



---

*Research article*

## Testing for individual and time effects in unbalanced panel data models with time-invariant regressors

Ke Liu and Hanzhong Liu\*

School of Economics and Statistics, Guangzhou University, Guangzhou 510006, China

\* **Correspondence:** Email: From8men@gzhu.edu.cn.

**Abstract:** In this paper, we use a moment-based method to test the existence of the individual and time effects in unbalanced panel data models with time-invariant regressors. Based on the difference of two variance estimators of idiosyncratic errors, three test statistics are proposed. The test statistics for individual (time) effect is robust when the time (individual) effect exists, and is robust for the correlation between explanatory variables and individual or time effect. Additionally, they do not require prior distributional assumptions on the error term. The asymptotic properties of estimators and the test statistics are given in this paper. The Monte Carlo simulations show that the test statistics have good power in finite samples at various situations and a real example is studied for illustration.

**Keywords:** moment-based method; unbalanced; time-invariant; panel data; individual and time effects

---

### 1. Introduction

The panel data model has been widely used in economics and finance, as it can provide more information compared to the time series and cross-section model [1–4]. Also, it can relieve the problem that the important variables are not included in the model by adding individual and time effects to represent the heterogeneity and commonality of units respectively. Therefore, it is very important to test whether the individual and time effects are in the model, see, e.g., Breush and Pagan [5], Honda [6], Baltagi and Li [7], Bera et al. [8] and Baltagi et al. [9], etc.

If the model contains variables that should not present the estimators of this model will be inconsistent and even the statistic inference will be unreliable. For example, when the true model is a linear panel data model with individual effect but a linear panel data model is incorrectly specified, the bias estimate of regression parameters will mislead policy makers. This motivated us to test whether the individual and time effects exist in the model.

Many scholars studied the testing methods of the existence of random effects. Breush and Pagan [5]

proposed the Lagrange multiplier (LM) test under the assumption of normality of error terms, but Honda [6] found that the LM test was also robust under the assumption of non-normality and proposed two one-way tests on the basis of LM test. Baltagi and Li [7] further applied the LM test to the unbalanced panel. Bera et al. [8] used LM to test the individual effect when the idiosyncratic errors have autocorrelation. However, the above models they studied contain only one effect in the error term. For example, when we need to test individual effect, the test results will be distorted if the time effect exists. Baltagi et al. [9] used the LM test method based on the existence of two effects in the model. But LM test requires normality hypothesis and uncorrelation among explanatory variables, random effects and error term. Therefore, when the individual or time effect in the model is related to the explanatory variables, the estimation results will be biased when the ordinary least square (OLS) or general least square (GLS) estimation is used. We know that the individual or time effect is often related to the explanatory variables in the real econometric problem. Wu and Li [10] proposed a test method for individual and time effects based on moment estimation without making assumptions about the distribution of error terms and the independence between random effects and explanatory variables. The main idea is to construct the difference between the two variance estimators of the error terms of the model. One of the variance estimators of the error term is consistent regardless of whether there is an individual or time effect. Another variance estimator of the error term is only consistent under the null hypothesis. Therefore, the difference between them can be used to construct the test statistics, that is to say, the test statistics are close to zero under the null hypothesis and large under the alternative one. Also, the proposed test statistics are robust when testing one effect and the other effect exists.

In real economics applications, when factors such as age and race are considered, time-invariant variables should be included in the regression model. Pesaran [11], Sebastian and Claudia [12] and many other scholars studied the panel data model with time-invariant regressors. On the basis of Wu and Li [10], Chen et al. [13] further proposed the individual and time effect tests for the two-way error panel model with time-invariant regressors, which also combined with Pesaran [11]'s proposed estimation method when the time-invariant variables are endogenous. However, previous literature has not studied the test for an unbalanced panel with time-invariant variables.

Due to the complex external environment, panel data is often unbalanced or incomplete and the missing data may be random or non random. It is necessary to directly study the estimation and statistical inference of the unbalanced panel in order to get a precise conclusion. In the unbalanced panel model, the estimation and test methods for model contained one or two effects and for the different types of missing data are different. For example, on the basis of the testing idea put forward by Wu and Li [10] mentioned above, if the data is monotonically missing (the data of an individual after a certain time point is missing) and it is a linear panel model with only individual effect, the test method proposed by Wu and Li [10] can still be used in the unbalanced panel [14]. However, if the time effect also exists in the model, it is necessary to classify the cross-section individuals with the same missing type into one group. The main reason is that the common time effect can only be removed within the group when using OLS or other estimation methods to obtain accurate estimators. Besides, the interference of individual (time) effects should be avoided during the construction of test statistics [15]. If the data is random missing, the cross-sectional units with the same missing type need to be divided together, that is, the individuals can be differentiated in pairs to eliminate the time effect [16]. To sum up, many scholars have studied the test of parametric models in the unbalanced panel, but the test for individual and time effects in the unbalanced panel has not been studied. Therefore, we use a moment-based

method to test the individual and time effects in the unbalanced panel data model with time-invariant regressors.

The contribution of this paper is the following two aspects. First, considering the difficulty to obtain the whole balanced panel data sample, we study a two-way error component time-invariant regressors model with an unbalanced panel. Also, an effective estimation method is given. Second, referring to Wu and Li [10]'s method, we have built the moment-based test statistics for individual and time effects in the unbalanced panel data model by using the difference of the two variance estimators of the error term. The test for individual (time) effect is robust no matter the existence of time (individual) effect and robust in regards to the correlation between the explanatory variable and individual or time effect. Furthermore, it is unnecessary to make a distributional assumption for the error in advance. The simulations demonstrate that the power of the test statistics is high in various situations.

The rest of this paper is organized as follows. In Section 2, we introduce the basic model and the estimation process. Tests for individual and time effects are shown in Sections 3 and 4 respectively. Section 5 is the jointly test for two effects in the basic model and Section 6 is Monte Carlo simulations and a real example analysis. The conclusion is in Section 7. The proof of the relevant theorems will be presented in Appendix.

## 2. Model and estimation

Consider a two-way error component panel data model with time-invariant regressors,

$$\begin{aligned} y_{it} &= \alpha + X'_{it}\beta + Z'_i\gamma + v_{it}, \\ v_{it} &= \mu_i + \eta_t + \epsilon_{it}, \end{aligned} \quad (2.1)$$

where  $y_{it}$  is the response variable at unit  $i$  and time  $t$ ,  $X_{it} = (X_{it,1}, \dots, X_{it,p})'$  is a  $p \times 1$  vector of time varying explanatory variables,  $X'_{it}$  is the transpose of  $X_{it}$ ,  $\beta$  is a  $p \times 1$  vector of parameter.  $Z_i$  is a  $m \times 1$  time-invariant exogenous variables\* and  $\gamma$  is also a parameter. The error term  $v_{it}$  contains unobservable individual effect  $\mu_i$  and time effect  $\eta_t$ , which are both random effects.  $\mu_i$  is assumed to be independent and identically distributed with mean zero and finite variance  $\sigma_{\mu}^2$ .  $\eta_t$ , like  $\mu_i$ , has mean zero and finite variance  $\sigma_{\eta}^2$ . Also, The idiosyncratic error  $\epsilon_{it}$  is assumed to be identified independent distribution with mean zero and variance  $\sigma_{\epsilon}^2$ .

It is difficult to obtain all the sample data due to several factors including migration. In application, we often get unbalanced panels. However, the traditional methods for balanced panels are not suitable for unbalanced panel data models. We have to propose other effective methods for incomplete data model. It is worth noting that, in unbalanced panel data, if we use normal centering transformation for all units, the time effect will not be removed completely. The reason is that the individuals have different length of time T. Therefore, we can regard each group as a balanced panel data that they also have the same length of time by referring the grouping method of Wu et al. [15]. So, we can solve this problem by centering in each group.

Divide  $n$  cross-section units into  $L$  disjoint groups  $\mathcal{N}_1, \dots, \mathcal{N}_L$  such that the observed time periods are identical for each  $i \in \mathcal{N}_l$  with  $l = 1, 2, \dots, L$ . There are  $n_l$  units with time  $T_l$  in group  $\mathcal{N}_l$ . The whole number of the sample is  $N = \sum_{l=1}^L n_l T_l$ . For each group  $\mathcal{N}_l$ ,

$$y_{li} = \iota_{T_l}\alpha + X_{li}\beta + (\iota_{T_l}Z'_i)\gamma + \iota_{T_l}\mu_i + \eta_l + \epsilon_{li}, l_i \in \mathcal{N}_l, i = 1, \dots, n_l, \quad (2.2)$$

\*We only consider that  $Z_i$  is exogenous for simplicity.

where  $y_{li} = (y_{li,t_{l,1}}, y_{li,t_{l,2}}, \dots, y_{li,t_{l,T_l}})'$ ,  $\iota_{T_l}$  is a vector of one with dimension  $T_l$  and the other right-hand-side variables are stacked accordingly. Next, estimate the unbalanced panel data model.

First, we estimate  $\beta$  by making a centering transformation in each group to remove the time effect,

$$\tilde{y}_{li} = \tilde{X}_{li}\beta + (\iota_{T_l}\tilde{Z}'_i)\gamma + \iota_{T_l}\tilde{\mu}_{li} + \tilde{\epsilon}_i, l_i \in \mathcal{N}_l, i = 1, \dots, n_l, \quad (2.3)$$

where  $\tilde{y}_{li} = y_{li} - \frac{1}{n_l} \sum_{i=1}^{n_l} y_{li}$ .  $\tilde{X}_{li}$ ,  $\tilde{Z}'_i$ ,  $\tilde{\mu}_{li}$  and  $\tilde{\epsilon}_i$  are defined similarly. We can find matrix  $Q_{T_l}$  such that  $(T_l^{-1/2}\iota_{T_l}, Q_{T_l})$  is a  $T_l \times T_l$  orthogonal matrix (e.g., Wu and Li [10]). Denote the  $j^{\text{th}}$  column vector of matrix  $Q_{T_l}$  by  $q_{lj} = (q_{lj,1}, \dots, q_{lj,T_l})'$ ,  $l = 1, \dots, L, j = 1, \dots, T_l - 1$ . Then, multiplying model (2.3) with  $Q'_{T_l}$ , we have

$$Q'_{T_l}\tilde{y}_{li} = Q'_{T_l}\tilde{X}_{li}\beta + Q'_{T_l}\tilde{\epsilon}_i, l_i \in \mathcal{N}_l, i = 1, \dots, n_l, \quad (2.4)$$

where  $Q'_{T_l}\iota_{T_l} = 0$ , so the time-invariant regressors and individual effect are also removed. We can use OLS to obtain the consistent estimator of  $\beta$ ,

$$\hat{\beta} = \arg \min_{\beta} \sum_{l=1}^L \sum_{i=1}^{n_l} \|Q'_{T_l}\tilde{y}_{li} - Q'_{T_l}\tilde{X}_{li}\beta\|^2 = \left( \sum_{l=1}^L \sum_{i=1}^{n_l} \tilde{X}'_{li}P_{T_l}\tilde{X}_{li} \right)^{-1} \left( \sum_{l=1}^L \sum_{i=1}^{n_l} \tilde{X}'_{li}P_{T_l}\tilde{y}_{li} \right), \quad (2.5)$$

where  $P_{T_l} = Q_{T_l}Q'_{T_l} = I_{T_l} - \frac{1}{T_l}\iota_{T_l}\iota'_{T_l}$  that  $I_{T_l}$  is  $T_l \times T_l$  identity matrix and  $Q'_{T_l}Q_{T_l} = I_{T_l-1}$ .

**Theorem 2.1.** Under some mild conditions,  $|\Sigma_1| > 0$  and as  $n \rightarrow \infty$  with fixed time period,  $\hat{\beta}$  has the following asymptotic normal distribution,

$$\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{d} N(0, \Sigma_1^{-1}\Sigma_2\Sigma_1^{-1}) \quad (2.6)$$

where  $\Sigma_1 = \sum_{l=1}^L m_l [E(X'_{li}P_{T_l}X_{li}) - EX'_{li}P_{T_l}EX_{li}]$ ,  $\Sigma_2 = \sum_{l=1}^L m_l E[(X_{li} - EX_{li})'P_{T_l}\epsilon_{li}\epsilon'_{li}(X_{li} - EX_{li})]$  and  $m_l = \lim_{n \rightarrow \infty} \frac{n_l}{n}$ , see Shao et al. [25]. Note that the result is based on the large individuals and short time period.

The proof of the asymptotic distribution of  $\hat{\beta}$  can refer Wu et al. [10] since the time-invariant variables are removed after the transformation.

Second, estimate  $\gamma$ . Peseran and Zhou [11] used the filtered method to obtain the estimator of  $\gamma$  by the regression of  $\hat{u}_{it}$  on  $Z_i$  and mentioned that the estimator is still consistent under unbalanced panel data model. Details of estimating procedure are as follow. For each group  $\mathcal{N}_l$ , the average over time of  $\hat{u}_{it}$ ,

$$\tilde{u}_{li} = \alpha + Z'_i\gamma + \bar{v}_{li}, i = 1, \dots, n_l, \quad (2.7)$$

where  $\tilde{u}_{li} = \frac{1}{T_l} \sum_{t=1}^{T_l} \hat{u}_{it}$ ,  $\bar{X}_{li} = \frac{1}{T_l} \sum_{t=1}^{T_l} X_{it}$ ,  $\bar{v}_{li} = \frac{1}{T_l} \sum_{t=1}^{T_l} \epsilon_{it}$  and  $\hat{u}_{it}$  can be estimated by  $y_{it} - X'_{it}\hat{\beta}$ . Next, we can make the centering transformation over cross-section units in each group. Thus, the estimator of  $\gamma$  is

$$\hat{\gamma} = \left[ \sum_{l=1}^L \sum_{i=1}^{n_l} (Z_{li} - \bar{Z})(Z_{li} - \bar{Z})' \right]^{-1} \sum_{l=1}^L \sum_{i=1}^{n_l} (Z_{li} - \bar{Z})(\tilde{u}_{li} - \bar{\tilde{u}}), \quad (2.8)$$

where  $\bar{Z} = \frac{1}{n_l} \sum_{i=1}^{n_l} Z_{li}$ ,  $\bar{\tilde{u}} = \frac{1}{n_l} \sum_{i=1}^{n_l} \tilde{u}_{li}$ . For the identifying assumption of  $\gamma$ , the time-invariant variables are not correlated to the individual effect in this paper and the endogenous regressors can be considered in the future study.

**Assumption 2.1.** The error term  $\epsilon_{it}$  is cross-sectional uncorrelated and unrelated with explanatory variables  $X_{it}$  for  $i = 1, \dots, n_t$  and  $t = 1, \dots, T$ ,  $E(\epsilon_{it}\epsilon_{js}|X, Z) = 0$  for  $t, s$  and  $i \neq j$ ,  $E(\epsilon_{it}X_{it,p}) = 0$  for  $i, j$  and  $p$ .

**Assumption 2.2.** The  $\epsilon_{it}$  can be heterogeneous,  $E(\epsilon_{it}\epsilon_{is}) = r_i(t, s)$ , where  $r_i(t, t) = r_i^2$  and  $r_i(t, s) < K$ ,  $K$  is the nonzero constant and  $K < \infty$ .

**Assumption 2.3.** The time-invariant variable  $Z_{it}$  is uncorrelated to the individual effect  $\mu_{it}$  and the error term  $\bar{\epsilon}_{it}$  and also  $\mu_{it}$  and  $\bar{\epsilon}_{it}$  are independently distributed. The fourth moment of  $Z_{it}$  is finite.  $E(Z_{it}\mu_{it}) = 0$ ,  $E(Z_{it}\bar{\epsilon}_{it}) = 0$ ,  $E(\|Z_{it}\|^2\mu_{it}^2) = 0$ ,  $E(\|Z_{it}\|^2\bar{\epsilon}_{it}^2) = 0$ ,  $E(Z_{it}\mu_{it}) = 0$  and  $E(\|Z_{it}\|^2\mu_{it}\bar{\epsilon}_{it}) = 0$  for all  $i, j$ .

These assumptions are similar to Peseran and Zhou [11].

**Theorem 2.2.** Under the Assumptions 2.1–2.3 and suppose that  $|\Sigma_3| > 0$ , as  $n \rightarrow \infty$  with fixed time period, the estimator  $\gamma$  has the following asymptotic distribution,

$$\sqrt{n}(\hat{\gamma} - \gamma) \xrightarrow{d} N(0, \Sigma_3^{-1}(\Sigma_4 + \Omega_1 + \Omega_2)\Sigma_3^{-1}) \quad (2.9)$$

where the definition of  $\Sigma_3$ ,  $\Sigma_4$ ,  $\Omega_1$  and  $\Omega_2$  can be seen in Eqs (A.8), (A.10), (A.14) and (A.15) in the Appendix.

### 3. Test for individual effect

In this section, we construct a test statistic for individual effect in model (2.2). The hypothesis is:

$$H_0^\mu : \sigma_\mu^2 = 0, H_1^\mu : \sigma_\mu^2 > 0.$$

After the transformation of original model, we obtain model (2.4) and the moment condition of  $Q'_{T_l}\tilde{\epsilon}_{it}$ :

$$\sum_{l=1}^L \sum_{i=1}^{n_l} E\|Q'_{T_l}\tilde{\epsilon}_{it}\|^2 = \sum_{l=1}^L \sum_{i=1}^{n_l} \sum_{j=1}^{T_l-1} E(q'_{lj}\tilde{\epsilon}_{it})^2 = c_1\sigma_\epsilon^2, \quad (3.1)$$

where  $c_1 = \sum_{l=1}^L c_{1l}$  with  $c_{1l} = (n_l - 1)(T_l - 1)$ . Then, the estimator of  $\sigma_\epsilon^2$  is:

$$\hat{\sigma}_{0\epsilon}^2 = \frac{1}{c_1} \sum_{l=1}^L \sum_{i=1}^{n_l} \|Q'_{T_l}(\tilde{y}_{it} - \tilde{X}_{it}\hat{\beta})\|^2 = c_1^{-1} \sum_{l=1}^L \sum_{i=1}^{n_l} (\tilde{y}_{it} - \tilde{X}_{it}\hat{\beta})' P_{T_l}(\tilde{y}_{it} - \tilde{X}_{it}\hat{\beta}), \quad (3.2)$$

where the  $\hat{\beta}$  is obtained from Eq (2.5).  $\hat{\sigma}_{0\epsilon}^2$  is a consistent estimator of  $\sigma_{0\epsilon}^2$  under the null hypothesis, alternative hypothesis, and regardless of the existence of  $\mu_{it}$  and  $\eta_{it}$  (see Wu et al. [10]).

We can construct the fourth-order moment of the variance of error  $\sigma_\epsilon$  by referring Wu and Zhu [17] and Wu et al. [15].

$$\sum_{l=1}^L \sum_{i=1}^{n_l} \sum_{j=1}^{T_l-1} E(q'_{lj}\tilde{\epsilon}_{it})^4 = c_2\lambda_\epsilon^4 + c_2c_3(\sigma_\epsilon^2)^2 \quad (3.3)$$

where  $\lambda_\epsilon^4 = E\epsilon_{it}^4$ .

$$c_2 = \sum_{l=1}^L \sum_{j=1}^{T_l-1} \sum_{t=1}^{T_l} q_{lj,t}^4 (n_l - 1)(n_l^2 - 3n_l + 3)/n_l^2, \quad (3.4)$$

$$c_3 = c_2^{-1} \sum_{l=1}^L 3(n_l - 1)^2(T_l - 1)/n_l - 3. \quad (3.5)$$

Thus, the estimator of the  $\lambda_\epsilon^4$ ,

$$\hat{\lambda}_\epsilon^4 = c_2^{-1} \sum_{l=1}^L \sum_{i=1}^{n_l} \sum_{j=1}^{T_l-1} E(q'_{lj} \tilde{\epsilon}_{li})^4 - c_3(\hat{\sigma}_{0\epsilon}^2)^2 = c_2^{-1} \sum_{l=1}^L \sum_{i=1}^{n_l} \sum_{j=1}^{T_l-1} (q'_{lj}(\tilde{y}_{li} - \tilde{X}_{li}\hat{\beta}))^4 - c_3(\hat{\sigma}_{0\epsilon}^2)^2 \quad (3.6)$$

is consistent under some mild conditions.

Next, we should obtain the consistent estimator of  $\beta$  under the null hypothesis. Under the null hypothesis, the original model (2.2) becomes

$$y_{li} = l_{T_l}\alpha + X_{li}\beta + (l_{T_l}\tilde{Z}'_i)\gamma + \eta_l + \epsilon_{li}, l_i \in \mathcal{N}_l, i = 1, \dots, n_l. \quad (3.7)$$

We just need to remove the time effect by centering in each group,

$$\tilde{y}_{li} = \tilde{X}_{li}\beta + (l_{T_l}\tilde{Z}'_i)\gamma + \tilde{\epsilon}_{li}, l_i \in \mathcal{N}_l, i = 1, \dots, n_l. \quad (3.8)$$

It holds that

$$\sum_{l=1}^L \sum_{i=1}^{n_l} E\|\tilde{\epsilon}_{li}\|^2 = c_2\sigma_\epsilon^2. \quad (3.9)$$

where  $c_4 = \sum_{l=1}^L c_{4l}$  with  $c_{4l} = (n_l - 1)T_l$ . Similarly,

$$\hat{\sigma}_{1\epsilon}^2 = \frac{1}{c_4} \sum_{l=1}^L \sum_{i=1}^{n_l} \|(\tilde{y}_{li} - \tilde{X}_{li}\hat{\beta} - (l_{T_l}\tilde{Z}'_i)\hat{\gamma})\|^2, \quad (3.10)$$

where the  $\hat{\gamma}$  is obtained in Eq (2.8). However,  $\hat{\sigma}_{1\epsilon}^2$  is only consistent under  $H_1^\mu$  (see Wu et al. [10]) while  $\hat{\sigma}_{0\epsilon}^2$  is both consistent under  $H_0^\mu$  and  $H_1^\mu$ . Therefore, we can construct a test statistic by following Hausman [18],

$$T_\mu = \omega_n^{-1/2} \sqrt{n}(\hat{\sigma}_{1\epsilon}^2 - \hat{\sigma}_{0\epsilon}^2) \quad (3.11)$$

where  $\omega_n = a_n\hat{\lambda}_\epsilon^4 + b_n(\hat{\sigma}_{0\epsilon}^2)^2$  is used to standardize the test statistic. The

$$a_n = \frac{1}{n} \sum_{l=1}^L n_l \left( \frac{n^2}{c_4^2} T_l + \frac{n^2}{c_1^2} (T_l + \frac{1}{T_l} - 2) - \frac{2n^2}{c_1 c_4} (T_l - 1) \right), \quad (3.12)$$

$$b_n = \frac{1}{n} \sum_{l=1}^L n_l \left( \frac{n^2}{c_4^2} T_l(T_l - 1) + \frac{n^2}{c_1^2} (T_l - 1)(T_l + \frac{3}{T_l} - 2) - \frac{2n^2}{c_1 c_4} (T_l - 1)^2 \right). \quad (3.13)$$

That  $T_\mu$  will be close to zero under the  $H_0^\mu$  but large under  $H_1^\mu$ .

**Assumption 3.1.** The individual effect  $\mu_i = n^{-\frac{1}{4}}\mu_{0i}$ , where  $\mu_{0i}$  is i.i.d. with zero mean and finite variance  $\sigma_1^2 > 0$ .

**Assumption 3.2.**  $E(\mu_i\epsilon_{it}) = 0$ ,  $n^{\frac{1}{2}}E(\mu_i^2\epsilon_{it}^2) < \infty$ ,  $E(n^{\frac{1}{4}}X_{it,p}\mu_i) < \infty$  for each  $i$ ,  $t$  and  $p$ .

The assumptions can refer to Wu et al. [15].

**Theorem 3.1.** *Suppose the explanatory variable  $X_{it}$ , time-invariant variable  $Z_{it}$  and error term  $\epsilon_{it}$  are i.i.d sequence with  $EX_{it}^4 < \infty$ ,  $|\Sigma_1| > 0$ ,  $EZ_{it}^4 < \infty$ ,  $|\Sigma_3| > 0$  and  $E\epsilon_{it}^5 < \infty$ . Under the Assumptions 3.1 and 3.2, we have the following asymptotic distribution,*

$$T_\mu - \Phi_n^{-1/2} \sigma_1^2 \rightarrow_d N(0, 1) \quad (3.14)$$

where  $\Phi_n = a_n \lambda_\epsilon^4 + b_n (\sigma_{0\epsilon}^2)^2$ ,  $a = \lim_{n \rightarrow \infty} a_n$ ,  $b = \lim_{n \rightarrow \infty} b_n$ .

The proof of the above theorem is in Appendix.

#### 4. Test for time effect

In this section, we will test the time effect of model (2.2). Consider the heteroscedasticity of  $\eta$ , the hypothesis is as follows,

$$H_0^\eta : \text{var}(\eta_1) = \dots = \text{var}(\eta_T) = 0, \quad H_1^\eta : \text{at least one of them is nonzero.}$$

where  $T$  is the largest number of time of all groups. Similar to the test for individual, under  $H_0^\eta$ , the model (2.2) reduces to,

$$y_{li} = l_{Tl} \alpha + X_{li} \beta + (\iota_{Tl} Z'_i) \gamma + \iota_{Tl} \mu_{li} \epsilon_{li}, \quad l_i \in \mathcal{N}_l, i = 1, \dots, n_l. \quad (4.1)$$

We obtain the variance estimator of error by eliminating the individual effect and the time-invariant variables by the same orthogonal transformation mentioned in the Section 2. We have

$$Q'_{Tl} y_{li} = Q'_{Tl} X_{li} \beta + Q'_{Tl} \epsilon_{li}, \quad l_i \in \mathcal{N}_l, i = 1, \dots, n_l. \quad (4.2)$$

It also holds that

$$\sum_{l=1}^L \sum_{i=1}^{n_l} E \|Q'_{Tl} \epsilon_{li}\|^2 = c_5 \sigma_\epsilon^2, \quad (4.3)$$

where  $c_5 = \sum_{l=1}^L c_{5l}$  with  $c_{5l} = (T_l - 1)n_l$ . So,

$$\hat{\sigma}_{2\epsilon}^2 = \frac{1}{c_5} \sum_{l=1}^L \sum_{i=1}^{n_l} \|Q'_{Tl} (y_{li} - X_{li} \hat{\beta})\|^2 = c_5^{-1} \sum_{l=1}^L \sum_{i=1}^{n_l} (y_{li} - X_{li} \hat{\beta})' P_{Tl} (y_{li} - X_{li} \hat{\beta}). \quad (4.4)$$

That  $\hat{\sigma}_{2\epsilon}^2$  is consistent under the null hypothesis  $H_0^\eta$  but inconsistent under  $H_1^\eta$  (see Wu et al. [10]). The difference between  $\hat{\sigma}_{2\epsilon}^2$  and  $\hat{\sigma}_{0\epsilon}^2$  should be small under  $H_0^\eta$ . A test statistic for time effect is

$$T_\eta = \frac{c_5}{\hat{\sigma}_{0\epsilon}^2} (\hat{\sigma}_{1\epsilon}^2 - \hat{\sigma}_{0\epsilon}^2) + \sum_{l=1}^L (T_l - 1). \quad (4.5)$$

Then, we set some assumptions to study the properties of  $T_\eta$  (refer to Wu et al. [15]).

**Assumption 4.1.** *The time effect  $\eta_t = n^{-1/2} \eta_{0t}$ , where  $\eta_{0t}$  is random variable with mean zero and finite variance  $E\eta_{0t}^2 > 0$ .*

**Theorem 4.1.** Suppose that  $EX_{it,p}^2 < \infty$ ,  $E\epsilon_{it}^2 < \infty$ ,  $|\Sigma_1| > 0$  and  $E(X_{it,p}\epsilon_{is}^2)$  for  $i, t, p, s$ . If the Assumption 4.1 holds, we have

$$T_\eta \longrightarrow \sum_{l=1}^L \|\sigma_\epsilon Q'_{T_l} \eta_l + N(0, I_{T_l-1} - \sigma_\epsilon^{-2} \Sigma_{8l})\|^2 \quad (4.6)$$

where  $\eta = (\eta_{0t_{1,1}}, \eta_{0t_{1,2}}, \dots, \eta_{0t_{1,T_1}})$ ,  $\Sigma_{8l} = \Sigma_{9l} + \Sigma'_{9l} - \Sigma_{10l}$  with  $\Sigma_{9l} = m_l Q'_{T_l} E(\epsilon_{it} \epsilon'_{it} P_{T_l}(X_{it} - EX_{it})) \Sigma_1^{-1} EX'_{it} Q_{T_l}$  and  $\Sigma_{10l} = m_l Q'_{T_l} EX_{it} \Sigma_1^{-1} \Sigma_2 \Sigma_1^{-1} \times EX'_{it} Q_{T_l}$ .

See Wu et al. [15] for similar discussion and the proof is omitted in this paper.

We can obtain that  $\Sigma_{8l} = 0$  if the  $EX_{it}$  is independent over time period. Then, the time effect test statistic  $T_\eta$  follows chi-distribution under the null hypothesis, which is very efficient in real application. In order to satisfy this condition, we consider two transformations as same as Wu et al. [15]. Firstly, centralize the  $X_{it}$  in the original model by deducing  $\frac{1}{n_l} \sum_{i=1}^{n_l} X_{it}$  in each group. Secondly, transform the explained variable  $y_{l,t}$  into  $y_{l,t} - X_{l,t} \hat{\beta}$ . Theorem 4.1 still holds after the above transformation.

## 5. Test jointly for two effects

In this section, we consider the jointly test for the two effects of model (2.2). The hypothesis is as follows,

$$H_0^{\mu\eta} : \sigma_\mu^2 = \text{var}(\eta_1) = \dots = \text{var}(\eta_T) = 0, H_1^{\mu\eta} : \text{at least one of them is nonzero.}$$

Under the null hypothesis  $H_0^{\mu\eta}$ , the original model becomes

$$y_{li} = \iota_{T_l} \alpha + X_{li} \beta + (\iota_{T_l} Z'_i) \gamma + \epsilon_{li}, l_i \in \mathcal{N}_l, i = 1, \dots, n_l. \quad (5.1)$$

We have

$$\sum_{l=1}^L \sum_{i=1}^{n_l} E\|\epsilon_{li}\|^2 = \sum_{l=1}^L n_l T_l \sigma_\epsilon^2 = N \sigma_\epsilon^2. \quad (5.2)$$

The estimator of the variance of the error term is

$$\hat{\sigma}_{3\epsilon}^2 = \frac{1}{N} \sum_{l=1}^L \sum_{i=1}^{n_l} \|y_{li} - \iota_{T_l} \hat{\alpha} - X_{li} \hat{\beta} - (\iota_{T_l} Z'_i) \hat{\gamma}\|^2, \quad (5.3)$$

where  $\hat{\beta}$  and  $\hat{\gamma}$  are obtained from (2.5) and (2.8). Referring to Pesaran and Zhou [11], the consistent estimator of  $\alpha$  is that  $\hat{\alpha} = \frac{1}{N} \sum_{l=1}^L \sum_{i=1}^{n_l} \iota'_{T_l} (y_{li} - X_{li} \hat{\beta} - (\iota_{T_l} Z'_i) \hat{\gamma})$ .  $\hat{\sigma}_{3\epsilon}^2$  is only consistent under  $H_0^{\mu\eta}$ . The test statistic can be similarly constructed based on the difference between  $\hat{\sigma}_{3\epsilon}^2$  and  $\hat{\sigma}_{0\epsilon}^2$ ,

$$T_{\eta\mu} = \omega_n^{-1/2} \sqrt{n} (\hat{\sigma}_{3\epsilon}^2 - \hat{\sigma}_{0\epsilon}^2) \quad (5.4)$$

where  $\omega_n$  is as same as (3.11).

**Theorem 5.1.** Under the Assumptions 3.1, 3.2 and 4.1 mentioned above,  $EX_{it}^4 < \infty$ ,  $|\Sigma_1| > 0$ ,  $EZ_{it}^4 < \infty$ ,  $|\Sigma_3| > 0$  and  $E\epsilon_{it}^5 < \infty$ , we have the following asymptotic distribution,

$$T_{\eta\mu} - \Phi_n^{-1/2} \sigma_1^2 \rightarrow_d N(0, 1) \quad (5.5)$$

where  $\Phi_n = a_n \lambda_\epsilon^4 + b_n (\sigma_{0\epsilon}^2)^2$  and  $\sigma_1$  is defined in Assumption 3.1.

The proof of the theorem is similar to Theorem 3.1 and omitted in this paper since we can also refer to Wu et al. [15].



## 6. Simulations

In order to evaluate the performance of the test method, Monte Carlo simulations are used to compute the empirical power of the test statistics based on 1000 replications.

### 6.1. Monte Carlo

We consider the data generating process as follows,

$$y_{it} = \alpha + X'_{it}\beta + Z'_i\gamma + \mu_i + \eta_t + \epsilon_{it}, i = 1, \dots, n, t = 1, \dots, T_i, \quad (6.1)$$

where the explanatory variables and time-invariant variables are two-dimensional,  $X_{it} = (X_{it,1}, X_{it,2})'$ ,  $Z_i = (Z_{i,1}, Z_{i,2})'$ ,  $\beta = (\beta_1, \beta_2)'$  with  $\beta_1 = 1$  and  $\beta_2 = 1$ ,  $\gamma = (\gamma_1, \gamma_2)'$  with  $\gamma_1 = 1$  and  $\gamma_2 = 1$ ,  $\{\mu_i\}$  and  $\{\eta_t\}$  are i.i.d. with mean zero, finite variance  $\text{var}(\mu_i) = \sigma_\mu^2$  and  $\text{var}(\eta_t) = \sigma_\eta^2$ . In this unbalanced panel data model, we assume that it has three groups with different time periods  $T_i$ , which is  $T = [4, 8, 12]$ . The number of cross-section units in each group is random (see Wu et al. [15]). Referring to Chen et al. [13], we also set time-varying variables generated by

$$X_{it,1} = k\mu_i g_{1t} + 3h\eta_t + w_{it}^{(1)}, X_{it,2} = k\mu_i g_{2t} + h\eta_t + w_{it}^{(2)}, \quad (6.2)$$

where  $k$  and  $h$  are constant and we can use them to control the correlations between  $X_{it}$  and individual effect  $\mu_i$  and time effect  $\eta$  respectively. The correlation form has no effect on the results. For simplicity, we only set this form.  $g_{1t}$  and  $g_{2t}$  are uniformly distributed  $U(1, 2)$ ,  $w_{it}^{(1)}$  and  $w_{it}^{(2)}$  are i.i.d. normal distribution with mean zero and one variance. In this paper, we assume that the time-invariant variable  $Z_i$  is exogenous.

$$Z_{i,1} = 1 + \bar{w}_{i1} + \epsilon_i^{(1)}, Z_{i,2} = 1 + \bar{w}_{i2} + \epsilon_i^{(2)}, \quad (6.3)$$

where  $\bar{w}_{i1} = T^{-1} \sum_{t=1}^T w_{it}^{(1)}$ ,  $\bar{w}_{i2} = T^{-1} \sum_{t=1}^T w_{it}^{(2)}$ ,  $\epsilon_i^{(1)}$  and  $\epsilon_i^{(2)}$  are i.i.d.  $N(0, 1)$ . Next, we describe the estimators of two parameters and the power of the test statistics in different situations. And we choose  $n = 50, 100, 150, 200$  to evaluate the performance of the test statistics as the cross-section units increase. First, we summarize the performance of the estimators in Table 1 when the individual and time effects exist simultaneously. The value of the  $\hat{\beta}_1$  and  $\hat{\beta}_2$  ( $\hat{\gamma}_1$  and  $\hat{\gamma}_2$ ) is very close, so we just show the result of  $\hat{\beta}_1$  and  $\hat{\gamma}_1$ . Both  $\hat{\gamma}$  and  $\hat{\beta}$  have small deviations and variances. The standard deviations of the two estimators also decrease when the sample size increases. We can see that the estimator is very close to the true value.

Second, we evaluate the empirical power of individual test statistic  $T_\mu$  in two different distributions of error term. They are  $\epsilon_{it} \sim N(0, 1)$  and  $\epsilon_{it} \sim \sqrt{\frac{1}{2}} \chi^2(1)$ . Table 2 shows that power of the  $T_\mu$  is still high even the time effect  $\eta$  exists and we omit the result here because the results are same as that when the time effect does not exist. The individual effect test statistic is robust under two types of distribution and also robust to the correlation between the explanatory variables and individual effect. Besides, the empirical power of  $T_\mu$  increases with the increase of sample size.

**Table 1.** Finite sample performance of the estimators of  $\hat{\beta}_1$  and  $\hat{\gamma}_1$  in case of  $k = 0, k = 1$  and  $h = 0$  with  $\sigma_\eta = 0.2$ . (The values in the parentheses are the standard deviation).

	$n$	$\sigma_\mu = 0.2$		$\sigma_\mu = 0.4$	
		k=0	k=1	k=0	k=1
$\hat{\beta}_1$	50	1.0037 (0.0748)	1.0039 (0.0746)	1.0037 (0.0748)	1.0041 (0.0748)
	100	0.9981 (0.0528)	0.9981 (0.0529)	0.9980 (0.0529)	0.9981 (0.0530)
	150	1.0001 (0.0436)	1.0003 (0.0433)	1.0001 (0.0436)	0.9999 (0.0437)
	200	1.0022 (0.0387)	1.0023 (0.0388)	1.0020 (0.0361)	1.0024 (0.0400)
$\hat{\gamma}_1$	50	0.9988 (0.0038)	0.9986 (0.0040)	0.9986 (0.0062)	1.1895 (0.2510)
	100	1.0012 (0.0018)	1.0014 (0.0017)	1.0016 (0.0029)	1.1120 (0.0030)
	150	0.9988 (0.0011)	0.9987 (0.0012)	0.9981 (0.0018)	0.9983 (0.0020)
	200	0.9993 (0.0009)	0.9994 (0.0008)	0.9995 (0.0014)	0.9996 (0.0015)

**Table 2.** Empirical powers of the tests  $T_\mu$  in case of  $k = 0, k = 1$  and  $h = 0$  with  $\sigma_\eta = 0$ .

$\varepsilon_{it} \sim N(0, 1)$	$n$	$\sigma_\mu$	$T_\mu$		$\varepsilon_{it} \sim \sqrt{\frac{1}{2}}\chi^2(1)$	$n$	$\sigma_\mu$	$T_\mu$	
			k=0	k=1				k=0	k=1
	50	0.1	0.0455	0.0500		50	0.1	0.0555	0.0550
		0.2	0.2410	0.2590			0.2	0.2450	0.2605
		0.3	0.6760	0.6355			0.3	0.6600	0.6215
		0.4	0.9425	0.8805			0.4	0.9195	0.8570
	100	0.1	0.0810	0.0855		100	0.1	0.0695	0.0710
		0.2	0.4475	0.4535			0.2	0.4110	0.4090
		0.3	0.9255	0.8790			0.3	0.8910	0.8410
		0.4	0.9990	0.9860			0.4	0.9857	0.9854
	150	0.1	0.0830	0.0850		150	0.1	0.1960	0.2090
		0.2	0.6010	0.5870			0.2	0.4140	0.4530
		0.3	0.9860	0.9625			0.3	0.9810	0.9470
		0.4	1.0000	0.9980			0.4	0.9995	0.9970
	200	0.1	0.1095	0.1085		200	0.1	0.0970	0.1050
		0.2	0.7190	0.7065			0.2	0.6790	0.6635
		0.3	0.9995	0.9870			0.3	0.9960	0.9825
		0.4	1.0000	1.0000			0.4	1.0000	1.0000

Next, we evaluate the time effect test statistic  $T_\eta$ . Table 3 shows that the power of  $T_\eta$  is higher than the individual effect test statistic  $T_\mu$ . When the individual effect  $\mu > 0$ , the power of the  $T_\eta$  is as same as the result of  $T_\mu > 0$  and we omit here. Besides, the power of the  $T_\eta$  increases as the sample size increases. Also,  $T_\eta$  is robust to the correlation between the explanatory variables and time effect.

Table 4 displays the empirical power of the  $T_{\mu\eta}$  when the explanatory variables are not related to the time effect ( $h = 0$ ). The test statistics becomes larger as the sample size increases. According to the Tables 2 and 3, the power of  $T_\eta$  is higher than  $T_\mu$ . So the value of  $\sigma_\eta$  will have a greater impact on  $T_{\mu\eta}$  than  $\sigma_\mu$ . When  $\sigma_\eta$  is large, such as  $\sigma_\eta = 0.4$ , the power of  $T_{\mu\eta}$  is high at different values of  $\sigma_\mu$ . When the sample size is small like  $n = 50, 100$ , the power is also very high. Thus, we just simulate  $n = 50$  and 100.

Table 5 displays the empirical power of the  $T_{\mu\eta}$  when the explanatory variables are not related to the individual effect ( $k = 0$ ).  $T_{\mu\eta}$  is robust under  $h = 0$  and  $h = 1$ . When  $\sigma_\mu = \sigma_\eta = 0$ ,  $T_{\mu\eta}$  is as same as Table 4 no matter  $h = 0$  or  $h = 1$  and  $k = 1$  or  $k = 1$  and  $T_{\mu\eta}$  just increases with the sample data increases. So, we omit this case here.

To sum up, three test statistics proposed in this paper are robust and have good power under several unique situations. Furthermore, many studies verified that these types of moment-based test methods have higher power than LM tests proposed by Breusch and Pagan, [5], Honda's test [6] and conditional LM tests proposed by Baltagi et al. [9], see, e.g., Wu and Li [10], Wu et al. [15], Chen et al. [13], etc, for more details.

**Table 3.** Empirical powers of the tests  $T_\eta$  in case of  $h = 0$ ,  $h = 1$  and  $k = 0$  with  $\sigma_\mu = 0$ .

n	$\sigma_\eta$	$T_\eta$	
		h=0	h=1
50	0.1	0.1575	0.1730
	0.2	0.5565	0.5405
	0.3	0.8955	0.8180
	0.4	0.9890	0.9250
100	0.1	0.2955	0.2950
	0.2	0.8605	0.8325
	0.3	0.9980	0.9730
	0.4	1.000	0.9950
150	0.1	0.4095	0.4010
	0.2	0.9675	0.9405
	0.3	1.0000	0.9965
	0.4	1.0000	0.9995
200	0.1	0.5345	0.5405
	0.2	0.9865	0.978
	0.3	1.0000	0.9785
	0.4	1.0000	1.0000

**Table 4.** Empirical powers of the tests  $T_{\mu\eta}$  in case of  $k = 1$  and  $k = 0$  with  $h = 0$ .

n	$\sigma_\eta$	$\sigma_\mu$	$T_{\mu\eta}$		n	$\sigma_\eta$	$\sigma_\mu$	$T_{\mu\eta}$	
			k=0	k=1				k=0	k=1
50	0.0	0.0	0.0400	0.0400	100	0.0	0.0	0.0310	0.0310
		0.1	0.0684	0.0685			0.1	0.0865	0.0855
		0.2	0.2345	0.2480			0.2	0.4320	0.4380
		0.3	0.6430	0.6130			0.3	0.9095	0.8650
	0.2	0.4	0.9240	0.8600		0.4	0.9975	0.9830	
		0.0	0.2590	0.2590		0.2	0.0	0.4400	0.4400
		0.1	0.3305	0.3335		0.1	0.5565	0.5595	
		0.2	0.5855	0.5770		0.2	0.8640	0.8575	
	0.4	0.3	0.8575	0.8300		0.3	0.9875	0.9765	
		0.4	0.9810	0.9480		0.4	0.9995	0.9975	
		0	0.9165	0.9165		0.4	0	0.9850	0.9850
		0.1	0.9370	0.9380		0.1	0.9925	0.9910	
	0.2	0.9755	0.9700	0.2	0.9985	0.9985			
	0.3	0.9935	0.9890	0.3	1.0000	1.0000			
	0.4	0.9990	0.9965	0.4	1.0000	1.0000			

**Table 5.** Empirical powers of the tests  $T_{\mu\eta}$  in case of  $h = 1$  and  $h = 0$  with  $k = 0$ .

n	$\sigma_\mu$	$\sigma_\eta$	$T_{\mu\eta}$		n	$\sigma_\mu$	$\sigma_\eta$	$T_{\mu\eta}$	
			h=0	h=1				h=0	h=1
50	0.0	0.1	0.0805	0.0860	100	0.0	0.1	0.0815	0.0915
		0.2	0.2590	0.2965			0.2	0.4390	0.4385
		0.3	0.6370	0.6075			0.3	0.8675	0.8200
		0.4	0.9165	0.8090			0.4	0.9850	0.9640
	0.2	0.1	0.3270	0.3345		0.2	0.1	0.5660	0.5695
		0.2	0.5855	0.5740		0.2	0.2	0.8640	0.8465
		0.3	0.8565	0.8000		0.3	0.3	0.9800	0.9645
		0.4	0.9755	0.9150		0.4	0.4	0.9985	0.9920
	0.4	0.1	0.9445	0.9455		0.4	0.1	0.9990	0.9990
		0.2	0.9810	0.9765		0.2	0.2	0.9995	0.9995
		0.3	0.9965	0.9905		0.3	0.3	1.0000	0.9995
		0.4	0.9990	0.9950		0.4	0.4	1.0000	1.0000

## 6.2. A real example

Economic development is of great importance to countries and regions. There are many factors affecting the economic growth, see, e.g., Oleg [19], Iuliana [20], Carlsen [21], etc. The impact of foreign direct investment (FDI) on economic growth has been one of the focuses in economics. It can provide some useful suggestions for policymakers. Richard et al. [22] studied nine OECD countries

and seven industries by using the cross-section data and concluded that FDI promotes economic growth through technology spillover. Borenztein et al. [23] studied 69 developing countries and found that FDI can only promote economic growth of host countries when advanced technology has sufficient absorptive capacity, and this impact depends on human capital. Besides, there are many other factors have effects on economic growth.

In order to study the relationship between FDI and economic growth, Kottardi and Stengos [24] considered the traditional linear model,

$$y_{it} = \beta_1(\text{FDI}/Y)_{it} + \beta_2 \ln(\text{DI}/Y)_{it} + \beta_3 n_{it} + \beta_4 h_{it} + \epsilon_{it}. \quad (6.4)$$

where  $y_{it}$  is per capita GDP growth rate in  $i^{\text{th}}$  province and  $t^{\text{th}}$  time point,  $(\text{FDI}/Y)_{it}$  is the ratio of foreign direct investment to total output,  $(\text{DI}/Y)_{it}$  is the ratio of local investment to total output,  $n_{it}$  is natural population growth rate,  $h_{it}$  is human capital and  $\epsilon_{it}$  is the idiosyncratic error. However, it is widely accepted that the effect of FDI on economic growth has the cross-sectional units heterogeneity due to the environment and some relevant policies on different provinces and regions. Also, the economic growth may have a trend over time, which cannot be captured by other variables. So, the model may contain the individual and time effects ( $\mu_i$  and  $\eta_t$ ). Then, one may wonder whether the individual and time effects exist. Therefore, the possible model for this example is as follows,

$$y_{it} = \beta_1(\text{FDI}/Y)_{it} + \beta_2 \ln(\text{DI}/Y)_{it} + \beta_3 n_{it} + \beta_4 h_{it} + \mu_i + \eta_t + \epsilon_{it}. \quad (6.5)$$

In this section, the model (6.5) can also be used to study the effect on economic growth of China. Our model utilized the panel data of 30 provincial regions in China from 1992 to 2017 (excluding Tibet) and the data are from China Statistical Yearbook and statistical yearbooks of various regions.

Since the proposed test method is very general, it can also be used for two-way error component panel data model without time-invariant variables. In order to illustrate the efficiency of the proposed test, we chose three types of unbalanced panel data sample with  $T = [10, 18, 26]$  from the original database and the number of cross-section units are random.

**Table 6.** Values of three test statistics.

	$T_\mu$	$T_\eta$	$T_{\mu\eta}$
Case 1	201.1669	2514.3201	498.8891
Case 2	278.1351	4467.8232	839.4362
Case 3	143.8810	1598.8783	374.0413

Table 6 gives the results of the test statistics in three different cases. The p-value of  $T_\mu$ ,  $T_\eta$  and  $T_{\mu\eta}$  are all less than 0.0001. So, we have to reject the null hypothesis at a significance level of 0.05. Similarly, the time effect test statistic  $T_\eta$  is larger than  $T_\mu$ , which has the same result we obtained in the simulation. Thus, we conclude that the effect of FDI on economic growth has the heterogeneity over provinces and the time effect should be included in the model to get the correct estimator when studying this problem.

## 7. Conclusions

In this paper, we construct three test statistics for individual effect, time effect and jointly effects in unbalanced panel data models which include the time-invariant variables. The test is based on the moment method using the difference of two variance estimators of error. When we test the individual effect, the test statistic is efficient no matter the existence of the time effect, and vice versa. Furthermore, three test statistics are robust to the correlations between the explanatory variables and individual or time effect. There is no need to make some distributional assumption on error term. The simulation results show that the proposed test statistics are robust under various situations and they all have good finite sample properties. We also studied the relationship between FDI and economic growth, and found that the effect of FDI on economic growth has heterogeneity and common time characteristics.

In fact, traditional linear models are not efficient in actual application. The nonparametric model has been widely used for several years and we can consider a more effective test method of various nonparametric panel models in the future. In addition, the time-invariant variables could be endogenous. We can study the estimation and test methods for this kind of situation such as embedding the instrumental variable method or system generalized method of moments.

## Acknowledgments

This work was supported by the Innovation Research for the Postgraduates of Guangzhou University under Grant number 2021GDJC-D01.

## Conflict of interest

The authors declare there is no conflicts of interest.

## References

1. S. Liu, X. Shen, T. Jiang, P. Failler, Impacts of the financialization of manufacturing enterprises on total factor productivity: empirical examination from China's listed companies, *Green Finance*, **3** (2021), 59–89. <https://doi.org/10.3934/GF.2021005>
2. Y. Yao, D. Hu, C. Yang, Y. Tan, The impact and mechanism of fintech on green total factor productivity, *Green Finance*, **3** (2021), 198–221. <https://doi.org/10.3934/GF.2021011>
3. T. Tuzcuoğlu, The impact of financial fragility on firm performance: an analysis of BIST companies, *Quant. Finance Econ.*, **4** (2020), 310–342. <https://doi.org/10.3934/QFE.2020015>
4. Y. Tan, Z. Li, S. Liu, M. I. Nazir, M. Haris, Competitions in different banking markets and shadow banking: evidence from China, *Int. J. Emerging Mark.*, **17** (2021), 1465–1483. <https://doi.org/10.1108/IJOEM-04-2020-0401>
5. T. S. Breusch, A. R. Pagan, The lagrange multiplier test and its applications to model specification in econometrics, *Rev. Econ. Stud.*, **47** (1980), 239–253. <https://doi.org/10.2307/2297111>
6. Y. Honda, Testing the error components model with non-normal disturbances, *Rev. Econ. Stud.*, **52** (1985), 681–690. <https://doi.org/10.2307/2297739>

7. B. H. Baltagi, Q. Li, A lagrange multiplier test for the error components model with incomplete panels, *Econom. Rev.*, **9** (1990), 681–690. <https://doi.org/10.2307/2297739>
8. K. A. Bera, S. Walter, Y. Mann, Tests for the error component model in the presence of local misspecification, *J. Econom.*, **101** (2001), 681–690. [https://doi.org/10.1016/S0304-4076\(00\)00071-3](https://doi.org/10.1016/S0304-4076(00)00071-3)
9. B. H. Baltagi, Y. Chang, Q. Li, Monte Carlo results on several new and existing tests for the error component model, *J. Econom.*, **54** (1992), 95–120. [https://doi.org/10.1016/0304-4076\(92\)90101-V](https://doi.org/10.1016/0304-4076(92)90101-V)
10. J. Wu, G. Li, Moment-based tests for individual and time effects in panel data models, *J. Econom.*, **178** (2014), 569–581. <https://doi.org/10.1016/j.jeconom.2013.08.020>
11. M. H. Pesaran, Q. Zhou, Estimation of time-invariant effects in static panel data models, *Econom. Rev.*, **37** (2018), 1137–1171. <https://doi.org/10.1080/07474938.2016.1222225>
12. K. Sebastian, S. Claudia, Estimation of linear dynamic panel data models with time-invariant regressors, *J. Appl. Econom.*, **34** (2019), 526–546. <https://doi.org/10.1002/jae.2681>
13. J. Chen, R. Yue, J. Wu, Testing for individual and time effects in the two-way error component model with time-invariant regressors, *Econ. Modell.*, **92** (2020), 216–229. <https://doi.org/10.1016/j.econmod.2020.01.002>
14. J. Wu, J. Qin, Q. Ding, A moment-based test for individual effects in the error component model with incomplete panels, *Stat. Probab. Lett.*, **104** (2015), 153–162. <https://doi.org/10.1016/j.spl.2015.05.013>
15. J. Wu, G. Li, Q. Xia, Moment-based tests for random effects in the two-way error component model with unbalanced panels, *Econ. Modell.*, **1** (2018), 61–76. <https://doi.org/10.1016/j.econmod.2018.05.003>
16. J. Chen, R. Yue, J. Wu, Hausman-type tests for individual and time effects in the panel regression model with incomplete data, *J. Korean Stat. Soc.*, **47** (2018), 347–363. <https://doi.org/10.1016/j.jkss.2018.04.002>
17. P. Wu, L. Zhu, An orthogonality-based estimation of moments for linear mixed models, *Scand. J. Stat.*, **37** (2010), 253–263. <https://doi.org/10.1111/j.1467-9469.2009.00673.x>
18. J. A. Hausman, Specification tests in econometrics, *Econometrica*, **46** (1978), 1251–1271.
19. O. S. Sukharev, Structure analysis of income and risk dynamics in models of economic growth, *Quant. Finance Econ.*, **4** (2020), 1–18. <https://doi.org/10.3934/QFE.2020001>
20. M. Iuliana, Is financial development good for economic growth? Empirical insights from emerging European countries, *Quant. Finance Econ.*, **4** (2020), 653–678. <https://doi.org/10.3934/QFE.2020030>
21. L. Carlsen, Decent work and economic growth in the European Union. A partial order analysis of Eurostat SDG 8 data, *Green Finance*, **3** (2021), 483–494. <https://doi.org/10.3934/GF.2021022>
22. B. Richard, B. Henrik, F. Rikard, Multinationals, endogenous growth, and technological spillovers: Theory and evidence, *Rev. Int. Econ.*, **13** (2005), 945–963. <https://doi.org/10.1111/j.1467-9396.2005.00546.x>
23. E. Borensztein, J. D. Gregorio, J. W. Lee, How does foreign direct investment affect economic growth, *J. Int. Econ.*, **45** (1998), 115–135. [https://doi.org/10.1016/S0022-1996\(97\)00033-0](https://doi.org/10.1016/S0022-1996(97)00033-0)

24. C. Kottaridi, T. Stengos, Foreign direct investment, human capital and non-linearities in economic growth, *J. Macroeconomics*, **32** (2010), 858–871. <https://doi.org/10.1016/j.jmacro.2010.01.004>
25. J. Shao, Z. Xiao, R. Xu, Estimation with unbalanced panel data having covariate measurement error, *J. Stat. Plann. Inference*, **141** (2011), 800–808. <https://doi.org/10.1016/j.jspi.2010.08.003>

## Appendix

The proof of the theorems is as follows.

*Proof.* For theorem 2.2. We can refer the proof of Pesaran [11]. For each group  $\mathcal{N}_l$ , according to the estimation,

$$\tilde{u}_l - \bar{u} = (\mu_l - \bar{\mu}) + (\bar{\epsilon}_l - \bar{\epsilon}) + (Z_l - \bar{Z})' \gamma - (\bar{X}_l - \bar{X})' (\hat{\beta} - \beta) \quad (\text{A.1})$$

where  $\bar{\mu} = \frac{1}{n_l} \sum_{i=1}^{n_l} \mu_{li}$ ,  $\bar{\epsilon} = \frac{1}{n_l} \sum_{i=1}^{n_l} \epsilon_{li}$ . For the identification, we have  $\sum_{l=1}^L \sum_{i=1}^{n_l} (Z_{li} - \bar{Z})(\bar{\mu} + \bar{\epsilon}) = 0$ . On the basis of the estimator  $\hat{\gamma}$ , we have

$$\hat{\gamma} - \gamma = \left[ \frac{1}{n} \sum_{l=1}^L \sum_{i=1}^{n_l} (Z_{li} - \bar{Z})(Z_{li} - \bar{Z})' \right]^{-1} \frac{1}{n} \sum_{l=1}^L \sum_{i=1}^{n_l} (Z_{li} - \bar{Z}) \zeta_{li} \quad (\text{A.2})$$

where  $\zeta_{li} = \mu_{li} + \bar{\epsilon}_l - (\bar{X}_{li} - \bar{X})' (\hat{\beta} - \beta)$ . For each group  $\mathcal{N}_l$ , noting that  $\hat{\beta}$  is consistent to  $\beta$ , we have

$$E[(Z_{li} - \bar{Z}) \zeta_{li} | Z, X, \mu] = (Z_{li} - \bar{Z}) \mu_{li} \quad (\text{A.3})$$

For  $i = 1, \dots, n_l$ ,  $Z$  and  $X$  is the matrix that contains  $Z_i$  and  $X_{it}$  respectively, and  $\mu$  is the vector that contains  $\mu_{li}$ . Under the Assumption 2.3,  $E((Z_{li} - \bar{Z}) \mu_{li}) = 0$ . We have  $E(\hat{\gamma}) = \gamma$  that  $\hat{\gamma}$  is consistent to  $\gamma$ .

Then, we will proof that the  $\hat{\gamma}$  is a  $\sqrt{n}$  consistent estimator of  $\gamma$ .

$$\begin{aligned} \frac{1}{n} \sum_{l=1}^L \sum_{i=1}^{n_l} (Z_{li} - \bar{Z}) \zeta_{li} &= \frac{1}{n} \sum_{l=1}^L \sum_{i=1}^{n_l} (Z_{li} - \bar{Z}) (\mu_{li} + \bar{\epsilon}_l) \\ &\quad - \left[ \frac{1}{n} \sum_{l=1}^L \sum_{i=1}^{n_l} (Z_{li} - \bar{Z}) (\bar{X}_{li} - \bar{X})' \right] (\hat{\beta} - \beta) \end{aligned} \quad (\text{A.4})$$

Under the Assumptions 2.2 and 2.3,  $\frac{1}{n} \sum_{l=1}^L \sum_{i=1}^{n_l} (Z_{li} - \bar{Z}) (\bar{X}_{li} - \bar{X})'$  will converge to finite value. In this paper, we assume that the individual effects in each group are the same, so the variance of the effect is constant  $\sigma_\mu^2$ .

$$\frac{1}{n} \sum_{l=1}^L \sum_{i=1}^{n_l} (Z_{li} - \bar{Z}) (\mu_{li} + \bar{\epsilon}_l) \rightarrow_d N(0, \omega_{l_i T_l} Q_{zz}) \quad (\text{A.5})$$

where  $\omega_{l_i T_l} = \sigma_\mu^2 + \frac{r_i^2}{T_l} + \frac{1}{T_l^2} \Sigma r_i(s, t)$ . Then we have

$$\frac{1}{n} \sum_{l=1}^L \sum_{i=1}^{n_l} (Z_{li} - \bar{Z}) \zeta_{li} = O_p(N^{-1/2}) + O_p(\hat{\beta} - \beta) \quad (\text{A.6})$$



We know that  $\hat{\beta} - \beta = O_p(N^{-1/2})$ . Then,

$$\hat{\gamma} - \gamma = O_p(N^{-1/2}) \quad (\text{A.7})$$

Next, obtain the asymptotic variance of the estimator  $\hat{\gamma}$ . According to the Eq (A.2), the first term of right side,

$$\begin{aligned} \frac{1}{n} \sum_{l=1}^L \sum_{i=1}^{n_l} (Z_{li} - \bar{Z})(Z_{li} - \bar{Z})' &= \sum_{l=1}^L \frac{n_l}{n} \frac{1}{n_l} \sum_{i=1}^{n_l} (Z_{li} - \bar{Z})(Z_{li} - \bar{Z})' \\ &= \sum_{l=1}^L m_l E \left[ (Z_{li} - EZ_{li})(Z_{li} - EZ_{li})' \right] + o_p(1) = \Sigma_3 + o_p(1) \end{aligned} \quad (\text{A.8})$$

where  $m_l = \lim_{n \rightarrow \infty} \frac{n_l}{n}$ , this setting is commonly used in Shao et al. [25].

$$\begin{aligned} \frac{1}{\sqrt{n}} \sum_{l=1}^L \sum_{i=1}^{n_l} (Z_{li} - \bar{Z})\zeta_{li} &= \frac{1}{\sqrt{n}} \sum_{l=1}^L \sum_{i=1}^{n_l} (Z_{li} - \bar{Z})\mu_{li} \\ &+ \frac{1}{\sqrt{n}} \sum_{l=1}^L \sum_{i=1}^{n_l} (Z_{li} - \bar{Z}) \left[ \bar{\epsilon}_{li} - (\bar{X}_{li} - \bar{X})' (\hat{\beta} - \beta) \right] \end{aligned} \quad (\text{A.9})$$

The variance of the first term of above formula,

$$\begin{aligned} \text{var} \left[ \frac{1}{\sqrt{n}} \sum_{l=1}^L \sum_{i=1}^{n_l} (Z_{li} - \bar{Z})\mu_{li} \right] &= \text{var} \left[ \sum_{l=1}^L \sqrt{\frac{n_l}{n}} \frac{1}{\sqrt{n_l}} \sum_{i=1}^{n_l} (Z_{li} - \bar{Z})\mu_{li} \right] \\ &= \sum_{l=1}^L m_l \sigma_{\mu} E((Z_{li} - EZ_{li})(Z_{li} - EZ_{li})') + o_p(1) \\ &= \Sigma_4 + o_p(1) \end{aligned} \quad (\text{A.10})$$

where  $\Sigma_4 = \text{var} \left[ \sum_{l=1}^L \sqrt{\frac{n_l}{n}} \frac{1}{\sqrt{n_l}} \sum_{i=1}^{n_l} (Z_{li} - \bar{Z})\mu_{li} \right]$ . First, we know that

$$\begin{aligned} \frac{1}{\sqrt{n}} \sum_{l=1}^L \sum_{i=1}^{n_l} (Z_{li} - \bar{Z})(\bar{X}_{li} - \bar{X})' &= \sum_{l=1}^L \sqrt{\frac{n_l}{n}} \frac{1}{\sqrt{n_l}} \sum_{i=1}^{n_l} (Z_{li} - \bar{Z})(\bar{X}_{li} - \bar{X})' \\ &= \sum_{l=1}^L m_l [E(Z_{li}\bar{X}_{li}') - E(Z_{li})E(\bar{X}_{li})'] + o_p(1) \\ &= \Sigma_5 + o_p(1) \end{aligned} \quad (\text{A.11})$$

where  $\Sigma_5 = \sum_{l=1}^L m_l [E(Z_{li}\bar{X}_{li}') - E(Z_{li})E(\bar{X}_{li})']$ . Then, the second term,

$$\begin{aligned} \frac{1}{\sqrt{n}} \sum_{l=1}^L \sum_{i=1}^{n_l} (Z_{li} - \bar{Z}) \left[ \bar{\epsilon}_{li} - (\bar{X}_{li} - \bar{X})' (\hat{\beta} - \beta) \right] \\ &= \sum_{l=1}^L \sqrt{\frac{n_l}{n}} \frac{1}{\sqrt{n_l}} \sum_{i=1}^{n_l} \left[ (Z_{li} - \bar{Z})T_l^{-1}t_{T_l}' - \Sigma_5 \Sigma_1^{-1} (X_{li} - EX_{li})' P_{T_l} \right] \epsilon_{li} \\ &= \sum_{l=1}^L m_l \frac{1}{\sqrt{n_l}} \sum_{i=1}^{n_l} \xi_{li} \end{aligned} \quad (\text{A.12})$$

where  $\xi_{li} = [(Z_{li} - \bar{Z})T_l^{-1}l'_{T_l} - \Sigma_5\Sigma_1^{-1}(X_{li} - EX_{li})'P_{T_l}] \epsilon_{li}$ . According to Assumption 2.3, we have

$$E(\xi_{li}) = 0 \quad (\text{A.13})$$

and

$$\begin{aligned} \text{var}\left(\sum_{l=1}^L m_l \frac{1}{\sqrt{n_l}} \sum_{i=1}^{n_l} \xi_{li}\right) &= \sum_{l=1}^L m_l E(\bar{\epsilon}_{li}^2 (Z_{li} - EZ_{li})(Z_{li} - EZ_{li})') \\ &\quad - \sum_{l=1}^L m_l E[(Z_{li} - EZ_{li})l'_{T_l} \epsilon_{li} \epsilon'_{li} P_{T_l} (X_{li} - EX_{li}) \Sigma_1^{-1} \Sigma_5] \\ &\quad - \Sigma_5 \Sigma_1^{-1} \sum_{l=1}^L m_l E[(Z_{li} - EZ_{li})l'_{T_l} \epsilon_{li} \epsilon'_{li} P_{T_l} (X_{li} - EX_{li})'] \\ &\quad + \Sigma_5 \Sigma_1^{-1} \Sigma_2 \Sigma_1^{-1} \Sigma_5' + o_p(1) = \Omega_1 + o_p(1) \end{aligned} \quad (\text{A.14})$$

The covariance of the two term in Eq (A.9),

$$\begin{aligned} &E\left[\frac{1}{\sqrt{n}} \sum_{l=1}^L \sum_{i=1}^{n_l} (Z_{li} - \bar{Z}) \mu_{li} \sum_{l=1}^L m_l \frac{1}{\sqrt{n_l}} \sum_{i=1}^{n_l} \xi_{li}\right] \\ &= \sum_{l=1}^L m_l E[\mu_{li} \bar{\epsilon}_{li} (Z_{li} - EZ_{li})(Z_{li} - EZ_{li})'] \\ &\quad - \sum_{l=1}^L m_l E[(Z_{li} - EZ_{li}) \mu_{li} \epsilon'_{li} P_{T_l} (X_{li} - EX_{li})] \Sigma_1^{-1} \Sigma_5 + o_p(1) \\ &= \Sigma_6 - \Sigma_7 \Sigma_1^{-1} \Sigma_5 + o_p(1) = \Omega_2 + o_p(1) \end{aligned} \quad (\text{A.15})$$

where  $\Sigma_6 = \sum_{l=1}^L m_l E[\mu_{li} \bar{\epsilon}_{li} (Z_{li} - EZ_{li})(Z_{li} - EZ_{li})']$  and  $\Sigma_7 = \sum_{l=1}^L m_l E[(Z_{li} - EZ_{li}) \mu_{li} \epsilon'_{li} P_{T_l} (X_{li} - EX_{li})]$ .

Following the central limit theorem, we have

$$\frac{1}{\sqrt{n}} \sum_{l=1}^L \sum_{i=1}^{n_l} (Z_{li} - \bar{Z}) \zeta_{li} \rightarrow_d N(0, \Sigma_4 + \Omega_1 + \Omega_2) \quad (\text{A.16})$$

Therefore, we can obtain the asymptotic distribution of  $\hat{\gamma}$ ,

$$\sqrt{n}(\hat{\gamma} - \gamma) \xrightarrow{d} N(0, \Sigma_3^{-1}(\Sigma_4 + \Omega_1 + \Omega_2)\Sigma_3^{-1}) \quad (\text{A.17})$$

**Lemma 7.1.** Suppose that  $E\epsilon_{it}^4 < \infty$ ,  $EX_{it}^4 < \infty$  and under the Assumption 2.1 for each  $i, t$ ,  $\hat{\sigma}_{0\epsilon}^2$  and  $\hat{\lambda}_\epsilon^4$  is consistent estimator of  $\sigma_\epsilon^2$  and  $\lambda_\epsilon^4$  respectively. The proof can be seen the proof of Lemma 2 of Wu et al. [15]. This is because the time-invariant variables are removed after the centering transformation.

*Proof.* For theorem 3.1. According to the lemma 7.1 and Eq (2.4), we have

$$\begin{aligned} \sqrt{n}\hat{\sigma}_{0\epsilon}^2 &= \frac{\sqrt{n}}{c_1} \sum_{l=1}^L \sum_{i=1}^{n_l} E\|Q'_{T_l} \hat{\epsilon}_{li}\|^2 \\ &= \frac{\sqrt{n}}{c_1} \sum_{l=1}^L \sum_{i=1}^{n_l} \epsilon'_{li} P_{T_l} \epsilon_{li} + o_p(1) \end{aligned} \quad (\text{A.18})$$

and

$$\begin{aligned}\sqrt{n}\hat{\sigma}_{1\epsilon}^2 &= \frac{\sqrt{n}}{c_4} \sum_{l=1}^L \sum_{i=1}^{n_l} E\|Q'_{T_l} \hat{\epsilon}_{i_l}\|^2 \\ &= \frac{\sqrt{n}}{c_4} \sum_{l=1}^L \sum_{i=1}^{n_l} E\|\tilde{X}_{i_l}(\beta - \hat{\beta}) + (\iota_{T_l} \tilde{Z}'_{i_l})(\gamma - \hat{\gamma}) + \iota_{T_l} \tilde{\mu}_{i_l} + \tilde{\epsilon}_{i_l}\|^2\end{aligned}\quad (\text{A.19})$$

Note that  $\sqrt{n}(\hat{\beta} - \beta) = O_p(1)$  and  $\sqrt{n}(\hat{\gamma} - \gamma) = O_p(1)$ . Besides, use some limit theorems and the proof in Appendix of Chen et al. [13] and we suppose the explanatory variable  $X_{it}$  and time-invariant variable  $Z_{it}$  are i.i.d sequence and  $EX_{it}^4 < \infty$ ,  $EZ_{it}^4 < \infty$ . We have

$$\sqrt{n}\hat{\sigma}_{1\epsilon}^2 = \frac{\sqrt{n}}{c_4} \sum_{l=1}^L \sum_{i=1}^{n_l} \epsilon'_{i_l} \epsilon_{i_l} + \sigma_1^2 + o_p(1) \quad (\text{A.20})$$

where  $\sigma_1^2$  is defined in Assumption 3.2.

$$\begin{aligned}\sqrt{n}\hat{\sigma}_{1\epsilon}^2 - \hat{\sigma}_{0\epsilon}^2 &= \frac{\sqrt{n}}{c_4} \sum_{l=1}^L \sum_{i=1}^{n_l} \epsilon'_{i_l} \epsilon_{i_l} - \frac{\sqrt{n}}{c_1} \sum_{l=1}^L \sum_{i=1}^{n_l} \epsilon'_{i_l} P_{T_l} \epsilon_{i_l} + \sigma_1^2 + o_p(1) \\ &= \frac{1}{\sqrt{n}} \sum_{l=1}^L \sum_{i=1}^{n_l} \varsigma_{i_l} + \sigma_1^2 + o_p(1)\end{aligned}\quad (\text{A.21})$$

where  $\varsigma_{i_l} = \frac{n}{c_4} \epsilon'_{i_l} \epsilon_{i_l} - \frac{n}{c_1} \epsilon'_{i_l} P_{T_l} \epsilon_{i_l}$  and  $\varsigma_{i_l}$  is independent cross units. It holds that  $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{l=1}^L \sum_{i=1}^{n_l} E\varsigma_{i_l} = 0$  and

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{l=1}^L \sum_{i=1}^{n_l} E(\varsigma_{i_l})^2 = 0$$

where  $\omega_n = a_n \lambda_\epsilon^4 + b_n (\sigma_{0\epsilon}^2)^2$  is used to standardize the test statistics. The

$$a = \lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{l=1}^L n_l \left( \frac{n^2}{c_4^2} T_l + \frac{n^2}{c_1^2} (T_l + \frac{1}{T_l} - 2) - \frac{2n^2}{c_1 c_4} (T_l - 1) \right) \quad (\text{A.22})$$

$$b = \lim_{n \rightarrow \infty} b_n = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{l=1}^L n_l \left( \frac{n^2}{c_4^2} T_l (T_l - 1) + \frac{n^2}{c_1^2} (T_l - 1) (T_l + \frac{3}{T_l} - 2) - \frac{2n^2}{c_1 c_4} (T_l - 1)^2 \right) \quad (\text{A.23})$$

The above proof process is the same as Wu et al. [15] and we also have the same result  $\omega_n^{-1/2} \sqrt{n}(\hat{\sigma}_{1\epsilon}^2 - \hat{\sigma}_{0\epsilon}^2) - \Phi_n^{-1/2} \sigma_1^2 \rightarrow_d N(0, 1)$ . Thus, the proof of theorem is complete.



AIMS Press

©2022 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)