multiplier (LM) approach are derived, since this LM test is based on the parameters' estimation under the null hypothesis and its computation is simple only requiring residuals. This paper focuses on a more general type of outlier that has specific impacts on subsequent observations, of which an individual outlier model is a particular case. The test statistics of a general type of outlier and an individual outlier are, respectively, calculated through Lagrange Multiplier (LM) approach. Furthermore, this paper would demonstrate outlier detection of fixed effects models and random effects models by the corresponding LM test statistics.

The rest of the paper is organized as follows. Section 2 briefly presents panel data models with individual effects. Section 3 proposes the variance intervention effects outlier model based on the remainder disturbance. Section 4 describes maximum likelihood estimator. Section 5 provides an LM testing approach for the detection and identification of a general type of outlier. Furthermore, fixed effects and random effects models are discussed with outliers and the corresponding LM test statistics are given. The LM test statistics of an individual outlier model as a particular case are given. Section 6 performs an application of the proposed method using a panel data and explains the advantages of the method. Finally, Section 7 provides the concluding remarks. Proofs of the main results are provided in the appendix.

2. Panel Data Models with Individual Effects

Firstly, the following panel data model with individual effects is considered:

$$y_{it} = \alpha_i + x_{it}'\beta + \varepsilon_{it}, \quad i = 1, 2, \ldots, N; \; t = 1, 2, \ldots, T, \tag{1}$$

where the subscript i denotes individuals and t denotes time. This means that i represents the cross-section dimension and t denotes the time series dimension. Here y_{it} is a dependent variable observed for individual i at time t, x_{it} is a $K \times 1$ column vector of observable independent variables ($x_{it}' = (x_{it1}, x_{it2}, \ldots, x_{itK})$), β is a $K \times 1$ column vector of regression parameters, α_i is the unobserved time-invariant individual effect, and ε_{it} denotes the remainder disturbance term that is uncorrelated over time and across cross-sectional units.

The panel data model can be represented in a matrix form, and then

$$y = x\beta + z_\alpha \alpha + \varepsilon, \tag{2}$$

where $y = (y_{11}, y_{12}, \ldots, y_{NT})'$ is an $NT \times 1$ vector of the dependent variable, $x = (x_{11}, x_{12}, \ldots, x_{NT})'$ is an $NT \times K$ matrix, K is the number of explanatory variables, $Z_\alpha = I_N \otimes l_T$, $l_T = (1, 1, \ldots, 1)'$, $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_N)'$ is the $N \times 1$ vector of individual effects coefficients, $\otimes$ is the Kronecker product, and ε is an $NT \times 1$ vector.

3. An Intervention Effects Outlier Model Based Variance

In a panel data model the standard assumption is that the majority of data follow a certain specified distribution. Unfortunately, a certain small percentage of the panel data take values unlikely to follow this same distribution. The residuals are likely to be contaminated with outliers in a panel data model. This means that the error term has the deviation of variance in the model. Since the main purpose of this paper is to propose a method of outlier detection based on the remainder disturbance, it will be assumed that there is no intervention effects problem based on variance if α_i is present, and therefore this paper will not deal with the inference with

the variance of α_i. The following is the panel data model for (1):

$$y_{it} = \alpha_i + x_{it}'\beta + \varepsilon_{it}, \quad i = 1, 2, \ldots, N; \; t = 1, 2, \ldots, T. \quad (3)$$

Suppose outliers interfere with the remainder disturbance ε_{it} and there are the following two conditions. One is variation across both individuals and time where $\mathrm{var}(\varepsilon_{jt}) = \sigma_{jt}^2$; the other is variation across only individuals where $\mathrm{var}(\varepsilon_{jt}) = \sigma_j^2$. Here deviations of the variance across only individual can be assumed, so the variance of the remainder disturbance will be

$$\begin{aligned} \mathrm{var}\left(\varepsilon_j\right) &= \sigma_\varepsilon^2 f_\varepsilon\left(\theta_{\varepsilon_j}\right) I_T, \quad \theta_{\varepsilon_j} \neq 0, \\ \mathrm{var}\left(\varepsilon_i\right) &= \sigma_\varepsilon^2 I_T, \quad i = 1, 2, \ldots, N, \; i \neq j, \end{aligned} \quad (4)$$

where $f_\varepsilon(\theta_\varepsilon)$ is arbitrary strictly positive twice continuously differentiable function satisfying the conditions $f_\varepsilon(\theta_\varepsilon) > 0$, $f_\varepsilon(0) = 1$, and $f_\varepsilon^{(1)}(0) \neq 0$. $f_\varepsilon^{(1)}(x)$ denotes the first derivative of $f_\varepsilon(x)$ with respect to x.

The panel data outlier model is based on the variance of remainder disturbance in a matrix form, and then

$$y = x\beta + z_\alpha \alpha + \varepsilon, \quad \varepsilon \sim \left(0, \Omega_\varepsilon\right), \quad (5)$$

where $\Omega_\varepsilon = \sigma_\varepsilon^2[\mathrm{diag}(I_T, I_T, \ldots, f_\varepsilon(\theta_{\varepsilon_j})I_T, I_T, \ldots, I_T)]$ and $\theta_{\varepsilon_j} \neq 0$. $f_\varepsilon(\theta_{\varepsilon_j})I_T$ denotes the jth diagonal matrix. It is the jth individual variance interventions' effects model. If the outliers impact not only the jth individual observations, but also the subsequent observations, then without loss of generality, it can be assumed that the outliers employ sequence impacts, which are respectively, $f_\varepsilon(\theta_{\varepsilon_{j_1}}), f_\varepsilon(\theta_{\varepsilon_{j_2}}), \ldots, f_\varepsilon(\theta_{\varepsilon_{j_k}})$; then $\Omega_\varepsilon = \sigma_\varepsilon^2[I_{NT} - \sum_{j\in J}(1 - f_\varepsilon(\theta_{\varepsilon_j}))D_j]$; here $J = \{j_1, j_2, \ldots, j_k\}$. D_j denotes a $NT \times NT$ matrix with its jth diagonal matrix being I_T and other matrixes being zeros.

Remark 1. The research scholars of science subjects such as applied mathematics, statistics, and engineering have made active and useful contributions to the detection of outliers and provided some methods to identify outliers, such as Grubbs test, t-test, Dixon test, and Nair test. Grubbs test and Dixon test are not generally effective against identification of multiple outliers. t-test calculation is more complex and generally used for small samples. Nair test requires variance which is assumed to be known. A number of observations do not contain only one outlier, but there are several outliers; thus, the traditional methods especially for detection of a single outlier have less resistance for the pollution of multiple outliers and are highly likely producing the shielding effect once when there are multiple outliers in the sample data.

Remark 2. The outlier model with intervention effects based variance can be applied in a wide range and considers the case in which there are multiple outliers in the observations. By hypothesis testing to construct a valid test statistic, this model develops a new method for the detection of multivariate outliers. It has a strong ability to resist the pollution from outliers. This method can avoid the shielding effect due to the presence of multivariate outliers and also can be used continuously for the detection of multiple outliers.

4. Maximum Likelihood (ML) Estimation

Under the assumption of normality, the log-likelihood function for the model (2) can be written as

$$\begin{aligned} L\left(y \mid x; \beta, \sigma_\varepsilon^2, \sigma_\alpha^2\right) &= c - \frac{1}{2}\ln|\Omega| - \frac{1}{2}u'\Omega^{-1}u \\ &= c - \frac{1}{2}\sum_{i=1}^{N}\ln\left|\Omega_i\right| - \frac{1}{2}\sum_{i=1}^{N}u_i'\Omega_i^{-1}u_i, \end{aligned} \quad (6)$$

where $u = y - x\beta$ and $u_i = y_i - x_i\beta$. Ω is the variance-covariance matrix of the error term u. The variance-covariance matrix can be computed as

$$\begin{aligned} \Omega &= E\left(uu'\right) = Z_\alpha E\left(\alpha\alpha'\right) Z_\alpha' + E\left(\varepsilon\varepsilon'\right) \\ &= \sigma_\alpha^2\left(I_N \otimes J_T\right) + \sigma_\varepsilon^2 I_{NT}, \end{aligned} \quad (7)$$

where $\Omega_i = E(u_i u_i') = \sigma_\alpha^2 l_T l_T' + \sigma_\varepsilon^2 I_T$ and $l_T = (1, 1, \ldots, 1)'$ is the $T \times 1$ column vector. In order to obtain the ML estimator of the regression coefficients, Ω^{-1} needs to be computed as

$$\begin{aligned} \Omega^{-1} &= \sigma_\varepsilon^{-2} I_{NT} - \sigma_\alpha^2\sigma_\varepsilon^{-2}\left(T\sigma_\alpha^2 + \sigma_\varepsilon^2\right)^{-1}\mathrm{diag}\left[J_T, \ldots, J_T\right], \\ \Omega_i^{-1} &= \sigma_\varepsilon^{-2}\left(I_T - \frac{\sigma_\alpha^2}{T\sigma_\alpha^2 + \sigma_\varepsilon^2} l_T l_T'\right). \end{aligned} \quad (8)$$

The maximum likelihood estimators of β, σ_α^2, and σ_μ^2 are obtained by solving the following normal equations:

$$\begin{aligned} \frac{\partial L}{\partial \beta} &= x'\Omega^{-1}y - \left(x'\Omega^{-1}x\right)\beta = x'\Omega^{-1}u = 0, \\ \frac{\partial L}{\partial \sigma_\varepsilon^2} &= -\frac{1}{2}\mathrm{tr}\,\Omega^{-1} + \frac{1}{2}u'\Omega^{-2}u = 0, \\ \frac{\partial L}{\partial \sigma_\alpha^2} &= -\frac{1}{2}\mathrm{tr}\,\Omega^{-1}\left(I_N \otimes J_T\right) + \frac{1}{2}u'\Omega^{-2}\left(I_N \otimes J_T\right)u = 0. \end{aligned} \quad (9)$$

It can be noted that these equations would be nonlinear and difficult to solve explicitly. Let $\tilde{\beta}$, $\tilde{\sigma}_\varepsilon^2$, and $\tilde{\sigma}_\alpha^2$ denote the ML estimates of β, σ_ε^2, and σ_α^2.

5. Outlier Detection and Testing

In the following context, outliers in panel data will be detected and the corresponding LM test for the variance intervention effects model will be derived.

5.1. LM Test with a General Type of Outlier for $H_0 : \theta_\varepsilon = 0$. The log-likelihood function for variance intervention effects model can be written as

$$L\left(y \mid x; \beta, \rho\right) = c - \frac{1}{2}\ln|\Omega| - \frac{1}{2}u'\Omega^{-1}u, \quad (10)$$

where $\rho = (\sigma_\varepsilon^2, \sigma_\alpha^2, \theta)' = \{\rho_r\}$.

The Hessian matrix of L can be obtained:

$$H = \begin{bmatrix} \dfrac{\partial^2}{\partial\beta\partial\beta'} & \dfrac{\partial^2}{\partial\beta\partial\rho'} \\ \dfrac{\partial^2}{\partial\rho\partial\beta'} & \dfrac{\partial^2}{\partial\rho\partial\rho'} \end{bmatrix}. \tag{11}$$

Following Magnus [25], Lejeune [26], and Baltagi [24], the information matrix is given by $\psi = -E_0[H \mid x; \beta, \rho]$ denoting expectation taken with respect to the true distribution. To calculate the information matrix, it is noted that $\partial^2/\partial\beta\partial\beta' = x'\Omega^{-1}x$ and the first derivatives of the likelihood function with respect to β, evaluated at the restricted MLE of β, are zero. Thus, $\operatorname{cov}(\partial L/\partial\beta, \partial L/\partial\rho) = 0$.

Thus the information matrix under the null hypothesis is

$$\psi = -E_0\left[H \mid x; \beta, \rho\right] = \begin{bmatrix} \dfrac{\partial^2}{\partial\beta\partial\beta'} & 0 \\ 0 & \dfrac{\partial^2}{\partial\rho\partial\rho'} \end{bmatrix}. \tag{12}$$

The information matrix is block-diagonal between β and ρ, and the part of the information matrix corresponding to β is ignored in computing the LM statistic since the null hypothesis only involves ρ. Therefore, the LM statistic may be written as

$$\mathrm{LM} = \widetilde{D}_\rho' \widetilde{\psi}_\rho^{-1} \widetilde{D}_\rho, \tag{13}$$

where $\widetilde{D}_\rho = \partial L/\partial\rho(\widetilde{\rho})$ is a vector of partial derivatives of the log-likelihood with respect to each element of ρ, evaluated at the restricted MLE $\widetilde{\rho}$. $\widetilde{\psi}_\rho = E[-\partial^2 L/\partial\rho\partial\rho'](\widetilde{\rho})$ is the information matrix, evaluated at the restricted MLE $\widetilde{\rho}$. Under the null hypothesis, this statistic is asymptotically distributed as a χ^2 with k_ρ degrees of freedom, k_ρ being the number of parameters in the vector ρ.

The first derivative $\widetilde{D}_\rho$ can be obtained as

$$\frac{\partial L}{\partial \rho_r} = -\frac{1}{2}\operatorname{tr}\left[\Omega^{-1}\left(\frac{\partial\Omega}{\partial\rho_r}\right)\right] + \frac{1}{2}\left[u'\Omega^{-1}\left(\frac{\partial\Omega}{\partial\rho_r}\right)\Omega^{-1}u\right]. \tag{14}$$

For the information matrix, the formula is given by Baltagi [24]:

$$\psi_{rs} = E\left[-\frac{\partial^2 L}{\partial\rho_r\partial\rho_s}\right] = \frac{1}{2}\operatorname{tr}\left[\Omega^{-1}\left(\frac{\partial\Omega}{\partial\rho_r}\right)\Omega^{-1}\left(\frac{\partial\Omega}{\partial\rho_s}\right)\right]. \tag{15}$$

Based on (10), log-likelihood function may be written as

$$L\left(y \mid x; \beta, \sigma_\varepsilon^2, \sigma_\alpha^2, \theta_{\varepsilon_j}\right) = c - \frac{1}{2}\ln|\Omega| - \frac{1}{2}u'\Omega^{-1}u. \tag{16}$$

Here

$$\Omega = \sigma_\alpha^2\left(I_N \otimes J_T\right) + \sigma_\varepsilon^2\left(I_{NT} - \sum_{i\in J}\left(1 - f_\varepsilon\left(\theta_{\varepsilon_j}\right)\right)D_j\right). \tag{17}$$

Testing for outliers in this model $H_0 : \Omega_\varepsilon = \sigma_\varepsilon^2 I_{NT} \sim H_1 : \Omega_\varepsilon \neq \sigma_\varepsilon^2 I_{NT}$ amounts to testing $H_0 : \theta_{\varepsilon_j} = 0 \sim H_1 : \theta_{\varepsilon_j} \neq 0$, $j \in J$. Let $\rho^1 = (\sigma_\varepsilon^2, \sigma_\alpha^2, \theta_{\varepsilon_{j_1}}, \ldots, \theta_{\varepsilon_{j_k}})'$, and under the null hypothesis H_0, the inverse matrix of the covariance matrix is

$$\Omega^{-1}\big|_{H_0} = \sigma_\varepsilon^{-2} I_{NT} - \sigma_\alpha^2\sigma_\varepsilon^{-2}\left(T\sigma_\alpha^2 + \sigma_\varepsilon^2\right)^{-1}\operatorname{diag}\left[J_T, \ldots, J_T\right]. \tag{18}$$

Theorem 3. *If $\widetilde{\sigma}_\varepsilon^2, \widetilde{\sigma}_\alpha^2$ denote the ML estimates of $\sigma_\varepsilon^2, \sigma_\alpha^2$, $J = \{j_1, j_2, \ldots, j_k\}$. Outliers interfere with the remainder disturbance ε_{it}. The LM test statistic is given by*

$$\begin{aligned} LM &= \widetilde{D}_\rho' \widetilde{\psi}_\rho^{-1} \widetilde{D}_\rho \\ &= \frac{2}{T\left(1 - 2\lambda + T\lambda^2\right)}\left(\sum_{r=1}^{k} v_r^2 + \frac{1}{N-k}\sum_{r,s=1}^{k} v_r v_s\right), \end{aligned} \tag{19}$$

where

$$\begin{aligned} \lambda &= \widetilde{\sigma}_\alpha^2\left(T\widetilde{\sigma}_\alpha^2 + \widetilde{\sigma}_\varepsilon^2\right)^{-1}, \\ v_r &= -\frac{T}{2}(1-\lambda) + \frac{1}{2}\widetilde{\sigma}_\varepsilon^{-2}\sum_{s=1}^{T}\widetilde{u}_{j_r s}^2 \\ &\quad + \left(-\widetilde{\sigma}_\varepsilon^{-2}\lambda + \frac{\widetilde{\sigma}_\varepsilon^{-2}T\lambda^2}{2}\right)\left(\sum_{s,t=1}^{T}\widetilde{u}_{j_r s}\widetilde{u}_{j_r t}\right). \end{aligned} \tag{20}$$

Under the null hypothesis H_0, this statistic is asymptotically distributed as χ^2 with k degrees of freedom.

5.2. LM Test of Fixed Effects or Random Effects Model with Outliers for $H_0 : \theta_\varepsilon = 0$. This section would consider fixed effects model of panel data with outliers and random effects model of panel data with the rth individual outliers.

Proposition 4. *If $k > 1$ and $\sigma_\alpha^2 = 0$, so the LM statistic for fixed effects model of panel data with a number of individuals outliers is given by*

$$LM = \widetilde{D}_\rho' \widetilde{\psi}_\rho^{-1} \widetilde{D}_\rho = \frac{2}{T}\left(\sum_{r=1}^{k}\tau_r^2 + \frac{1}{N-k}\sum_{r,s=1}^{k}\tau_r\tau_s\right), \tag{21}$$

where $\tau_r = -(T/2) + (1/2)\widetilde{\sigma}_\varepsilon^{-2}\sum_{s=1}^{T}\widetilde{u}_{j_r s}^2$.

Under the null hypothesis H_0, this statistic is asymptotically distributed as χ^2 with k degrees of freedom.

Proposition 5. *If $k = 1$ and $\sigma_\alpha^2 > 0$, so the LM statistic for random effects model of panel data with the rth individual outliers is given by*

$$LM = \widetilde{D}_\rho' \widetilde{\psi}_\rho^{-1} \widetilde{D}_\rho = \frac{2Nv_r^2}{T\left(1 - 2\lambda + T\lambda^2\right)(N-1)}, \tag{22}$$

where $v_r = -(T/2)(1-\lambda) + (1/2)\widetilde{\sigma}_\varepsilon^{-2}\sum_{s=1}^{T}\widetilde{u}_{rs}^2 + (-\widetilde{\sigma}_\varepsilon^{-2}\lambda + (\widetilde{\sigma}_\varepsilon^{-2}T\lambda^2/2))(\sum_{s,t=1}^{T}\widetilde{u}_{rs}\widetilde{u}_{rt})$.

Under the null hypothesis H_0, this statistic is asymptotically distributed as χ^2 with one degree of freedom.

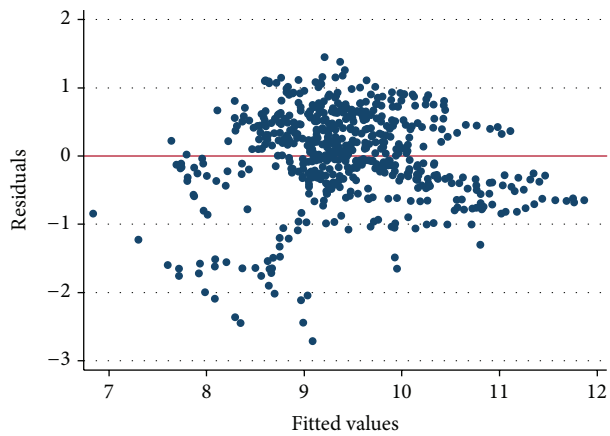

Figure 1: Residual plots of the basic model.

Proposition 6. *If $k = 1$ and $\sigma_\alpha^2 = 0$, so the LM statistic for fixed effects model of panel data with the rth individual outliers is given by*

$$LM = \widetilde{D}_\rho' \widetilde{\psi}_\rho^{-1} \widetilde{D}_\rho = \frac{2N\tau_r^2}{T(N-1)}, \tag{23}$$

where $\tau_r = -(T/2) + (1/2)\widetilde{\sigma}_\varepsilon^{-2} \sum_{s=1}^{T} \widetilde{u}_{rs}^2$.

Under the null hypothesis H_0, this statistic is asymptotically distributed as χ^2 with one degree of freedom.

6. Application Analysis

In this section, an application of the above proposed method using a panel data will be performed and the advantages of the proposed method will be explained.

To evaluate the performance and advantages of the proposed method to detect outliers in panel data, a sample is used as the test dataset to explore main variables' effects on regional environment. The dataset consists of 570 observations for 30 provinces in China from 1992 to 2010. The basic model is

$$\ln \mathrm{CO2}_{i,t} = \alpha_i + \beta_1 \ln \mathrm{EX}_{i,t} + \beta_2 \ln \mathrm{IM}_{i,t} + \beta_3 \ln \mathrm{GP}_{i,t} + \varepsilon_{i,t} \tag{24}$$

with the regional carbon emissions CO_2 as dependent environmental variable, regional export volume EX and import volume IM as main independent trade variables, and regional gross domestic product GP as main regional development level variables. The source data of independent variables are mainly downloaded from National Bureau of Statistics of China and carbon emissions are obtained by following the method of Lin and Sun [27] based on the quantity of fossil fuel consumption data and the CO2 emissions factors of various types of energy from the Intergovernmental Panel on Climate Change (IPCC) reference approach.

6.1. Outlier Detection of the Tested Panel Data. Although the residual plots (Figure 1) of the basic model show that some data points can be initially judged as abnormal, detection of the outliers only from the residual plots is not strict enough because the standard of diagnostic of outliers from residual plots is very vague and even sometimes we will encounter some tougher cases to judge outliers from residual plot. This paper would apply the LM test method to identify outliers accurately.

Table 1: Outlier detection results of the basic model through LM test.

Individual	$\sum_{s=1}^{T} \widetilde{u}_{rs}^2$	$\sum_{s,t=1}^{T} \widetilde{u}_{rs}\widetilde{u}_{rt}$	LM-value	Outlier (yes/no)
#1	56.22	1052.80	1.38	No
#2	41.05	772.05	0.74	No
#3	17.56	330.87	5.23	No
#4	25.51	472.84	0.05	No
#5	7.01	98.36	42.88	Yes
#6	0.10	0.00	6.51	No
#7	0.42	4.76	4.57	No
#8	0.94	10.85	1.29	No
#9	52.52	996.65	7.29	No
#10	1.37	20.05	2.00	No
#11	5.11	95.80	7.22	No
#12	8.73	160.28	2.29	No
#13	16.05	282.91	9.26	No
#14	0.49	3.97	2.47	No
#15	3.38	56.06	0.66	No
#16	25.91	490.25	6.13	No
#17	8.19	153.57	6.16	No
#18	8.93	163.99	2.23	No
#19	8.92	146.43	10.90	Yes
#20	1.47	17.22	0.01	No
#21	68.49	1230.12	275.30	Yes
#22	31.74	599.70	4.51	No
#23	32.49	609.28	0.69	No
#24	3.79	63.53	0.51	No
#25	34.95	648.53	1.55	No
#26	2.73	48.91	4.93	No
#27	10.66	196.21	1.69	No
#28	1.82	23.83	0.01	No
#29	2.60	26.70	10.25	Yes
#30	0.34	2.12	3.42	No

For those nonidentified outlier tests, which mean that we do not know in advance which data is outlier point, the test statistic $\mathrm{LM}_{\max} = \mathrm{LM}_i$. The Bonferroni inequality will be used to approximate the function, and then the process to identify nonidentified outliers with the confidence level as α will be as follows: if $\mathrm{LM}_{\max} \geq \chi^2_{\alpha/n}(1)$ (n is sample size), the data estimated will be outliers.

Let the confidence level $\alpha = 0.05$, and then the critical value $\chi^2_{0.05/30}(1) = 9.88$ for detection of nonidentified outliers. Thus, we can find that individuals such as #5, #19, #21, and #29 in Table 1 are detected as outliers as their LM test statistics are more than the critical value.

When data analysis, it can be probably happened that univariate does not meet the characteristics of structure and

Table 2: Regression results of the basic and correction models.

Model	β_1			β_2			β_3			n
	Coef.	Std. err.	$P > \lvert t \rvert$	Coef	Std. err.	$P > \lvert t \rvert$	Coef	Std. err.	$P > \lvert t \rvert$	
Basic	0.50	0.05	0.00	0.03	0.05	0.58	−0.67	0.09	0.00	570
Correction 1	0.46	0.05	0.00	0.08	0.05	0.09	−0.70	0.08	0.00	494
Correction 2	0.48	0.03	0.00	0.05	0.03	0.05	−0.73	0.05	0.00	494

Table 3: Unit root test results of correction models.

Model	LLC	HT	IPS	Fisher-ADF
Correction 1				
Statistic	−8.7425	−6.3660	−5.8678	−9.8327
P value	0.0000	0.0000	0.0000	0.0000
Correction 2				
Statistic	−8.8354	−5.3912	−6.6452	−9.2965
P value	0.0000	0.0000	0.0000	0.0000

The null hypothesis of LLC, HT, IPS, and Fisher-ADF test is the data which contain a unit root.

correlation among the variables. With the LM test method, significant outliers which could disrupt such relationship can be effectively identified, while the traditional test methods are not in a position to deal with this. Outliers can provide us with valuable information; for example, some significant discoveries often can be obtained on the condition of observations beyond the degree of dispersion of the next random errors. Also, outliers could help us to get an approach of model modification and optimization.

6.2. Correction Model. The following would propose an adjustment model based on above results of the panel model outlier test. After establishing correction model, the outlier test results will be validated by the contrast of the merits of correction model and the original model. The following are the adjustment plan of correction models.

6.2.1. Correction Model 1. As the results of Table 1 show that individuals #5, #19, #21, and #29 are outliers and #21 especially has a higher degree of abnormality, the correction model 1 would exclude the above abnormal individuals. This corrected sample size is 494 and regression results of this model are shown in Table 2.

6.2.2. Correction Model 2. Individuals #5, #19, #21, and #29 have been detected as outliers. Let such abnormal data reflect in the following model, and then the correction model 2 will be as follows:

$$\begin{gathered}\ln \mathrm{CO2}_{i,t} = \alpha_i + \beta_1 \ln \mathrm{EX}_{i,t} + \beta_2 \ln \mathrm{IM}_{i,t} + \beta_3 \ln \mathrm{GP}_{i,t} + \varepsilon_{i,t}, \\ \varepsilon_{it} \sim N\left(0, \sigma_\varepsilon^2\right), \quad i \neq 5, 19, 21, 29, \\ \varepsilon_{5t} \sim N\left(0, \frac{\sigma_\varepsilon^2}{\kappa_1}\right), \qquad \varepsilon_{19t} \sim N\left(0, \frac{\sigma_\varepsilon^2}{\kappa_2}\right), \\ \varepsilon_{21t} \sim N\left(0, \frac{\sigma_\varepsilon^2}{\kappa_3}\right), \qquad \varepsilon_{29t} \sim N\left(0, \frac{\sigma_\varepsilon^2}{\kappa_4}\right).\end{gathered} \tag{25}$$

When taking into account the presence of outliers, the regression results of correction models are compared with the basic model and the estimated value is changing. P value of the basic model shows that the coefficient of ln IM does not pass the test, while the coefficient of the correction model 1 passes the test at 10% confidence level and the correction model 2 passes the test at 5% confidence level. The coefficients of other two variables (ln EX, ln GP) in two correction models pass the test at 1% confidence level. The results in Table 2 show that the indicators of P value and std. err. are improving in two correction models; thus, it can be explained that the presence of outliers does affect the accuracy of the estimated value and correction based on the suggested method allows the regression results to be more accurate.

To further evaluate the regression results of correction models, the residuals will be tested whether they are stable or not, following the methods such as LLC test [28], HT test [29], IPS test [30], and Fisher-ADF test [31]. Table 3 shows the results of the diagnostic test for residuals of panel data. The results show that the correction models have smooth residuals of panel data at the 1% significance level, which indicates that regression results of correction models are better than the basic model with evaluated robustness.

This application case shows the efficiency of LM test to identify outliers and reflects the good effect of detection through avoidance of shielding effect.

7. Conclusion

The presence of individual outliers' properties in panel data can affect parameter estimates and statistical test in a dramatic way. There are some literatures available on time series model with respect to outlier, whereas there are very few studies on panel data models. In case of the existence of individual effect, the traditional mean-shift model cannot be applied in detecting the outliers for panel data with individual effect. Therefore, this paper developed the variance intervention effects model based on the remainder disturbance to detect outliers. Note that the equations of MLE with outliers'

model would be nonlinear and difficult to solve explicitly and directly. These methods of the outlier detection are available through the Lagrange Multiplier (LM) test in panel data. The accurate distribution of the test statistics is not available. Although the test statistics based on LM are asymptotically distributed as χ^2, some researches prove that LM test is also very effective even in small samples. This paper presented a more effective LM testing approach for a general type of outliers' detection based on variance intervention effects model. Furthermore, the corresponding LM test statistics of fixed effects model and random effects model with outliers are obtained. The LM test statistics of an individual outlier model as a particular case are given. This maximum likelihood estimation (ML) can be performed in Stata by using the regress command. Through the implemented estimation, the ML estimates of σ_ε^2 and σ_α^2 are obtained under the null hypothesis. The LM test statistics of outlier detection for each individual are calculated as proposed. Following the procedure, this paper recognized and identified individual outliers by using the results of the LM test statistics. Outlier detection and identification can provide some additional and valuable information which improve the robust direction of statistical model.

Appendix

This appendix derives the LM test for the intervention model with the variance of ε_{it}, and the null hypothesis is given by $H_0 : \theta_{\varepsilon_j} = 0$. Under H_0, the following are obtained:

$$\left.\frac{\partial \Omega}{\partial \sigma_\varepsilon^2}\right|_{H_0} = I_{NT}, \tag{A.1}$$

$$\left.\frac{\partial \Omega}{\partial \sigma_\alpha^2}\right|_{H_0} = I_N \otimes J_T, \tag{A.2}$$

$$\left.\frac{\partial \Omega}{\partial \theta_{\varepsilon_j}}\right|_{H_0} = \sigma_\varepsilon^2 f_\varepsilon'(0) D_j, \tag{A.3}$$

$$\begin{aligned}
\left.\frac{\partial L}{\partial \sigma_\varepsilon^2}\right|_{H_0} &= -\frac{1}{2}\operatorname{tr}\left\{\left\{\tilde{\sigma}_\varepsilon^{-2} I_{NT} - \tilde{\sigma}_\alpha^2 \tilde{\sigma}_\varepsilon^{-2}\left(T\tilde{\sigma}_\alpha^2 + \tilde{\sigma}_\varepsilon^2\right)^{-1} \times \operatorname{diag}\left[J_T, \ldots, J_T\right]\right\} I_{NT}\right\} \\
&\quad + \frac{1}{2}\tilde{u}'\left\{\tilde{\sigma}_\varepsilon^{-2} I_{NT} - \tilde{\sigma}_\alpha^2 \tilde{\sigma}_\varepsilon^{-2}\left(T\tilde{\sigma}_\alpha^2 + \tilde{\sigma}_\varepsilon^2\right)^{-1} \times \operatorname{diag}\left[J_T, \ldots, J_T\right]\right\} I_{NT} \\
&\quad \times \left\{\tilde{\sigma}_\varepsilon^{-2} I_{NT} - \tilde{\sigma}_\alpha^2 \tilde{\sigma}_\varepsilon^{-2}\left(T\tilde{\sigma}_\alpha^2 + \tilde{\sigma}_\varepsilon^2\right)^{-1} \times \operatorname{diag}\left[J_T, \ldots, J_T\right]\right\} \tilde{u} \\
&= -\frac{1}{2}\operatorname{tr}\left\{\left\{\tilde{\sigma}_\varepsilon^{-2} I_{NT} - \tilde{\sigma}_\alpha^2 \tilde{\sigma}_\varepsilon^{-2}\left(T\tilde{\sigma}_\alpha^2 + \tilde{\sigma}_\varepsilon^2\right)^{-1} \times \operatorname{diag}\left[J_T, \ldots, J_T\right]\right\} I_{NT}\right\} \\
&\quad + \frac{1}{2}\tilde{u}'\left\{\tilde{\sigma}_\varepsilon^{-2} I_{NT} - \tilde{\sigma}_\alpha^2 \tilde{\sigma}_\varepsilon^{-2}\left(T\tilde{\sigma}_\alpha^2 + \tilde{\sigma}_\varepsilon^2\right)^{-1} \times \operatorname{diag}\left[J_T, \ldots, J_T\right]\right\}^2 \tilde{u} \\
&= -\frac{NT}{2}\tilde{\sigma}_\varepsilon^{-2} + \frac{NT\tilde{\sigma}_\alpha^2 \tilde{\sigma}_\varepsilon^{-2}\left(T\tilde{\sigma}_\alpha^2 + \tilde{\sigma}_\varepsilon^2\right)^{-1}}{2} \\
&\quad + \frac{1}{2}\tilde{u}'\left\{\tilde{\sigma}_\varepsilon^{-2} I_{NT} - \tilde{\sigma}_\alpha^2 \tilde{\sigma}_\varepsilon^{-2}\left(T\tilde{\sigma}_\alpha^2 + \tilde{\sigma}_\varepsilon^2\right)^{-1} \times \operatorname{diag}\left[J_T, \ldots, J_T\right]\right\}^2 \tilde{u} \\
&= 0,
\end{aligned} \tag{A.4}$$

$$\begin{aligned}
\left.\frac{\partial L}{\partial \sigma_\alpha^2}\right|_{H_0} &= -\frac{1}{2}\operatorname{tr}\left\{\left\{\tilde{\sigma}_\varepsilon^{-2} I_{NT} - \tilde{\sigma}_\alpha^2 \tilde{\sigma}_\varepsilon^{-2}\left(T\tilde{\sigma}_\alpha^2 + \tilde{\sigma}_\varepsilon^2\right)^{-1} \times \operatorname{diag}\left[J_T, \ldots, J_T\right]\right\}\left(I_N \otimes J_T\right)\right\} \\
&\quad + \frac{1}{2}\tilde{u}'\left\{\tilde{\sigma}_\varepsilon^{-2} I_{NT} - \tilde{\sigma}_\alpha^2 \tilde{\sigma}_\varepsilon^{-2}\left(T\tilde{\sigma}_\alpha^2 + \tilde{\sigma}_\varepsilon^2\right)^{-1} \times \operatorname{diag}\left[J_T, \ldots, J_T\right]\right\}\left(I_N \otimes J_T\right) \\
&\quad \times \left\{\tilde{\sigma}_\varepsilon^{-2} I_{NT} - \tilde{\sigma}_\alpha^2 \tilde{\sigma}_\varepsilon^{-2}\left(T\tilde{\sigma}_\alpha^2 + \tilde{\sigma}_\varepsilon^2\right)^{-1} \times \operatorname{diag}\left[J_T, \ldots, J_T\right]\right\} \tilde{u} \\
&= -\frac{1}{2}\operatorname{tr}\left\{\left\{\tilde{\sigma}_\varepsilon^{-2} I_{NT} - \tilde{\sigma}_\alpha^2 \tilde{\sigma}_\varepsilon^{-2}\left(T\tilde{\sigma}_\alpha^2 + \tilde{\sigma}_\varepsilon^2\right)^{-1} \times \operatorname{diag}\left[J_T, \ldots, J_T\right]\right\}\left(I_N \otimes J_T\right)\right\} \\
&\quad + \frac{1}{2}\tilde{u}'\left\{\tilde{\sigma}_\varepsilon^{-2} I_{NT} - \tilde{\sigma}_\alpha^2 \tilde{\sigma}_\varepsilon^{-2}\left(T\tilde{\sigma}_\alpha^2 + \tilde{\sigma}_\varepsilon^2\right)^{-1} \times \operatorname{diag}\left[J_T, \ldots, J_T\right]\right\}^2 \left(I_N \otimes J_T\right)\tilde{u} \\
&= -\frac{NT}{2}\tilde{\sigma}_\varepsilon^{-2} + NT^2\frac{\tilde{\sigma}_\alpha^2 \tilde{\sigma}_\varepsilon^{-2}\left(T\tilde{\sigma}_\alpha^2 + \tilde{\sigma}_\varepsilon^2\right)^{-1}}{2} \\
&\quad + \frac{1}{2}\tilde{u}'\left\{\tilde{\sigma}_\varepsilon^{-2} I_{NT} - \tilde{\sigma}_\alpha^2 \tilde{\sigma}_\varepsilon^{-2}\left(T\tilde{\sigma}_\alpha^2 + \tilde{\sigma}_\varepsilon^2\right)^{-1} \times \operatorname{diag}\left[J_T, \ldots, J_T\right]\right\}^2 \left(I_N \otimes J_T\right)\tilde{u} \\
&= 0,
\end{aligned} \tag{A.5}$$

$$\begin{aligned}
\left.\frac{\partial L}{\partial \theta_{\varepsilon_{j_r}}}\right|_{H_0} &= -\frac{1}{2}\operatorname{tr}\left\{\left\{\tilde{\sigma}_\varepsilon^{-2} I_{NT} - \tilde{\sigma}_\alpha^2 \tilde{\sigma}_\varepsilon^{-2}\left(T\tilde{\sigma}_\alpha^2 + \tilde{\sigma}_\varepsilon^2\right)^{-1} \times \operatorname{diag}\left[J_T, \ldots, J_T\right]\right\}\tilde{\sigma}_\varepsilon^2 f_\varepsilon'(0) D_{j_r}\right\} \\
&\quad + \frac{1}{2}\tilde{u}'\left\{\tilde{\sigma}_\varepsilon^{-2} I_{NT} - \tilde{\sigma}_\alpha^2 \tilde{\sigma}_\varepsilon^{-2}\left(T\tilde{\sigma}_\alpha^2 + \tilde{\sigma}_\varepsilon^2\right)^{-1}\right.
\end{aligned}$$

$$\times \operatorname{diag}[J_T,\ldots,J_T]\Big\}\left(\tilde{\sigma}_\varepsilon^2 f_\varepsilon'(0) D_{j_r}\right)$$
$$\times\Big\{\tilde{\sigma}_\varepsilon^{-2} I_{NT} - \tilde{\sigma}_\alpha^2\tilde{\sigma}_\varepsilon^{-2}\left(T\tilde{\sigma}_\alpha^2+\tilde{\sigma}_\varepsilon^2\right)^{-1}$$
$$\times \operatorname{diag}[J_T,\ldots,J_T]\Big\}\tilde{u}$$
$$= -\frac{T}{2} f_\varepsilon'(0) + \frac{T}{2}\tilde{\sigma}_\alpha^2\left(T\tilde{\sigma}_\alpha^2+\tilde{\sigma}_\varepsilon^2\right)^{-1} f_\varepsilon'(0)$$
$$+\frac{1}{2}\tilde{u}'\Big\{f_\varepsilon'(0) D_{j_r} - \tilde{\sigma}_\alpha^2\left(T\tilde{\sigma}_\alpha^2+\tilde{\sigma}_\varepsilon^2\right)^{-1}$$
$$\times f_\varepsilon'(0)\operatorname{diag}[J_T,\ldots,J_T] D_{j_r}\Big\}$$
$$\times\Big\{\tilde{\sigma}_\varepsilon^{-2} I_{NT} - \tilde{\sigma}_\alpha^2\tilde{\sigma}_\varepsilon^{-2}\left(T\tilde{\sigma}_\alpha^2+\tilde{\sigma}_\varepsilon^2\right)^{-1}$$
$$\times \operatorname{diag}[J_T,\ldots,J_T]\Big\}\tilde{u}$$
$$= -\frac{T}{2} f_\varepsilon'(0) + \frac{T}{2}\tilde{\sigma}_\alpha^2\left(T\tilde{\sigma}_\alpha^2+\tilde{\sigma}_\varepsilon^2\right)^{-1} f_\varepsilon'(0)$$
$$+\frac{1}{2}\tilde{u}'\Big\{\tilde{\sigma}_\varepsilon^{-2} f_\varepsilon'(0) D_{j_r} - 2\tilde{\sigma}_\alpha^2\tilde{\sigma}_\varepsilon^{-2} f_\varepsilon'(0)$$
$$\times\left(T\tilde{\sigma}_\alpha^2+\tilde{\sigma}_\varepsilon^2\right)^{-1}$$
$$\times \operatorname{diag}[J_T,\ldots,J_T] D_{j_r}$$
$$+T\tilde{\sigma}_\alpha^4\tilde{\sigma}_\varepsilon^{-2}\left(T\tilde{\sigma}_\alpha^2+\tilde{\sigma}_\varepsilon^2\right)^{-2}$$
$$\times f_\varepsilon'(0)\operatorname{diag}[J_T,\ldots,J_T] D_{j_r}\Big\} u$$
$$= -\frac{T}{2} f_\varepsilon'(0) + \frac{T}{2}\tilde{\sigma}_\alpha^2\left(T\tilde{\sigma}_\alpha^2+\tilde{\sigma}_\varepsilon^2\right)^{-1} f_\varepsilon'(0)$$
$$+\frac{1}{2}\tilde{\sigma}_\varepsilon^{-2} f_\varepsilon'(0)\sum_{s=1}^{T}\tilde{u}_{j_r s}^2$$
$$-\tilde{\sigma}_\alpha^2\tilde{\sigma}_\varepsilon^{-2} f_\varepsilon'(0)\left(T\tilde{\sigma}_\alpha^2+\tilde{\sigma}_\varepsilon^2\right)^{-1}\sum_{s,t=1}^{T}\left(\tilde{u}_{j_r s}\cdot\tilde{u}_{j_r t}\right)$$
$$+\frac{T}{2}\tilde{\sigma}_\alpha^4\tilde{\sigma}_\varepsilon^{-2}\left(T\tilde{\sigma}_\alpha^2+\tilde{\sigma}_\varepsilon^2\right)^{-2} f_\varepsilon'(0)\sum_{s,t=1}^{T}\left(\tilde{u}_{j_r s}\tilde{u}_{j_r t}\right). \tag{A.6}$$

Using (15), the elements of the information matrix are given by

$$\psi_{11} = E\left[-\frac{\partial^2 L}{\partial\left(\sigma_\varepsilon^2\right)^2}\right]$$
$$= \frac{1}{2}\operatorname{tr}\left[\left(\sigma_\varepsilon^{-2} I_{NT} - \sigma_\alpha^2\sigma_\varepsilon^{-2}\left(T\sigma_\alpha^2+\sigma_\varepsilon^2\right)^{-1}\left(I_N\otimes J_T\right)\right)^2\right]$$
$$= \frac{NT}{2}\sigma_\varepsilon^{-4} - NT\sigma_\alpha^2\sigma_\varepsilon^{-4}\left(T\sigma_\alpha^2+\sigma_\varepsilon^2\right)^{-1}$$
$$+\frac{NT^2}{2}\sigma_\alpha^4\sigma_\varepsilon^{-4}\left(T\sigma_\alpha^2+\sigma_\varepsilon^2\right)^{-2},$$

$$\psi_{22} = E\left[-\frac{\partial^2 L}{\partial\left(\sigma_\alpha^2\right)^2}\right]$$
$$= \frac{1}{2}\operatorname{tr}\Big[\left(\sigma_\varepsilon^{-2} I_{NT} - \sigma_\alpha^2\sigma_\varepsilon^{-2}\left(T\sigma_\alpha^2+\sigma_\varepsilon^2\right)^{-1}\left(I_N\otimes J_T\right)\right)$$
$$\times\left(I_N\otimes J_T\right)$$
$$\times\left(\sigma_\varepsilon^{-2} I_{NT} - \sigma_\alpha^2\sigma_\varepsilon^{-2}\left(T\sigma_\alpha^2+\sigma_\varepsilon^2\right)^{-1}\left(I_N\otimes J_T\right)\right)$$
$$\times\left(I_N\otimes J_T\right)\Big]$$
$$= \frac{1}{2}\sigma_\varepsilon^{-4} NT^2 - \sigma_\alpha^2\sigma_\varepsilon^{-4} NT^3\left(T\sigma_\alpha^2+\sigma_\varepsilon^2\right)^{-1}$$
$$+\frac{1}{2}\sigma_\alpha^4\sigma_\varepsilon^{-4} NT^4\left(T\sigma_\alpha^2+\sigma_\varepsilon^2\right)^{-2},$$

$$\psi_{12} = E\left[-\frac{\partial^2 L}{\partial\left(\sigma_\varepsilon^2\right)\partial\left(\sigma_\alpha^2\right)}\right]$$
$$= \frac{1}{2}\operatorname{tr}\Big[\left(\sigma_\varepsilon^{-2} I_{NT} - \sigma_\alpha^2\sigma_\varepsilon^{-2}\left(T\sigma_\alpha^2+\sigma_\varepsilon^2\right)^{-1}\left(I_N\otimes J_T\right)\right)$$
$$\times\left(\sigma_\varepsilon^{-2} I_{NT} - \sigma_\alpha^2\sigma_\varepsilon^{-2}\left(T\sigma_\alpha^2+\sigma_\varepsilon^2\right)^{-1}\left(I_N\otimes J_T\right)\right)$$
$$\times\left(I_N\otimes J_T\right)\Big]$$
$$= \frac{1}{2}\sigma_\varepsilon^{-4} NT - \sigma_\alpha^2\sigma_\varepsilon^{-4}\left(T\sigma_\alpha^2+\sigma_\varepsilon^2\right)^{-1} NT^2$$
$$+\frac{1}{2}\sigma_\alpha^4\sigma_\varepsilon^{-4}\left(T\sigma_\alpha^2+\sigma_\varepsilon^2\right)^{-2} NT^3$$
$$= \psi_{21},$$

$$\psi_{1j_r} = E\left[-\frac{\partial^2 L}{\partial\left(\sigma_\varepsilon^2\right)\partial\left(\theta_{\varepsilon_{j_r}}\right)}\right]$$
$$= \frac{1}{2}\operatorname{tr}\Big[\left(\sigma_\varepsilon^{-2} I_{NT} - \sigma_\alpha^2\sigma_\varepsilon^{-2}\left(T\sigma_\alpha^2+\sigma_\varepsilon^2\right)^{-1}\left(I_N\otimes J_T\right)\right)^2$$
$$\times\sigma_\varepsilon^2 f_\varepsilon'(0) D_{j_r}\Big]$$
$$= \frac{T}{2}\sigma_\varepsilon^{-2} f_\varepsilon'(0) - \sigma_\alpha^2\sigma_\varepsilon^{-2} T\left(T\sigma_\alpha^2+\sigma_\varepsilon^2\right)^{-1} f_\varepsilon'(0)$$
$$+\frac{1}{2}\sigma_\alpha^4\sigma_\varepsilon^{-2} T^2\left(T\sigma_\alpha^2+\sigma_\varepsilon^2\right)^{-2} f_\varepsilon'(0)$$
$$= \psi_{j_r 1},$$

$$\psi_{2j_r} = E\left[-\frac{\partial^2 L}{\partial\left(\sigma_\alpha^2\right)\partial\left(\theta_{\varepsilon_{j_r}}\right)}\right]$$
$$= \frac{1}{2}\operatorname{tr}\Big[\left(\sigma_\varepsilon^{-2} I_{NT} - \sigma_\alpha^2\sigma_\varepsilon^{-2}\left(T\sigma_\alpha^2+\sigma_\varepsilon^2\right)^{-1}\left(I_N\otimes J_T\right)\right)$$
$$\times\left(I_N\otimes J_T\right)$$

$$\times \left(\sigma_\varepsilon^{-2} I_{NT} - \sigma_\alpha^2 \sigma_\varepsilon^{-2} \left(T\sigma_\alpha^2 + \sigma_\varepsilon^2\right)^{-1} \left(I_N \otimes J_T\right)\right)$$
$$\times \sigma_\varepsilon^2 f_\varepsilon'(0) D_{j_r}\Big]$$
$$= \frac{1}{2} \operatorname{tr}\Big[\Big(\sigma_\varepsilon^{-4} \left(I_N \otimes J_T\right) - 2\sigma_\alpha^2 \sigma_\varepsilon^{-4} \left(T\sigma_\alpha^2 + \sigma_\varepsilon^2\right)^{-1}$$
$$\times \left(I_N \otimes J_T\right)^2$$
$$+ \sigma_\alpha^4 \sigma_\varepsilon^{-4} \left(T\sigma_\alpha^2 + \sigma_\varepsilon^2\right)^{-2} \left(I_N \otimes J_T\right)^3\Big)$$
$$\times \sigma_\varepsilon^2 f_\varepsilon'(0) D_{j_r}\Big]$$
$$= \frac{1}{2} T\sigma_\varepsilon^{-2} f_\varepsilon'(0) - T^2 \sigma_\alpha^2 \sigma_\varepsilon^{-2} f_\varepsilon'(0) \left(T\sigma_\alpha^2 + \sigma_\varepsilon^2\right)^{-1}$$
$$+ \frac{1}{2} T^3 \sigma_\alpha^4 \sigma_\varepsilon^{-2} f_\varepsilon'(0) \left(T\sigma_\alpha^2 + \sigma_\varepsilon^2\right)^{-2}$$
$$= \psi_{j_r 2},$$

$$\psi_{j_r j_r} = E\left[-\frac{\partial^2 L}{\partial \left(\theta_{\varepsilon_{j_r}}\right)^2}\right]$$
$$= \frac{1}{2} \operatorname{tr}\Big[\Big(\sigma_\varepsilon^{-2} I_{NT} - \sigma_\alpha^2 \sigma_\varepsilon^{-2} \left(T\sigma_\alpha^2 + \sigma_\varepsilon^2\right)^{-1} \left(I_N \otimes J_T\right)\Big)$$
$$\times \sigma_\varepsilon^2 f_\varepsilon'(0) D_{j_r}$$
$$\times \left(\sigma_\varepsilon^{-2} I_{NT} - \sigma_\alpha^2 \sigma_\varepsilon^{-2} \left(T\sigma_\alpha^2 + \sigma_\varepsilon^2\right)^{-1} \left(I_N \otimes J_T\right)\right)$$
$$\times \sigma_\varepsilon^2 f_\varepsilon'(0) D_{j_r}\Big]$$
$$= \frac{1}{2} \operatorname{tr}\Big[\Big(\sigma_\varepsilon^{-2} I_{NT} - \sigma_\alpha^2 \sigma_\varepsilon^{-2} \left(T\sigma_\alpha^2 + \sigma_\varepsilon^2\right)^{-1} \left(I_N \otimes J_T\right)\Big)^2$$
$$\times \sigma_\varepsilon^4 \left(f_\varepsilon'(0)\right)^2 \left(D_{j_r}\right)^2\Big]$$
$$= \frac{1}{2} \operatorname{tr}\Big[\left(f_\varepsilon'(0)\right)^2 \left(D_{j_r}\right)^2 - 2\sigma_\alpha^2 \left(f_\varepsilon'(0)\right)^2$$
$$\times \left(T\sigma_\alpha^2 + \sigma_\varepsilon^2\right)^{-1} \left(I_N \otimes J_T\right) \left(D_{j_r}\right)^2$$
$$+ \sigma_\alpha^4 \left(f_\varepsilon'(0)\right)^2 \left(T\sigma_\alpha^2 + \sigma_\varepsilon^2\right)^{-2} \left(I_N \otimes J_T\right)^2 \left(D_{j_r}\right)^2\Big]$$
$$= \frac{1}{2} T \left(f_\varepsilon'(0)\right)^2 - T\sigma_\alpha^2 \left(f_\varepsilon'(0)\right)^2 \left(T\sigma_\alpha^2 + \sigma_\varepsilon^2\right)^{-1}$$
$$+ \frac{1}{2} T^2 \sigma_\alpha^4 \left(f_\varepsilon'(0)\right)^2 \left(T\sigma_\alpha^2 + \sigma_\varepsilon^2\right)^{-2},$$

$$\psi_{j_r j_s} = E\left[-\frac{\partial^2 L}{\partial \left(\theta_{\varepsilon_{j_r}}\right) \partial \left(\theta_{\varepsilon_{j_s}}\right)}\right]$$
$$= \frac{1}{2} \operatorname{tr}\Big[\Big(\sigma_\varepsilon^{-2} I_{NT} - \sigma_\alpha^2 \sigma_\varepsilon^{-2} \left(T\sigma_\alpha^2 + \sigma_\varepsilon^2\right)^{-1} \left(I_N \otimes J_T\right)\Big)$$
$$\times \sigma_\varepsilon^2 f_\varepsilon'(0) D_{j_r}$$
$$\times \left(\sigma_\varepsilon^{-2} I_{NT} - \sigma_\alpha^2 \sigma_\varepsilon^{-2} \left(T\sigma_\alpha^2 + \sigma_\varepsilon^2\right)^{-1} \left(I_N \otimes J_T\right)\right)$$
$$\times \sigma_\varepsilon^2 f_\varepsilon'(0) D_{j_s}\Big]$$
$$= 0 \ (r \neq s)$$
$$= \psi_{j_s j_r}. \tag{A.7}$$

The derivative vector with respect to ρ^1 under H_0 is given by

$$\widehat{D}_\rho = \begin{pmatrix} 0 \\ 0 \\ \dfrac{\partial L}{\partial \theta_{\varepsilon_{j_1}}} \\ \vdots \\ \dfrac{\partial L}{\partial \theta_{\varepsilon_{j_k}}} \end{pmatrix}. \tag{A.8}$$

The information matrix evaluated under H_0 is given by

$$\widehat{\psi}_\rho = \begin{pmatrix} \widehat{\psi}_{11} & \widehat{\psi}_{12} & \widehat{\psi}_{1j_1} & \cdots & \widehat{\psi}_{1j_k} \\ \widehat{\psi}_{21} & \widehat{\psi}_{22} & \widehat{\psi}_{2j_1} & \cdots & \widehat{\psi}_{2j_k} \\ \widehat{\psi}_{j_1 1} & \widehat{\psi}_{j_1 2} & \widehat{\psi}_{j_1 j_1} & \cdots & \widehat{\psi}_{j_1 j_k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \widehat{\psi}_{j_k 1} & \widehat{\psi}_{j_k 2} & \widehat{\psi}_{j_k j_1} & \cdots & \widehat{\psi}_{j_k j_k} \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}. \tag{A.9}$$

Let $\lambda = \widetilde{\sigma}_\alpha^2 (T\widetilde{\sigma}_\alpha^2 + \widetilde{\sigma}_\varepsilon^2)^{-1}$,

$$A_{11} = \begin{pmatrix} \widehat{\psi}_{11} & \widehat{\psi}_{12} \\ \widehat{\psi}_{21} & \widehat{\psi}_{22} \end{pmatrix} = \begin{pmatrix} \widetilde{\sigma}_\varepsilon^{-4} \left(\dfrac{NT}{2} - NT\lambda + \dfrac{NT^2}{2}\lambda^2\right) & \widetilde{\sigma}_\varepsilon^{-4} \left(\dfrac{NT}{2} - NT^2\lambda + \dfrac{NT^3}{2}\lambda^2\right) \\ \widetilde{\sigma}_\varepsilon^{-4} \left(\dfrac{NT}{2} - NT^2\lambda + \dfrac{NT^3}{2}\lambda^2\right) & \widetilde{\sigma}_\varepsilon^{-4} \left(\dfrac{NT^2}{2} - NT^3\lambda + \dfrac{NT^4}{2}\lambda^2\right) \end{pmatrix}$$

$$
\begin{aligned}
&= \frac{1}{2}\widetilde{\sigma}_{\varepsilon}^{-4} NT \begin{pmatrix} \left(1-2\lambda+T\lambda^2\right) & (1-T\lambda)^2 \\ (1-T\lambda)^2 & T(1-T\lambda)^2 \end{pmatrix}, \\
A_{12} &= \begin{pmatrix} \widehat{\psi}_{1j_1} & \cdots & \widehat{\psi}_{1j_k} \\ \widehat{\psi}_{2j_1} & \cdots & \widehat{\psi}_{2j_k} \end{pmatrix} \\
&= \begin{pmatrix} \widetilde{\sigma}_{\varepsilon}^{-2} f_{\varepsilon}'(0)\left(\frac{T}{2}-T\lambda+\frac{T^2\lambda^2}{2}\right) & \cdots & \widetilde{\sigma}_{\varepsilon}^{-2} f_{\varepsilon}'(0)\left(\frac{T}{2}-T\lambda+\frac{T^2\lambda^2}{2}\right) \\ \widetilde{\sigma}_{\varepsilon}^{-2} f_{\varepsilon}'(0)\left(\frac{T}{2}-T^2\lambda+\frac{T^3\lambda^2}{2}\right) & \cdots & \widetilde{\sigma}_{\varepsilon}^{-2} f_{\varepsilon}'(0)\left(\frac{T}{2}-T^2\lambda+\frac{T^3\lambda^2}{2}\right) \end{pmatrix} \\
&= \frac{1}{2}\widetilde{\sigma}_{\varepsilon}^{-2} f_{\varepsilon}'(0)\, T \begin{pmatrix} \left(1-2\lambda+T\lambda^2\right) & \cdots & \left(1-2\lambda+T\lambda^2\right) \\ (1-T\lambda)^2 & \cdots & (1-T\lambda)^2 \end{pmatrix}, \\
A_{21} &= \begin{pmatrix} \widehat{\psi}_{j_1 1} & \widehat{\psi}_{j_1 2} \\ \vdots & \vdots \\ \widehat{\psi}_{j_k 1} & \widehat{\psi}_{j_k 2} \end{pmatrix} \\
&= \begin{pmatrix} \widetilde{\sigma}_{\varepsilon}^{-2} f_{\varepsilon}'(0)\left(\frac{T}{2}-T\lambda+\frac{T^2\lambda^2}{2}\right) & \widetilde{\sigma}_{\varepsilon}^{-2} f_{\varepsilon}'(0)\left(\frac{T}{2}-T^2\lambda+\frac{T^3\lambda^2}{2}\right) \\ \vdots & \vdots \\ \widetilde{\sigma}_{\varepsilon}^{-2} f_{\varepsilon}'(0)\left(\frac{T}{2}-T\lambda+\frac{T^2\lambda^2}{2}\right) & \widetilde{\sigma}_{\varepsilon}^{-2} f_{\varepsilon}'(0)\left(\frac{T}{2}-T^2\lambda+\frac{T^3\lambda^2}{2}\right) \end{pmatrix} \\
&= \frac{1}{2}\widetilde{\sigma}_{\varepsilon}^{-2} f_{\varepsilon}'(0)\, T \begin{pmatrix} \left(1-2\lambda+T\lambda^2\right) & (1-T\lambda)^2 \\ \vdots & \vdots \\ \left(1-2\lambda+T\lambda^2\right) & (1-T\lambda)^2 \end{pmatrix}, \\
A_{22} &= \begin{pmatrix} \widehat{\psi}_{j_1 j_1} & \cdots & \widehat{\psi}_{j_1 j_k} \\ \vdots & \ddots & \vdots \\ \widehat{\psi}_{j_k j_1} & \cdots & \widehat{\psi}_{j_k j_k} \end{pmatrix} \\
&= \frac{1}{2} T \left(f_{\varepsilon}'(0)\right)^2 \begin{pmatrix} \left(1-2\lambda+T\lambda^2\right) & 0 & \cdots & 0 \\ 0 & \ddots & \cdots & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & \left(1-2\lambda+T\lambda^2\right) \end{pmatrix}.
\end{aligned}
\tag{A.10}
$$

Next, the corresponding Lagrange Multiplier (LM) test for null hypotheses H_0 is derived:

$$
\begin{aligned}
\mathrm{LM} &= \widetilde{D}_{\rho}' \widetilde{\psi}_{\rho}^{-1} \widetilde{D}_{\rho} = \left(\widetilde{D}_{\theta_{\varepsilon_j}}\right)' \widetilde{\psi}_{\theta_{\varepsilon_j}}^{-1} \left(\widetilde{D}_{\theta_{\varepsilon_j}}\right) \\
&= \left(\widetilde{D}_{\theta_{\varepsilon_j}}\right)' \left(A_{22} - A_{21} A_{11}^{-1} A_{12}\right)^{-1} \left(\widetilde{D}_{\theta_{\varepsilon_j}}\right),
\end{aligned}
\tag{A.11}
$$

$$
A_{11}^{-1} = \frac{2\widetilde{\sigma}_{\varepsilon}^{4}}{NT(1-T)} \begin{pmatrix} -T & 1 \\ 1 & \dfrac{\left(-T\lambda^2+2\lambda-1\right)}{(1-T\lambda)^2} \end{pmatrix},
\tag{A.12}
$$

$$
\begin{aligned}
A_{21} A_{11}^{-1} A_{12} &= \frac{1}{2} \frac{\left(f_{\varepsilon}'(0)\right)^2 T}{N} \\
&\quad \times \begin{pmatrix} 1-2\lambda+T\lambda^2 & \cdots & 1-2\lambda+T\lambda^2 \\ \vdots & \ddots & \vdots \\ 1-2\lambda+T\lambda^2 & \cdots & 1-2\lambda+T\lambda^2 \end{pmatrix},
\end{aligned}
\tag{A.13}
$$

$$
\begin{aligned}
A_{22} - A_{21} A_{11}^{-1} A_{12} &= \frac{\left(f_{\varepsilon}'(0)\right)^2 T \left(1-2\lambda+T\lambda^2\right)}{2N} \\
&\quad \times \begin{pmatrix} N-1 & -1 & \cdots & -1 \\ -1 & N-1 & -1 & \vdots \\ \vdots & -1 & \ddots & -1 \\ -1 & \cdots & -1 & N-1 \end{pmatrix},
\end{aligned}
\tag{A.14}
$$

$$\left(A_{22}-A_{21}A_{11}^{-1}A_{12}\right)^{-1}=\widetilde{\psi}_{\theta_{\varepsilon_j}}^{-1}=\frac{2}{(f_\varepsilon'(0))^2T\left(1-2\lambda+T\lambda^2\right)}\cdot\Gamma, \tag{A.15}$$

where

$$\Gamma=\begin{pmatrix}1+\frac{1}{N-k} & \frac{1}{N-k} & \cdots & \frac{1}{N-k}\\ \frac{1}{N-k} & \ddots & \frac{1}{N-k} & \vdots\\ \vdots & \frac{1}{N-k} & \ddots & \frac{1}{N-k}\\ \frac{1}{N-k} & \cdots & \frac{1}{N-k} & 1+\frac{1}{N-k}\end{pmatrix}. \tag{A.16}$$

Using (A.6), the following is obtained as

$$\begin{aligned}\left.\frac{\partial L}{\partial\theta_{\varepsilon_{j_r}}}\right|_{H_0}&=-\frac{T}{2}f_\varepsilon'(0)(1-\lambda)+\frac{1}{2}\widetilde{\sigma}_\varepsilon^{-2}f_\varepsilon'(0)\sum_{s=1}^{T}\widetilde{u}_{j_rs}^2\\&\quad-\lambda\widetilde{\sigma}_\varepsilon^{-2}f_\varepsilon'(0)\sum_{s,t=1}^{T}\widetilde{u}_{j_rs}\cdot\widetilde{u}_{j_rt}\\&\quad+\frac{T}{2}\lambda^2\widetilde{\sigma}_\varepsilon^{-2}f_\varepsilon'(0)\sum_{s,t=1}^{T}\widetilde{u}_{j_rs}\widetilde{u}_{j_rt}\\&=f_\varepsilon'(0)\left[-\frac{T}{2}(1-\lambda)\right.\\&\quad+\frac{1}{2}\widetilde{\sigma}_\varepsilon^{-2}\sum_{s=1}^{T}\widetilde{u}_{j_rs}^2+\left(-\widetilde{\sigma}_\varepsilon^{-2}\lambda+\frac{\widetilde{\sigma}_\varepsilon^{-2}T\lambda^2}{2}\right)\\&\quad\left.\times\left(\sum_{s,t=1}^{T}\left(\widetilde{u}_{j_rs}\widetilde{u}_{j_rt}\right)\right)\right].\end{aligned} \tag{A.17}$$

Let

$$\begin{aligned}v_r&=-\frac{T}{2}(1-\lambda)+\frac{1}{2}\widetilde{\sigma}_\varepsilon^{-2}\sum_{s=1}^{T}\widetilde{u}_{j_rs}^2\\&\quad+\left(-\widetilde{\sigma}_\varepsilon^{-2}\lambda+\frac{\widetilde{\sigma}_\varepsilon^{-2}T\lambda^2}{2}\right)\left(\sum_{s,t=1}^{T}\left(\widetilde{u}_{j_rs}\widetilde{u}_{j_rt}\right)\right),\end{aligned} \tag{A.18}$$

and then

$$\left.\frac{\partial L}{\partial\theta_{\varepsilon_{j_r}}}\right|_{H_0}=f_\varepsilon'(0)v_r. \tag{A.19}$$

Using (A.11), get

$$\begin{aligned}\mathrm{LM}&=\left(\widetilde{D}_{\theta_{\varepsilon_j}}\right)'\widetilde{\psi}_{\theta_{\varepsilon_j}}^{-1}\left(\widetilde{D}_{\theta_{\varepsilon_j}}\right)\\&=\left(f_\varepsilon'(0)v_1\quad f_\varepsilon'(0)v_2\quad\cdots\quad f_\varepsilon'(0)v_k\right)\\&\quad\times\widetilde{\psi}_{\theta_{\varepsilon_j}}^{-1}\begin{pmatrix}f_\varepsilon'(0)v_1\\f_\varepsilon'(0)v_2\\\vdots\\f_\varepsilon'(0)v_k\end{pmatrix}\\&=\frac{2}{T\left(1-2\lambda+T\lambda^2\right)}\\&\quad\times\left[\left(1+\frac{1}{N-k}\right)\sum_{r=1}^{k}v_r^2+\frac{1}{N-k}\sum_{r,s=1,r\neq s}^{k}v_rv_s\right]\\&=\frac{2}{T\left(1-2\lambda+T\lambda^2\right)}\\&\quad\times\left[\left(1+\frac{1}{N-k}\right)\sum_{r=1}^{k}v_r^2\right.\\&\quad\left.+\frac{1}{N-k}\left(\sum_{r,s=1}^{k}v_rv_s-\sum_{r=1}^{k}v_r^2\right)\right]\\&=\frac{2}{T\left(1-2\lambda+T\lambda^2\right)}\left(\sum_{r=1}^{k}v_r^2+\frac{1}{N-k}\sum_{r,s=1}^{k}v_rv_s\right).\end{aligned} \tag{A.20}$$

Under the null hypothesis H_0, this statistic is asymptotically distributed as χ^2 with k degrees of freedom.

Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (Grant no. 71302054), China Scholarship Council (201308210076), and Dongbei University of Finance & Economics (DUFE2014J13). The author would like to thank the editors and referees whose comments led to an appreciable clarification of the argument of this paper.

References

[1] A. J. Fox, "Outliers in time series," *Journal of the Royal Statistical Society Series B: Methodological*, vol. 34, pp. 350–363, 1972.

[2] R. D. Martin and V. J. Yohai, "Influence functionals for time series," *The Annals of Statistics*, vol. 14, no. 3, pp. 781–818, 1986.

[3] I. Chang, G. C. Tiao, and C. Chen, "Estimation of time series parameters in the presence of outliers," *Technometrics*, vol. 30, no. 2, pp. 193–204, 1988.

[4] V. Verardi and C. Croux, "Robust regression in Stata," *The Stata Journal*, vol. 9, no. 3, pp. 439–453, 2009.

[5] M. Febrero, P. Galeano, and W. González-Manteiga, "Outlier detection in functional data by depth measures, with application to identify abnormal NO_x levels," *Environmetrics*, vol. 19, no. 4, pp. 331–345, 2008.

[6] H. Tan, J. Feng, G. Feng, W. Wang, and Y. J. Zhang, "Traffic volume data outlier recovery via tensor model," *Mathematical Problems in Engineering*, vol. 2013, Article ID 164810, 8 pages, 2013.

[7] J. I. P. Di Blasi, J. Martínez Torres, P. J. García Nieto, J. R. Alonso Fernández, C. Díaz Muñiz, and J. Taboada, "Analysis and detection of outliers in water quality parameters from different automated monitoring stations in the Miño river basin (NW Spain)," *Ecological Engineering*, vol. 60, pp. 60–66, 2013.

[8] D. M. Hawkins, *Identification of Outliers*, Chapman and Hall, London, UK, 1980.

[9] R. D. Cook and S. Weisberg, *Residuals and Influence in Regression*, Chapman and Hall, London, UK, 1982.

[10] R. Schall and T. T. Dunne, "A unified approach to outliers in the general linear model," *Sankhy: The Indian Journal of Statistics B*, vol. 50, no. 2, pp. 157–167, 1988.

[11] V. Barnett and T. Lewis, *Outliers in Statistical Data*, John Wiley & Sons, 1994.

[12] I. H. Langford and T. Lewis, "Outliers in multilevel data," *Journal of the Royal Statistical Society. Series A: Statistics in Society*, vol. 161, no. 2, pp. 121–160, 1998.

[13] P. J. Rousseeuw and A. M. Leroy, *Robust Regression and Outlier Detection*, John Wiley & Sons, New York, NY, USA, 2003.

[14] L. Shi and G. Chen, "Detection of outliers in multilevel models," *Journal of Statistical Planning and Inference*, vol. 138, no. 10, pp. 3189–3199, 2008.

[15] W.-D. Chen, "Detecting and identifying interventions with the Whittle spectral approach in a long memory panel data model," *Journal of Applied Statistics*, vol. 35, no. 7-8, pp. 879–892, 2008.

[16] G. Willems, H. Joe, and R. Zamar, "Diagnosing multivariate outliers detected by robust estimators," *Journal of Computational and Graphical Statistics*, vol. 18, no. 1, pp. 73–91, 2009.

[17] M. Riani, A. C. Atkinson, and A. Cerioli, "Finding an unknown number of multivariate outliers," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, vol. 71, no. 2, pp. 447–466, 2009.

[18] A. Cerioli, "Multivariate outlier detection with high-breakdown estimators," *Journal of the American Statistical Association*, vol. 105, no. 489, pp. 147–156, 2010.

[19] X. F. Yan, "Multivariate outlier detection based on self-organizing map and adaptive nonlinear map and its application," *Chemometrics and Intelligent Laboratory Systems*, vol. 107, no. 2, pp. 251–257, 2011.

[20] K.-V. Yuen and H.-Q. Mu, "A novel probabilistic method for robust parametric identification and outlier detection," *Probabilistic Engineering Mechanics*, vol. 30, pp. 48–59, 2012.

[21] F. Rapallo, "Outliers and patterns of outliers in contingency tables with algebraic statistics," *Scandinavian Journal of Statistics*, vol. 39, no. 4, pp. 784–797, 2012.

[22] S. Kuhnt, F. Rapallo, and A. Rehage, "Outlier detection in contingency tables based on minimal patterns," *Statistics and Computing*, vol. 24, no. 3, pp. 481–491, 2014.

[23] M. C. Bramati and C. Croux, "Robust estimators for the fixed effects panel data model," *Econometrics Journal*, vol. 10, no. 3, pp. 521–540, 2007.

[24] B. H. Baltagi, *Econometric Analysis of Panel Data*, John Wiley & Sons, 2005.

[25] J. R. Magnus, "Maximum likelihood estimation of the GLS model with unknown parameters in the disturbance covariance matrix," *Journal of Econometrics*, vol. 7, no. 3, pp. 281–312, 1978.

[26] B. Lejeune, "A full heteroscedastic one-way error components model for incomplete panel: maximum likelihood estimation and Lagrange multiplier testing," CORE Discussion Papers 1996006, 1996.

[27] B. Lin and C. Sun, "Evaluating carbon dioxide emissions in international trade of China," *Energy Policy*, vol. 38, no. 1, pp. 613–621, 2010.

[28] A. Levin, C.-F. Lin, and C. J. Chu, "Unit root tests in panel data: asymptotic and finite-sample properties," *Journal of Econometrics*, vol. 108, no. 1, pp. 1–24, 2002.

[29] R. D. Harris and E. Tzavalis, "Inference for unit roots in dynamic panels where the time dimension is fixed," *Journal of Econometrics*, vol. 91, no. 2, pp. 201–226, 1999.

[30] K. S. Im, M. H. Pesaran, and Y. Shin, "Testing for unit roots in heterogeneous panels," *Journal of Econometrics*, vol. 115, no. 1, pp. 53–74, 2003.

[31] I. Choi, "Unit root tests for panel data," *Journal of International Money and Finance*, vol. 20, no. 2, pp. 249–272, 2001.