International Journal of Statistics and Probability; Vol. 10, No. 3; May 2021
ISSN 1927-7032 E-ISSN 1927-7040
Published by Canadian Center of Science and Education

Empirical Likelihood Ratio Test for Seemingly Unrelated Regression Models

Chuanhua Wei[1] & Xiaoxiao Ma[1]

[1] School of Science, Minzu University of China, Beijing 100081, China

Correspondence: Correspondence: Chuanhua Wei, School of Science, Minzu University of China, Beijing 100081, P.R.China.

Received: November 26, 2020 Accepted: January 23, 2021 Online Published: March 4, 2021

doi:10.5539/ijsp.v10n3p1 URL: https://doi.org/10.5539/ijsp.v10n3p1

Abstract

This paper considers the problem of testing independence of equations in a seemingly unrelated regression model. A novel empirical likelihood test approach is proposed, and under the null hypothesis it is shown to follow asymptotically a chi-square distribution. Finally, simulation studies and a real data example are conducted to illustrate the performance of the proposed method.

Keywords: seemingly unrelated regression, empirical likelihood, independence

1. Introduction

The Seemingly Unrelated Regression (SUR) of Zellner (1962) is an important tool to analyze multiple equations with correlated disturbances. SUR models have been studied extensively by statistician and econometrician and applied in many areas, more details can be found in Srivastava and Giles (1987) and Fiebig (2001).

Due to the correlation of the model errors in regression equations, the SUR model allows one to estimate the regression coefficients more efficiently than each of the regression equations is estimated separately with the correlation is ignored. It is by now clear that for the traditional linear SUR model, the Generalized Least Squares (GLS) estimator is more efficient than its Ordinary Least Squares (OLS) counterpart. They are equivalence if the error covariance of the SUR model is diagonal. Therefore, the problem of testing independence of equations of a SUR model is important. Many testing approaches have been proposed for this problem. Breusch and Pagan (1980) proposed a Lagrange multiplier test statistic. Dufour and Khalaf (2002) extended the exact independence test method of Harvey and Phillips (1982) to the multi-equation framework. Tsay (2004) constructed a multivariate independent test statistic for SUR model with serially correlated errors.

Different to the above methods, we propose a empirical likelihood test statistic. The empirical likelihood of Owen (1988,1990) is an effective nonparametric inference method. More references can be found in Owen (2001).

The paper is organized as follows. The empirical log-likelihood ratio test statistic is given in Section 2. Section 3 conducts some simulation studies to illustrate the performance of the proposed method. An empirical example is also provided to demonstrate the usefulness of this test. Finally, conclusion is given in Section 4. The Appendix provides the proofs of the main results.

2. Test Statistic and Its Properties

Consider the following SUR model that comprises the p regression equations

$$\mathbf{Y}_i = \mathbf{X}_i\boldsymbol{\beta}_i + \boldsymbol{\varepsilon}_i, \quad i = 1, 2, \cdots, p, \tag{2.1}$$

with $\mathbf{Y}_i = (y_{i1}, y_{i2}, \cdots, y_{in})^{\mathrm{T}}$ is a $n \times 1$ vector of responses, $\mathbf{X}_i = (\mathbf{x}_{i1}, \mathbf{x}_{i2}, \cdots, \mathbf{x}_{in})^{\mathrm{T}}$ is a full-rank $n \times q_i$ matrix of regressors with $\mathbf{x}_{ik}^{\mathrm{T}} = (x_{ik1}, x_{ik2}, \cdots, x_{ikq_i})$, $\boldsymbol{\beta}_i = (\beta_{i1}, \beta_{i2}, \cdots, \beta_{iq_i})^{\mathrm{T}}$ is a vector of q_i-dimensional unknown parameters, and $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1}, \varepsilon_{i2}, \cdots, \varepsilon_{in})^{\mathrm{T}}$ is a $n \times 1$ error vector with $E\varepsilon_{ik} = 0, k = 1, 2, \cdots, n$.

The model (2.1) can be rewritten in vectors and matrixces,

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \tag{2.2}$$

where

$$\mathbf{Y}=\begin{bmatrix}\mathbf{Y}_1\\ \mathbf{Y}_2\\ \vdots\\ \mathbf{Y}_p\end{bmatrix},\ \mathbf{X}=\begin{bmatrix}\mathbf{X}_1 & \mathbf{0} & \cdots & \mathbf{0}\\ \mathbf{0} & \mathbf{X}_2 & \cdots & \mathbf{0}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{X}_p\end{bmatrix},\ \boldsymbol{\beta}=\begin{bmatrix}\boldsymbol{\beta}_1\\ \boldsymbol{\beta}_2\\ \vdots\\ \boldsymbol{\beta}_p\end{bmatrix},\ \boldsymbol{\varepsilon}=\begin{bmatrix}\boldsymbol{\varepsilon}_1\\ \boldsymbol{\varepsilon}_2\\ \vdots\\ \boldsymbol{\varepsilon}_p\end{bmatrix}$$

so that $\mathbf{X}$ is a $(np)\times q$ matrix, $\mathbf{Y}$ and $\boldsymbol{\varepsilon}$ each have dimension $(np)\times 1$ and $\boldsymbol{\beta}$ has dimension $q\times 1$, with $q=\sum_{i=1}^{p} q_i$. The basic assumptions underlying the disturbances of model (2.1) are

$$E(\varepsilon_{ik}\varepsilon_{jl})=\begin{cases}\sigma_{ij}, & k=l,\\ 0, & \text{otherwise},\end{cases}$$

for $1\le i,j\le p$ and $1\le k,l\le n$. Then, we have $Var(\boldsymbol{\varepsilon}_i)=E\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i^{\mathrm{T}}=\sigma_{ii}\mathbf{I}_n$, and $Cov(\boldsymbol{\varepsilon}_i,\boldsymbol{\varepsilon}_j)=E\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_j^{\mathrm{T}}=\sigma_{ij}\mathbf{I}_n$, with $\mathbf{I}_n$ is the identity matrix of order n. Therefore, the $np\times 1$ disturbance vector $\boldsymbol{\varepsilon}$ has the following variance-covariance matrix

$$\boldsymbol{\Omega}=E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{\mathrm{T}})=\boldsymbol{\Sigma}\otimes\mathbf{I}_n, \tag{2.3}$$

with

$$\boldsymbol{\Sigma}=\begin{bmatrix}\sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p}\\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p}\\ \vdots & \vdots & \ddots & \vdots\\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp}\end{bmatrix}.$$

We consider the problem of testing independence of p equations in model (2.1), which may be expressed as $H_0:\sigma_{ij}=0$ for $1\le i<j\le p$, or equivalently

$$H_0:\boldsymbol{\Sigma}=\begin{bmatrix}\sigma_{11} & 0 & \cdots & 0\\ 0 & \sigma_{22} & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \sigma_{pp}\end{bmatrix}. \tag{2.4}$$

Letting $\mathbf{U}_k=(\varepsilon_{1k}\varepsilon_{2k},\varepsilon_{1k}\varepsilon_{3k},\cdots,\varepsilon_{1k}\varepsilon_{pk},\varepsilon_{2k}\varepsilon_{3k},\cdots,\varepsilon_{2k}\varepsilon_{pk},\cdots,\varepsilon_{(p-1)k}\varepsilon_{pk})^{\mathrm{T}}, k=1,2,\cdots,n$, it is obvious that there are $N=\frac{p(p-1)}{2}$ elements in $\mathbf{U}_k$. For example, for the two equations case $p=2$, we have $\mathbf{U}_k=\varepsilon_{1k}\varepsilon_{2k}$ and $N=1$, for the three equations case $p=3$, $\mathbf{U}_k=(\varepsilon_{1k}\varepsilon_{2k},\varepsilon_{1k}\varepsilon_{3k},\varepsilon_{2k}\varepsilon_{3k})^{\mathrm{T}}$ and $N=3$. It is obvious that testing for diagonality of the $\boldsymbol{\Sigma}$ is equivalent to testing whether $E\mathbf{U}_k=\mathbf{0}, k=1,2,\cdots,n$. By Owen(1990), this can be done using the empirical likelihood method. Let $p_1,p_2,\cdots,p_n$ be nonnegative numbers summing to unity. Then the corresponding empirical log-likelihood ratio can be defined as

$$\bar{l}_n=-2\max\left\{\sum_{k=1}^{n}\log(np_k):\sum_{k=1}^{n}p_k\mathbf{U}_k=0, p_k\ge 0, \sum_{k=1}^{n}p_k=1\right\}. \tag{2.5}$$

However, $\varepsilon_{ik}'s$ in $\mathbf{U}_k$ are unknown, then $\bar{l}_n$ cannot be used directly. To solve the problem, we can replace ε_{ik} by its estimator

$$\hat{\varepsilon}_{ik}=y_{ik}-\mathbf{x}_{ik}^{\mathrm{T}}\hat{\boldsymbol{\beta}}_i,$$

with $\hat{\boldsymbol{\beta}}_i=(\mathbf{X}_i^{\mathrm{T}}\mathbf{X}_i)^{-1}\mathbf{X}_i^{\mathrm{T}}\mathbf{Y}_i$ is the least-squares estimator of the coefficients contained in the ith equation of model (1.1). Then, use $\hat{\varepsilon}_{ik}$ to replace ε_{ik} in $\mathbf{U}_k$, the estimated empirical log-likelihood ratio is then defined by

$$l_n=-2\max\left\{\sum_{k=1}^{n}\log(np_k):\sum_{k=1}^{n}p_k\boldsymbol{\xi}_k=0, p_k\ge 0, \sum_{k=1}^{n}p_k=1\right\}, \tag{2.6}$$

where $\boldsymbol{\xi}_k=(\hat{\varepsilon}_{1k}\hat{\varepsilon}_{2k},\hat{\varepsilon}_{1k}\hat{\varepsilon}_{3k},\cdots,\hat{\varepsilon}_{1k}\hat{\varepsilon}_{pk},\hat{\varepsilon}_{2k}\hat{\varepsilon}_{3k},\cdots,\hat{\varepsilon}_{2k}\hat{\varepsilon}_{pk},\cdots,\hat{\varepsilon}_{(p-1)k}\hat{\varepsilon}_{pk})$.

By the Lagrange multiplier technique, the empirical log-likelihood ratio can be represented as

$$l_n=2\sum_{k=1}^{n}\log\left(1+\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\xi}_k\right), \tag{2.7}$$

where $\boldsymbol{\lambda}=(\lambda_1,\lambda_2,\cdots,\lambda_n)^T$ is the solution of the equation

$$\frac{1}{n}\sum_{k=1}^{n}\frac{\boldsymbol{\xi}_k}{1+\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\xi}_k}=0. \tag{2.8}$$

The following theorem indicates that l_n is asymptotically distributed as a χ^2-distribution.

Theorem 2.1. Suppose the assumptions 1-2 given in Appendix hold, under the null hypothesis, as $n\to\infty$, we have

$$l_n \xrightarrow{D} \chi^2_N,$$

where χ^2_N is a χ^2-distribution with $N = \frac{p(p-1)}{2}$ degrees of freedom.

Remark 2.1 For the testing problem (2.4), Breusch and Pagan (1980) proposed a Lagrange multiplier test statistic. This is based upon the sample correlation coefficients of the OLS residuals:

$$\mathrm{LM} = n\sum_{i=1}^{p-1}\sum_{j=i}^{n}\hat{\rho}_{ij}^2,$$

where $\hat{\rho}_{ij}$ is the sample estimate of the pair-wise correlation of the residuals. Specifically,

$$\hat{\rho}_{ij} = \hat{\rho}_{ji} = \frac{\sum\limits_{k=1}^{n}\hat{\varepsilon}_{ik}\hat{\varepsilon}_{jk}}{(\sum\limits_{k=1}^{n}\hat{\varepsilon}_{ik}^2)^{1/2}(\sum\limits_{k=1}^{n}\hat{\varepsilon}_{jk}^2)^{1/2}}$$

Under the null hypothesis, LM has an asymptotic χ^2_N distribution, too.

3. Numerical Studies

3.1 Simulation Studies

In this subsection, we conducted some simulations to illustrate the finite sample properties of the proposed test procedure. In our simulations, the data are generated from the following SUR model

$$y_{ik} = x_{ik1}\beta_{i1} + x_{ik2}\beta_{i2} + \varepsilon_i, i = 1, 2, 3, k = 1, 2, \cdots, n$$

where $x_{ik1} \sim N(0, 1)$, $x_{ik2} \sim U(-2, 2)$, and $x_{ik3} \sim N(2, 1)$. The parameters are set as $\beta_{11} = 1, \beta_{12} = 2, \beta_{21} = 2, \beta_{22} = 3, \beta_{31} = -1, \beta_{32} = 3$. The model error $\varepsilon_{ik} \sim N(0, \sigma_{ii})$ and

$$\mathrm{Cov}[(\varepsilon_{1k}, \varepsilon_{2k}, \varepsilon_{3k})^{\mathrm{T}}] = \begin{bmatrix} \sigma_{11} & \rho_{12}\sqrt{\sigma_{11}\sigma_{22}} & \rho_{13}\sqrt{\sigma_{11}\sigma_{33}} \\ \rho_{12}\sqrt{\sigma_{11}\sigma_{22}} & \sigma_{22} & \rho_{23}\sqrt{\sigma_{22}\sigma_{33}} \\ \rho_{13}\sqrt{\sigma_{11}\sigma_{33}} & \rho_{23}\sqrt{\sigma_{22}\sigma_{33}} & \sigma_{33} \end{bmatrix}.$$

where $\sigma_{11} = 0.25, \sigma_{22} = 0.64, \sigma_{33} = 0.49$.

In order to examine the empirical size of the proposed empirical likelihood (EL) test and the Lagrange multiplier (LM) test statistic, we set $\boldsymbol{\rho} = (\rho_{12}, \rho_{13}, \rho_{23}) = (0, 0, 0)$, and $n = 30, 50, 100, 150, 200, 300, 400$, 1000 replications were run and the rejection rate under a given significance level $\alpha(0.01, 0.05, 0.10)$ was computed as the empirical size of the test, and the results are reported in Table 3.1. From the results, we can see that the empirical size of the proposed EL test is quite large for small samples. The size distortion of the LM test is smaller than that of the EL test for small samples. The sizes of both the EL test and the LM test converge to the correct nominal levels when n grows, as would be expected. The fact that the size distortion of the EL test is relatively large indicates that the approximation of the finite sample distribution in small samples using the asymptotic χ^2 is relatively poor. The phenomenon was also reported by Dong and Giles (2007), Liu *et al.* (2008) and Liu *et al.* (2011) in other testing problems. According to Owen (2001), this may be improved by using Fisher's F-distribution, or Bartlett correction, or bootstrap sample.

To assess the power of the EL and the LM tests, we took the values of $\boldsymbol{\rho}$ to be each of the following values, (0.1,0,0), (0,0.5,0), (0,0,-0.9), (0.2,0.3,0), (-0.5,0,0.4), (0,-0.5,-0.8), (0.1,-0.2,0.1), (0.1,0.2,0.8), (-0.5,0.4,-0.6), and $n = 30, 50$. Results are presented in Table 3.2. we can see that the power of the EL is bigger than that of the LM test for significance levels of 10%, 5%, and 1%.

Table 3.1. Empirical sizes of EL and LM tests

n	$\alpha =$	0.01	$\alpha =$	0.05	$\alpha =$	0.10
	EL	LM	EL	LM	EL	LM
30	0.067	0.008	0.142	0.048	0.215	0.095
50	0.032	0.014	0.106	0.040	0.155	0.096
100	0.021	0.009	0.067	0.049	0.121	0.090
150	0.013	0.012	0.053	0.050	0.107	0.104
200	0.011	0.009	0.053	0.047	0.119	0.096
300	0.011	0.012	0.047	0.047	0.111	0.102
400	0.011	0.009	0.053	0.051	0.104	0.095

Table 3.2. Power comparison of the EL test with the LM test

n	$\boldsymbol{\rho}$	$\alpha =$	0.01	$\alpha =$	0.05	$\alpha =$	0.10
		EL	LM	EL	LM	EL	LM
30	(0.1,0,0)	0.092	0.011	0.183	0.061	0.258	0.126
	(0,0.5,0)	0.594	0.286	0.765	0.579	0.832	0.691
	(0,0,-0.9)	1.000	1.000	1.000	1.000	1.000	1.000
	(0.2,0.3,0)	0.307	0.096	0.508	0.289	0.606	0.406
	(-0.5,0,0.4)	0.858	0.578	0.96	0.858	0.977	0.924
	(0,-0.5,-0.8)	1.000	1.000	1.000	1.000	1.000	1.000
	(0.1,0.2,-0.1)	0.183	0.036	0.308	0.144	0.419	0.231
	(0.1,0.2,0.8)	0.993	0.963	0.997	0.998	0.999	0.998
	(-0.5,0.4,-0.6)	0.936	0.898	0.978	0.959	0.986	0.989
50	(0.1,0,0)	0.062	0.019	0.158	0.086	0.247	0.154
	(0,0.5,0)	0.800	0.659	0.923	0.874	0.962	0.934
	(0,0,-0.9)	1.000	1.000	1.000	1.000	1.000	1.000
	(0.2,0.3,0)	0.460	0.258	0.616	0.512	0.741	0.655
	(-0.5,0,0.4)	0.988	0.941	0.996	0.991	0.997	0.994
	(0,-0.5,-0.8)	1.000	1.000	1.000	1.000	1.000	1.000
	(0.1,0.2,-0.1)	0.226	0.062	0.396	0.232	0.489	0.360
	(0.1,0.2,0.8)	0.999	1.000	1.000	1.000	1.000	1.000
	(-0.5,0.4,-0.6)	0.992	0.991	1.000	1.000	1.000	0.999

3.2 A Real Example

Baltagi and Griffin (1983) considered the following gasoline demand equation

$$\ln \frac{\text{Cas}}{\text{Car}} = \alpha + \beta_1 \ln \frac{\text{Y}}{\text{N}} + \beta_1 \ln \frac{\text{P}_{\text{MG}}}{\text{P}_{\text{GDP}}} + \beta_1 \ln \frac{\text{Car}}{\text{N}} + u,$$

where Gas/Car is motor gasoline consumption per auto, Y/N is real per capita income, PMG/PGDP is real motor gasoline price and Car/N denotes the stock of cars per capita. This panel consists of annual observations across 18 OECD countries, covering the period 1960-78. The data for this example can be found in package **plm** of the open source software **R**. Baltagi (2008) (P 244) considered the problem of testing the independence of the first two countries: Austria and Belgium. The observed values of Breusch-Pagan (1980) Lagrange multiplier test statistic and the Likelihood Ratio test statistic for this problem are 0.947 and 1.778, respectively. The observed value of the proposed empirical likelihood test statistic is 2.343. All the three test statistics are distributed as χ_1^2 under the null hypothesis, and do not reject the null hypothesis.

4. Conclusion

This paper proposes a novel approach for the independence test for the disturbances of the SUR models based on the empirical-likelihood method. The proposed test statistic under the null hypothesis is shown to has an asymptotic chi-square distribution. Compared to the Lagrange multiplier test statistic, the simulation experiment demonstrates that the proposed method performs satisfactorily. Furthermore, our approach can be applied to the case that the model errors of one equation of SUR model are correlated.

Appendix: Proof of the main results

We begin with the following assumptions required to derive the main results. These assumptions are quite mild and can be easily satisfied.

Assumption 1. $E(\mathbf{x}_{ik}\varepsilon_{ik}) = \mathbf{0}$ for $1 \le i \le p, 1 \le k \le n$.

Assumption 2. $E(\mathbf{X}_i^{\mathrm{T}}\mathbf{X}_i)$ is nonsingular, $1 \le i \le p$.

In order to prove that main results, we first introduce several lemmas.

Lemma 1 Under the assumptions 1-2 and the null hypothesis, we have

$$\frac{1}{\sqrt{n}}\sum_{k=1}^{n}\boldsymbol{\xi}_k \xrightarrow{D} N(\mathbf{0}, \mathbf{\Omega}),$$

with $\sigma_k^2 = \sigma_{kk}$ and

$$\mathbf{\Omega} = \begin{bmatrix} \sigma_1^2\sigma_2^2 & 0 & \cdots & 0 \\ 0 & \sigma_1^2\sigma_3^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{p-1}^2\sigma_p^2 \end{bmatrix}.$$

Proof: By the result of Tsay (2004), we can obtain Lemma 1.

Lemma 2 Under the assumptions 1-2 and the null hypothesis, we have

$$\frac{1}{n}\sum_{k=1}^{n}\boldsymbol{\xi}_k\boldsymbol{\xi}_k^{\mathrm{T}} \xrightarrow{p} \mathbf{\Omega}.$$

Proof: Firstly, we consider $\frac{1}{n}\sum_{k=1}^{n}\hat{\varepsilon}_{ik}\hat{\varepsilon}_{jk}\hat{\varepsilon}_{sk}\hat{\varepsilon}_{lk}$, one element of $\frac{1}{n}\sum_{k=1}^{n}\boldsymbol{\xi}_k\boldsymbol{\xi}_k^{\mathrm{T}}$. Let $e_{ik} = \mathbf{x}_{ik}^{\mathrm{T}}(\boldsymbol{\beta}_i - \hat{\boldsymbol{\beta}}_i)$, by the definition of $\boldsymbol{\xi}_k$, we have

$$\begin{aligned} \frac{1}{n}\sum_{k=1}^{n}\hat{\varepsilon}_{ik}\hat{\varepsilon}_{jk}\hat{\varepsilon}_{sk}\hat{\varepsilon}_{lk} &= \frac{1}{N}\sum_{k=1}^{n}(e_{ik}+\varepsilon_{ik})(e_{jk}+\varepsilon_{jk})(e_{sk}+\varepsilon_{sk})(e_{lk}+\varepsilon_{lk}) \\ &= \frac{1}{n}\sum_{k=1}^{n}\varepsilon_{ik}\varepsilon_{jk}\varepsilon_{sk}\varepsilon_{lk} + \sum_{i=1}^{15} I_i. \end{aligned}$$

We let $I_1 = \frac{1}{n}\sum_{k=1}^{n}e_{ik}e_{jk}e_{sk}e_{lk}$, By Lemma 3 in Owen (1990), we have

$$\begin{aligned} |I_1| &= \frac{1}{n}\sum_{k=1}^{n}|\mathbf{x}_{ik}^{\mathrm{T}}(\boldsymbol{\beta}_i - \hat{\boldsymbol{\beta}}_i)\mathbf{x}_{jk}^{\mathrm{T}}(\boldsymbol{\beta}_j - \hat{\boldsymbol{\beta}}_j)\mathbf{x}_{sk}^{\mathrm{T}}(\boldsymbol{\beta}_s - \hat{\boldsymbol{\beta}}_s)\mathbf{x}_{lk}^{\mathrm{T}}(\boldsymbol{\beta}_l - \hat{\boldsymbol{\beta}}_l)| \\ &\le \|\boldsymbol{\beta}_i - \hat{\boldsymbol{\beta}}_i\|\|\boldsymbol{\beta}_j - \hat{\boldsymbol{\beta}}_j\|\|\boldsymbol{\beta}_s - \hat{\boldsymbol{\beta}}_s\|\|\boldsymbol{\beta}_l - \hat{\boldsymbol{\beta}}_l\|\frac{1}{n}\sum_{k=1}^{n}\|\mathbf{x}_{ik}\|\|\mathbf{x}_{jk}\|\|\mathbf{x}_{sk}\|\|\mathbf{x}_{lk}\| \\ &\le \|\boldsymbol{\beta}_i - \hat{\boldsymbol{\beta}}_i\|\|\boldsymbol{\beta}_j - \hat{\boldsymbol{\beta}}_j\|\|\boldsymbol{\beta}_s - \hat{\boldsymbol{\beta}}_s\|\|\boldsymbol{\beta}_l - \hat{\boldsymbol{\beta}}_l\| \max_{1\le i\le p, 1\le k\le n}\|\mathbf{x}_{ik}\|^4 \\ &= O_p(n^{-1/2})O_p(n^{-1/2})O_p(n^{-1/2})O_p(n^{-1/2})o_p(n^2) \\ &= o_p(1). \end{aligned}$$

Hence, $I_1 = o_p(1)$. By the similar way, we can prove that $I_i = o_p(1), i = 2, 3, \cdots, 15$. Thus,

$$\frac{1}{n}\sum_{k=1}^{n}\hat{\varepsilon}_{ik}\hat{\varepsilon}_{jk}\hat{\varepsilon}_{sk}\hat{\varepsilon}_{lk} = \frac{1}{n}\sum_{k=1}^{n}\varepsilon_{ik}\varepsilon_{jk}\varepsilon_{sk}\varepsilon_{lk} + o_p(1),$$

and

$$\frac{1}{n}\sum_{k=1}^{n}\boldsymbol{\xi}_k\boldsymbol{\xi}_k^{\mathrm{T}} = \frac{1}{n}\sum_{k=1}^{n}\mathbf{U}_k\mathbf{U}_k^{\mathrm{T}} + o_p(1).$$

Finally, under the null hypothesis, and by the law of large numbers, we have

$$\frac{1}{n}\sum_{k=1}^{n}\boldsymbol{\xi}_k\boldsymbol{\xi}_k^{\mathrm{T}} \xrightarrow{p} \mathbf{\Omega}.$$

Lemma 3 Under the Assumptions 1-2, we have

$$\max_{1\le k\le n} \|\boldsymbol{\xi}_k\| = o_p(n^{1/2}),$$

where $\|\cdot\|$ is the Euclidean norm with $\|\mathbf{a}\| = (a_1^2 + \cdots + a_k^2)^{1/2}$ and $\mathbf{a} = (a_1, \cdots, a_k)^{\mathrm{T}}$.

Proof:

$$\begin{aligned}
\max_{1\le k\le n} |\hat{\varepsilon}_{ik}\hat{\varepsilon}_{jk}| &= \max_{1\le k\le n} \left| \left[\mathbf{x}_{ik}^{\mathrm{T}}(\boldsymbol{\beta}_i - \hat{\boldsymbol{\beta}}_i) + \varepsilon_{ik}\right] \left[\mathbf{x}_{jk}^{\mathrm{T}}(\boldsymbol{\beta}_j - \hat{\boldsymbol{\beta}}_j) + \varepsilon_{jk}\right] \right| \\
&\le \left(\|\boldsymbol{\beta}_i - \hat{\boldsymbol{\beta}}_i\| \max_{1\le k\le n} \|\mathbf{x}_{ik}\|\right) \times \left(\|\boldsymbol{\beta}_j - \hat{\boldsymbol{\beta}}_j\| \max_{1\le k\le n} \|\mathbf{x}_{ik}\|\right) + \max_{1\le k\le n} |\varepsilon_{ik}\varepsilon_{jk}| \\
&\quad + \left(\|\boldsymbol{\beta}_i - \hat{\boldsymbol{\beta}}_i\| \max_{1\le k\le n} \|\mathbf{x}_{ik}\|\right) \max_{1\le k\le n} |\varepsilon_{ik}| + \left(\|\boldsymbol{\beta}_j - \hat{\boldsymbol{\beta}}_j\| \max_{1\le k\le n} \|\mathbf{x}_{ik}\|\right) \max_{1\le k\le n} |\varepsilon_{jk}|.
\end{aligned}$$

By Lemma 3 in Owen (1990), we can prove that

$$\max_{1\le i\le p, 1\le k\le n} \|\mathbf{x}_{ik}\| = o_p(n^{1/2}), \max_{1\le i\le p, 1\le k\le n} |\varepsilon_{ik}| = o_p(n^{1/2}), \max_{1\le k\le n} |\varepsilon_{ik}\varepsilon_{jk}| = o_p(n^{1/2}),$$

Combining $\|\boldsymbol{\beta}_i - \hat{\boldsymbol{\beta}}_i\| = O_p(n^{-1/2})$ and $\|\boldsymbol{\beta}_k - \hat{\boldsymbol{\beta}}_k\| = O_p(n^{-1/2})$, we have

$$\max_{1\le k\le n} |\hat{\varepsilon}_{ik}\hat{\varepsilon}_{jk}| = o_p(n^{1/2}).$$

Therefore, we have

$$\max_{1\le k\le n} \|\boldsymbol{\xi}_k\| = o_p(n^{1/2}).$$

Proof of Theorem 2.1. Using the same strategy as the proof of Theorem 3.2 in Owen (1991), we can prove that

$$\|\boldsymbol{\lambda}\| = O_p(n^{-1/2}). \tag{a.1}$$

It follows from Lemma 3 and (a.1) that

$$\max_{1\le k\le n} |\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\xi}_k| = O_p(n^{-1/2})o_p(n^{1/2}) = o_p(1).$$

Hence, by Taylor's expansion, we have

$$l_n = 2\sum_{k=1}^{n} \log(1 + \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\xi}_k) = 2\sum_{k=1}^{n} \left(\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\xi}_k - \frac{1}{2}(\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\xi}_k)^2\right) + r_n, \tag{a.2}$$

with

$$|r_n| \le C\|\boldsymbol{\lambda}\|^3 \max_k \|\boldsymbol{\xi}_k\| \sum_{k=1}^{n} \|\boldsymbol{\xi}_k\|^2 = o_p(1) \quad .$$

Based on the equation (2.8), by Lemma 3 and (a.1), we have

$$\boldsymbol{\lambda} = \left(\sum_{k=1}^{n} \boldsymbol{\xi}_k\boldsymbol{\xi}_k^{\mathrm{T}}\right)^{-1} \sum_{k=1}^{n} \boldsymbol{\xi}_k + o_p(n^{-1/2}), \tag{a.3}$$

and

$$\sum_{k=1}^{n} \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\xi}_i = \sum_{k=1}^{n} \left(\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\xi}_k\right)^2 + o_p(1). \tag{a.4}$$

By (a.1)-(a.4), we know that

$$\begin{aligned}
l_n &= \sum_{k=1}^{n} \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\xi}_i\boldsymbol{\xi}_k^{\mathrm{T}}\boldsymbol{\lambda} + o_p(1) \\
&= \left(\frac{1}{\sqrt{n}}\sum_{k=1}^{n} \boldsymbol{\xi}_k\right)^{\mathrm{T}} \left(\frac{1}{n}\sum_{k=1}^{n} \boldsymbol{\xi}_k\boldsymbol{\xi}_k^{\mathrm{T}}\right)^{-1} \left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} \boldsymbol{\xi}_k\right) + o_p(1).
\end{aligned}$$

Finally, combining Lemmas 1 and 2, we have $l_n \xrightarrow{D} \chi_N^2$ as $n \to \infty$. The theorem is then proved.

Conflicts of Interest: The author declares no conflict of interest.

References

Baltagi, B. H. (2008). Econometrics (Fourth Edition), Springer, Berlin Heidelberg.

Baltagi, B. H., & Griffin, J. M. (1983). Gasoline Demand in the OECD: An Application of Pooling and Testing Procedures. *European Economic Review, 22*, 117-137. https://doi.org/10.1016/0014-2921(83)90077-6

Breusch, T. S., & Pagan, A. R. (1980). The Lagrange Multiplier Test and Its Applications to Model Specification in Econometrics. *Review of Economic Studies, 47*, 239-253. https://doi.org/10.2307/2297111

Dong, L. B., & Giles, D. A. (2007). An Empirical Likelihood Ratio Test for Normality. *Communications in Statistics-Simulation and Computation, 36*(1), 197-215. https://doi.org/10.1080/03610910601096544

Dufour, J. M., & Khalaf, L. (2002). Exact tests for contemporaneous correlation of disturbances in seemingly unrelated regressions. *Journal of Econometrics, 106*, 143-170. https://doi.org/10.1016/S0304-4076(01)00093-8

Fiebig, D. G. (2001). Seemingly Unrelated Regression, Chapter 5 in Baltagi, B. H. (ed.), A Companion to Theoretical Econometrics (Blackwell: Massachusetts).

Harvey, A. C., & Phillips, G. D. A. (1982). Testing for contemporaneous correlation of disturbances in systems of regression equations. *Bulletin of Economic Research, 34,* 79-81. https://doi.org/10.1111/j.1467-8586.1982.tb00438.x

Liu, F., & Chen, G. M., & Chen, M. (2008). Testing serial correlation in partial linear errors-in-variables models based on empirical likelihood. *Communications in Statistics-Theory and Methods, 37*(12), 1905-1918. https://doi.org/10.1080/03610920801893780

Liu, X. H., & Wang, Z. Z., & Hu, X. M., & Wang, G. F. (2011). Testing serial correlation in partially linear single-index errors-in-variables models. *Communications in Statistics-Theory and Methods, 40,* 2554-2573. https://doi.org/10.1080/03610926.2010.489174

Owen, A. B. (1988). Empirical likelihood ratio confidence intervals for a single functional. *Biometrika, 75*, 237-249. https://doi.org/10.1093/biomet/75.2.237

Owen, A. B. (1990). Empirical likelihood ratio confidence regions. *The Annals of Statistics, 18*, 90-120. https://doi.org/10.1214/aos/1176347494

Owen, A. B. (1991). Empirical likelihood for linear models. *The Annals of Statistics, 19,* 1725-1747. https://doi.org/10.1214/aos/1176348368

Owen, A. B. (2001). Empirical Likelihood. Chapman and Hall, NewYork.

Srivastava, V. K., & Giles, D. E. A. (1987). Seemingly Unrelated Regression Equations Models: Estimation and Inference, Marcel Dekker, New York.

Tsay, W. J. (2004). Testing for contemporaneous correlation of disturbances in seemingly unrelated regressions with serial dependence. *Economics Letters, 83*, 69-76. https://doi.org/10.1016/j.econlet.2003.10.005

Zellner, A. (1962). An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American Statistical Association, 57*, 348-368. https://doi.org/10.1080/01621459.1962.10480664

Copyrights