Journal of Econometrics 178 (2014) 569–581

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

Moment-based tests for individual and time effects in panel data models

Jianhong Wu [a], Guodong Li [b,*]

[a] School of Statistics and Mathematics, Zhejiang Gongshang University, Hangzhou, China

[b] Department of Statistics and Actuarial Science, University of Hong Kong, Pokfulam Road, Hong Kong

ARTICLE INFO

Article history:
Received 30 September 2011
Received in revised form
8 April 2013
Accepted 20 August 2013
Available online 28 August 2013

JEL classification:
C12
C23

Keywords:
Estimation of moment
Hausman-type test
Individual effect
Panel data
Time effect

ABSTRACT

This paper proposes two Hausman-type tests respectively for individual and time effects in a two-way error component regression model by comparing estimators of the variance of the idiosyncratic error at different robust levels. They are both robust to the presence of the other effect, and the test for the individual effect has a larger asymptotic power than the corresponding ANOVA F test when the effects are correlated with covariates. Tests jointly for both effects are also discussed. Monte Carlo evidence shows their good size properties and better power properties than competing tests, and the application to the crime rate study gives further support.

1. Introduction

Consider the two-way error component regression model,

$$y_{it} = \alpha + X_{it}'\beta + \mu_i + \eta_t + u_{it},$$
$$i = 1, 2, \ldots, n,\ t = 1, 2, \ldots, T, \quad (1)$$

where X_{it} is the p-dimensional vector of covariates, α is a scalar, β is the vector of coefficients of covariates, and u_{it} corresponds to the idiosyncratic error. The individual effect μ_i and the time effect η_t can capture the heterogeneity of individuals and time points, and hence model (1) is able to better explain many real data, see Hsiao (2003), Wooldridge (2002) and Baltagi (2008). It is an important topic in this literature to test for the existence of the individual effect and the time effect since their involvements, if unnecessary, will make the inference complicated and even inefficient when T or n is fixed.

The Lagrange multiplier (LM) test has been widely discussed in this literature since Breusch and Pagan (1980). Honda (1985) derived two uniformly most powerful one-sided tests by modifying the test statistics in Breusch and Pagan (1980). These tests are based on the one-way error component models, i.e. the null hypotheses correspond to the case without any effect, and the sizes may be distorted due to the presence of the time (individual) effect when the individual (time) effect is tested. Bera et al. (2001) considered LM tests for the individual effect in the presence of serial correlation in the idiosyncratic error, and they are also based on the one-way error component model. Baltagi et al. (1992) proposed some LM tests based on the more general two-way error component model (1), and they are robust to the presence of the redundant effect since its variance is treated as a parameter and estimated from the data. Such robustness depends on the reliability of the estimated variances of the redundant effect and, for this reason, neither n or T can be too small in Baltagi et al. (1992). Secondly, the construction of the aforementioned LM tests needs the assumption of normality and independence among the covariates, effects and the idiosyncratic error. They may still be valid when the assumption of normality is relaxed, see Honda (1985) and Baltagi et al. (1992). However, when the covariates are correlated with the individual effect and/or the time effect (Cornwell and Trumbull, 1994), the commonly used feasible generalized least squares (GLS) estimation for the coefficients β is biased (Hausman, 1978), and these tests are also expected to be biased.

This paper makes use of two transformations on model (1) to wipe out respectively these two effects: the centering for the time effect as in Baltagi (2008) and an orthogonal transformation for the

* Corresponding author. Tel.: +852 2859 1986; fax: +852 2858 9041.
E-mail addresses: wjhstat1@gmail.com (J. Wu), gdli@hku.hk (G. Li).

0304-4076/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jeconom.2013.08.020

individual effect as in Wu and Zhu (2010). Based on these transformed models, we can work out three different estimators of the variance of the idiosyncratic error, and they are consistent respectively under the presence of the time effect only, the individual effect only and both effects. Following the idea of Hausman's specification test (Hausman, 1978), we can construct two tests, one for the individual effect in Section 2 and the other for the time effect in Section 3. As in Baltagi et al. (1992), these two proposed tests are robust to the presence of one effect when the other is tested, while robustness here is achieved by employing the corresponding transformation to difference out the redundant effect. This mechanism makes sure that the tests can be applied to more general cases than random effects such as the fixed effect and the case that these two effects are random but correlated with the covariates. We also demonstrate that the commonly used F tests in the analysis of covariance (ANOVA F tests) are asymptotically equivalent to variations of our tests. When the covariates are correlated with the effects, comparing with the corresponding ANOVA F test, the proposed test for the individual effect is shown to be asymptotically more powerful, while that for the time effect has the same asymptotic power.

Similarly, we also construct a Hausman-type test for individual and time effects jointly in Section 4, however, it is less powerful in detecting the existence of the time effect asymptotically. Fortunately, these two proposed test statistics in Sections 2 and 3 are shown to be asymptotically independent when the idiosyncratic error is independent of the covariates. This makes it convenient to combine them to jointly check the presence of both individual and time effects. As an example, we give a combined test in Section 4 with the asymptotic distribution under the null hypothesis being a mixed chi-square distribution as in Baltagi et al. (1992).

Monte Carlo evidence in Section 5 shows that the proposed tests have good size properties and better power properties than competing tests such as those in Baltagi et al. (1992) when the covariates are correlated with the effects. We apply our tests to the crime rate study in Section 6, and show that our tests are more informative about the existence of county heterogeneity and time heterogeneity of the crime rate than the existing tests. A short discussion is given in Section 7. Proofs of theorems and corollaries are given in the Appendix. Estimators and their asymptotic normalities of the higher order moments of the idiosyncratic error and the individual effect are also presented in the Appendix.

2. Test for the individual effect

Let $\mathbf{y}_i = (y_{i1}, y_{i2}, \dots, y_{iT})'$, $\mathbf{X}_i = (X_{i1}, X_{i2}, \dots, X_{iT})'$, $\mathbf{u}_i = (u_{i1}, u_{i2}, \dots, u_{iT})'$, $\boldsymbol{\eta} = (\eta_1, \eta_2, \dots, \eta_T)'$, and $\boldsymbol{\iota}_T$ be a T-dimensional vector with all elements equal to one. Model (1) can be rewritten into the vector form,

$$\mathbf{y}_i = \alpha\boldsymbol{\iota}_T + \mathbf{X}_i\beta + \mu_i\boldsymbol{\iota}_T + \boldsymbol{\eta} + \mathbf{u}_i, \quad i = 1, \dots, n, \tag{2}$$

where the idiosyncratic errors $\{u_{it}\}$ are independent and identically distributed ($i.i.d.$) across individuals and time points, $\{\mathbf{X}_i\}$ are $i.i.d.$ across individuals, and $E(X_{it}u_{is}) = 0$ for $s \geq t$, i.e. covariates X_{it} are predetermined. The asymptotic results in this paper are based on the assumption that n tends to infinity and T is fixed, and this is a commonly used setting in the literature, see Baltagi (2008). Denote by σ_u^2 the variance of the idiosyncratic error u_{it}, and it will play a key role in this paper.

We first consider the test for the individual effect, and it is usually assumed to be a random variable with mean zero and finite variance σ_μ^2. The hypotheses of the test can be formalized as follows,

$$H_0^\mu : \sigma_\mu^2 = 0 \quad \text{vs} \quad H_1^\mu : \sigma_\mu^2 > 0.$$

To construct a test robust to the presence of the time effect $\boldsymbol{\eta}$, we first wipe it out by centering each term in model (2), and it results in

$$\widetilde{\mathbf{y}}_i = \widetilde{\mathbf{X}}_i\beta + \widetilde{\mu}_i\boldsymbol{\iota}_T + \widetilde{\mathbf{u}}_i, \quad i = 1, \dots, n, \tag{3}$$

where $\widetilde{\mathbf{y}}_i = \mathbf{y}_i - n^{-1}\sum_{j=1}^n \mathbf{y}_j$, $\widetilde{\mathbf{X}}_i = \mathbf{X}_i - n^{-1}\sum_{j=1}^n \mathbf{X}_j$, $\widetilde{\mu}_i = \mu_i - n^{-1}\sum_{j=1}^n \mu_j$, and $\widetilde{\mathbf{u}}_i = \mathbf{u}_i - n^{-1}\sum_{j=1}^n \mathbf{u}_j$. Note that $\|\boldsymbol{\iota}_T\| = T^{1/2}$, where $\|\cdot\|$ is the Euclidean norm. Then we can find a matrix Q such that $(T^{-1/2}\boldsymbol{\iota}_T, Q)$ is a $T \times T$ orthogonal matrix. Conducting an orthogonal transformation on model (3) with the matrix $(T^{-1/2}\boldsymbol{\iota}_T, Q)$, we have that

$$\boldsymbol{\iota}_T'\widetilde{\mathbf{y}}_i = \boldsymbol{\iota}_T'\widetilde{\mathbf{X}}_i\beta + T\widetilde{\mu}_i + \boldsymbol{\iota}_T'\widetilde{\mathbf{u}}_i, \tag{4}$$

$$Q'\widetilde{\mathbf{y}}_i = Q'\widetilde{\mathbf{X}}_i\beta + Q'\widetilde{\mathbf{u}}_i, \tag{5}$$

where $i = 1, \dots, n$. The individual effect μ_i is only present in model (4), which is a single equation while model (5) is a $(T-1)$-dimensional equation. It is then reasonable to use model (5) only to obtain a consistent estimator of σ_u^2. It holds that $QQ' = I_T - T^{-1}\boldsymbol{\iota}_T\boldsymbol{\iota}_T'$ and $Q'Q = I_{T-1}$, where I_m is an m-dimensional identity matrix.

Consider the ordinary least squares (OLS) estimation for model (5),

$$\widehat{\beta} = \underset{\beta}{\operatorname{argmin}} \sum_{i=1}^n \| Q'\widetilde{\mathbf{y}}_i - Q'\widetilde{\mathbf{X}}_i\beta \|^2 = \left(\sum_{i=1}^n \widetilde{\mathbf{X}}_i'P_{\iota_T}^{\perp}\widetilde{\mathbf{X}}_i\right)^{-1} \sum_{i=1}^n \widetilde{\mathbf{X}}_i'P_{\iota_T}^{\perp}\widetilde{\mathbf{y}}_i, \tag{6}$$

where $P_{\iota_T}^{\perp} = QQ' = I_T - T^{-1}\boldsymbol{\iota}_T\boldsymbol{\iota}_T'$ is independent of the matrix Q. Note that the above estimator is the same as the so-called Within estimator in Baltagi (2008). Denote $q_1 = T^{-1/2}\boldsymbol{\iota}_T$, $Q = (q_2, q_3, \dots, q_T)$ and $q_l = (q_{1l}, q_{2l}, \dots, q_{Tl})'$ for $1 \leq l \leq T$. It holds that

$$E\|Q'\widetilde{\mathbf{u}}_i\|^2 = E\left[\sum_{l=2}^T (q_l'\widetilde{\mathbf{u}}_i)^2\right] = a_n\sigma_u^2,$$

and then, from model (5), we can estimate the variance of the idiosyncratic error u_{it} by

$$\widehat{\sigma}_{0u}^2 = a_n^{-1} \cdot \frac{1}{n}\sum_{i=1}^n \|Q'(\widetilde{\mathbf{y}}_i - \widetilde{\mathbf{X}}_i\widehat{\beta})\|^2 = a_n^{-1} \cdot \frac{1}{n}\sum_{i=1}^n (\widetilde{\mathbf{y}}_i - \widetilde{\mathbf{X}}_i\widehat{\beta})'P_{\iota_T}^{\perp}(\widetilde{\mathbf{y}}_i - \widetilde{\mathbf{X}}_i\widehat{\beta}), \tag{7}$$

where $a_n = (T-1)(n-1)/n$. Under some regularity conditions, it holds that, regardless of the presence of individual and time effects,

$$\sqrt{n}(\widehat{\beta} - \beta) \to_d N(0, \Sigma_1^{-1}\Sigma_2\Sigma_1^{-1})$$

as $n \to \infty$, and $\widehat{\sigma}_{0u}^2$ is consistent, where $\Sigma_1 = E(\mathbf{X}_i'P_{\iota_T}^{\perp}\mathbf{X}_i) - E(\mathbf{X}_i')P_{\iota_T}^{\perp}E(\mathbf{X}_i)$ and $\Sigma_2 = E[(\mathbf{X}_i - E\mathbf{X}_i)'P_{\iota_T}^{\perp}\mathbf{u}_i\mathbf{u}_i'P_{\iota_T}^{\perp}(\mathbf{X}_i - E\mathbf{X}_i)]$.

Under the null hypothesis of $\sigma_\mu^2 = 0$, it holds that $\mu_1 = \cdots = \mu_n = 0$, and model (3) reduces to $\widetilde{\mathbf{y}}_i = \widetilde{\mathbf{X}}_i\beta + \widetilde{\mathbf{u}}_i$. This leads to another estimator of σ_u^2,

$$\widehat{\sigma}_{1u}^2 = b_n^{-1} \cdot \frac{1}{n}\sum_{i=1}^n \|\widetilde{\mathbf{y}}_i - \widetilde{\mathbf{X}}_i\widehat{\beta}\|^2$$

since $E\|\widetilde{\mathbf{u}}_i\|^2 = b_n\sigma_u^2$ with $b_n = T(n-1)/n$, and $\widehat{\beta}$ is given as in (6). Note that, under H_1^μ, $\widehat{\sigma}_{1u}^2$ is no longer consistent, however, $\widehat{\sigma}_{0u}^2$ is still consistent. Hence, a statistically significant difference between $\widehat{\sigma}_{1u}^2$ and $\widehat{\sigma}_{0u}^2$ can be interpreted as evidence against the null hypothesis H_0^μ. By following the idea of the Hausman specification test (Hausman, 1978), we can construct a test statistic as follows,

$$T_\mu = \Phi_n^{-1/2} \cdot \sqrt{n}(\widehat{\sigma}_{1u}^2 - \widehat{\sigma}_{0u}^2) = [0.5T(T-1)]^{1/2}\sqrt{n}\left(\frac{\widehat{\sigma}_{1u}^2}{\widehat{\sigma}_{0u}^2} - 1\right),$$

where the scalar $\Phi_n = 2(\widehat{\sigma}_{0u}^2)^2/[T(T-1)]$ is used to standardize the statistic.

To study the asymptotic behavior of the proposed test T_μ, we further assume that

(A1) $\{\mu_i\}$ are *i.i.d.* with mean zero and variance $\sigma_\mu^2 = n^{-1/2}\sigma_1^2$ with a constant $\sigma_1^2 \geq 0$, and

(A2) $E(\mu_i \mathbf{u}_i) = 0$, $n^{1/2}E(\mu_i^2\|\mathbf{u}_i\|^2) < \infty$, and $n^{1/4}E(\mu_i\mathbf{X}_i) = \Omega$.

Assumption (A1) gives the local alternatives, and the case with $\sigma_1^2 = 0$ corresponds to the null hypothesis H_0^μ. Assumption (A2) further restricts the distribution of μ_i such that we can study the asymptotic power under the possible correlations between the individual effect and covariates. Denote by $\{\mu_i^*\}$ an *i.i.d.* sequence with $E(\mu_i^*) = 0$, $\mathrm{var}(\mu_i^*) = \sigma_1^2$, $E(\mu_i^*\mathbf{u}_i) = 0$, $E(\mu_i^*\|\mathbf{u}_i\|)^2 < \infty$ and $E(\mu_i^*\mathbf{X}_i) = \Omega$. Let $\mu_i = n^{-1/4}\mu_i^*$ for $1 \leq i \leq n$, and then $\{\mu_i\}$ will satisfy both Assumptions (A1) and (A2).

Theorem 1. *Suppose that* $E(u_{it}^4) < \infty$, $E\|\mathbf{X}_i\|^4 < \infty$, $|\Sigma_1| > 0$ *and* $cov(\mathbf{X}_i'\boldsymbol{\iota}_T, \mathbf{u}_i'\boldsymbol{\iota}_T) = 0$. *If Assumptions* (A1) *and* (A2) *hold, then*

$$T_\mu \rightarrow_d \sigma_1^2[0.5T(T-1)]^{1/2}/\sigma_u^2 + \zeta_1$$

as $n \to \infty$, *where* ζ_1 *follows the standard normal distribution.*

From the above theorem, we can refer to the standard normal distribution for the critical values or p-values, and the test T_μ is nontrivial. Note that the quantity $E(u_{it}^4)$ is not present in the above theorem. Actually, it is canceled out in the derivation, however, the conditions of $E(u_{it}^4) < \infty$ and $E\|\mathbf{X}_i\|^4 < \infty$ are still needed to make some quantities in the derivation meaningful.

Besides the quantity $\widehat{\sigma}_{1u}^2$, we may consider another consistent estimator of σ_u^2 under H_0^μ,

$$\widetilde{\sigma}_{1u}^2 = b_n^{-1} \cdot \frac{1}{n}\sum_{i=1}^n \|\widetilde{\mathbf{y}}_i - \widetilde{\mathbf{X}}_i\widetilde{\beta}_1\|^2,$$

where $\widetilde{\beta}_1 = \mathrm{argmin}_\beta \sum_{i=1}^n \|\widetilde{\mathbf{y}}_i - \widetilde{\mathbf{X}}_i\beta\|^2$. This results in a new test statistic

$$T_\mu^* = [0.5T(T-1)]^{1/2}\sqrt{n}\left(\frac{\widetilde{\sigma}_{1u}^2}{\widehat{\sigma}_{0u}^2} - 1\right) = c_{1n} \cdot F_\mu - d_{1n},$$

where $c_{1n} = [0.5nT(T-1)]^{1/2}(1-T^{-1})(n-1)/[(n-1)(T-1)-p]$, $d_{1n} = [0.5nT(T-1)]^{1/2}/T$ and

$$F_\mu = \frac{\sum_{i=1}^n\{\|\widetilde{\mathbf{y}}_i - \widetilde{\mathbf{X}}_i\widetilde{\beta}_1\|^2 - \| Q'(\widetilde{\mathbf{y}}_i - \widetilde{\mathbf{X}}_i\widehat{\beta})\|^2\}/(n-1)}{\sum_{i=1}^n \| Q'(\widetilde{\mathbf{y}}_i - \widetilde{\mathbf{X}}_i\widehat{\beta})\|^2/[(n-1)(T-1)-p]}.$$

Notice that c_{1n} and d_{1n} are two positive constants, and the F_μ statistic in fact is the ANOVA F test statistic. The ANOVA F test uses the critical values that are quantiles from the F distribution with $n-1$ and $(n-1)(T-1)-p$ degrees of freedom, and this is justified under the assumption that u_{it} is normally distributed, see Baltagi (2008). Let $\Sigma_3 = \mathrm{var}(\mathbf{X}_i) = E[(\mathbf{X}_i - E\mathbf{X}_i)'(\mathbf{X}_i - E\mathbf{X}_i)]$.

Corollary 1. *If* $|\Sigma_3| > 0$ *and the conditions of Theorem 1 hold, then*

$$T_\mu^* \rightarrow_d (\sigma_1^2 - T^{-1}\boldsymbol{\iota}_T'\Omega\Sigma_3^{-1}\Omega'\boldsymbol{\iota}_T)[0.5T(T-1)]^{1/2}/\sigma_u^2 + \zeta_1$$

as $n \to \infty$, *where* Ω *and* ζ_1 *are defined as in Assumption* (A2) *and Theorem 1, respectively.*

Because $\Omega = 0$ as $\sigma_1^2 = 0$ and T_μ^* is simply an affine transformation of F_μ, a test based on T_μ^* and a critical value from the standard normal distribution is asymptotically equivalent to the ANOVA F test. It can be shown that $0 \leq T^{-1}\boldsymbol{\iota}_T'\Omega\Sigma_3^{-1}\Omega'\boldsymbol{\iota}_T \leq \sigma_1^2$, and the quantity $T^{-1}\boldsymbol{\iota}_T'\Omega\Sigma_3^{-1}\Omega'\boldsymbol{\iota}_T$ equals to zero if and only if $\mathrm{cov}(\mu_i, \boldsymbol{\iota}_T'\mathbf{X}_i) = n^{-1/4}\boldsymbol{\iota}_T'\Omega = 0$. We may conclude that our test T_μ will be asymptotically more powerful than the ANOVA F test when the individual effect is correlated with covariates, see the simulation results in Section 5 for more evidences. Note that, unlike T_μ^*, the test statistic T_μ does not follow an F distribution up to an affine transformation.

Many efforts recently have been spent on the heteroscedasticity of individual and time effects, i.e. the variances of $\{\mu_i\}$ and $\{\eta_t\}$ are all different, see Baltagi et al. (2006, 2010) and Montes-Rojas and Sosa-Escudero (2011). Note that the heteroscedasticity of $\{\eta_t\}$ has no effect on the results in this section. For the heteroscedasticity of $\{\mu_i\}$, the null hypothesis can be set to

$$H_0^\mu : \mathrm{var}(\mu_1) = \cdots = \mathrm{var}(\mu_n) = 0,$$

and Assumptions (A1) and (A2) are replaced by a new one as follows.

(A3) Individual effects $\{\mu_i\}$ are independent random variables with $E(\mu_i) = 0$, $\sigma_1^2 = \lim_{n\to\infty} n^{-1/2}\sum_{i=1}^n \mathrm{var}(\mu_i) < \infty$, $E(\mu_i\mathbf{u}_i) = 0$, $n^{-1/2}\sum_{i=1}^n E(\mu_i^2\|\mathbf{u}_i\|^2) = O(1)$ and $n^{-3/4}\sum_{i=1}^n E(\mu_i\mathbf{X}_i) = O(1)$.

Corollary 2. *Suppose that* $E(u_{it}^4) < \infty$, $E\|\mathbf{X}_i\|^4 < \infty$, $|\Sigma_1| > 0$ *and* $cov(\mathbf{X}_i'\boldsymbol{\iota}_T, \mathbf{u}_i'\boldsymbol{\iota}_T) = 0$. *If Assumption* (A3) *holds, then*

$$T_\mu \rightarrow_d \sigma_1^2[0.5T(T-1)]^{1/2}/\sigma_u^2 + \zeta_1$$

as $n \to \infty$, *where* ζ_1 *is defined as in Theorem 1.*

When the null hypothesis H_0^μ is accepted, we may employ model (3) to estimate the coefficients β, and it results in $\widetilde{\beta}_1$. It can be shown that the estimator $\widetilde{\beta}_1$ is asymptotically as efficient as the OLS estimator of model (2) with the absence of both individual and time effects.

3. Test for the time effect

Suppose that $\{\eta_t\}$ are random variables with mean zero and finite variance. To take into account the heteroscedasticity of $\{\eta_t\}$, the test for the time effect can be formalized as follows,

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

Similar to the case for the individual effect, we first wipe out the individual effect μ_i by applying the orthogonal transformation on model (2) directly, and it results in

$$Q'\mathbf{y}_i = Q'\mathbf{X}_i\beta + Q'\boldsymbol{\eta} + Q'\mathbf{u}_i \quad (8)$$

and $\boldsymbol{\iota}_T'\mathbf{y}_i = T\alpha + \boldsymbol{\iota}_T'\mathbf{X}_i\beta + T\mu_i + \boldsymbol{\iota}_T'\boldsymbol{\eta} + \boldsymbol{\iota}_T'\mathbf{u}_i$, where $i = 1, \ldots, n$. Under the null hypothesis H_0^η, model (8) reduces to $Q'\mathbf{y}_i = Q'\mathbf{X}_i\beta + Q'\mathbf{u}_i$, and we can alternatively estimate the variance of the idiosyncratic error u_{it} by

$$\widehat{\sigma}_{2u}^2 = \frac{1}{T-1}\cdot\frac{1}{n}\sum_{i=1}^n \|Q'(\mathbf{y}_i - \mathbf{X}_i\widehat{\beta})\|^2 = \frac{1}{T-1}\cdot\frac{1}{n}\sum_{i=1}^n (\mathbf{y}_i - \mathbf{X}_i\widehat{\beta})'P_{\iota_T}^{\perp}(\mathbf{y}_i - \mathbf{X}_i\widehat{\beta})$$

since $E\|Q'\mathbf{u}_i\|^2 = (T-1)\sigma_u^2$, and $\widehat{\beta}$ is given as in (6). Notice that, under H_1^η, $\widehat{\sigma}_{2u}^2$ is inconsistent, however, $\widehat{\sigma}_{0u}^2$ in (7) is consistent.

Based on the difference between these two estimators, we can construct a test statistic as follows,

$$T_\eta = \frac{T-1}{\widehat{\sigma}_{0u}^2} \cdot n(\widehat{\sigma}_{2u}^2 - \widehat{\sigma}_{0u}^2) + (T-1).$$

To study the asymptotic behavior of T_η, we give the condition on $\boldsymbol{\eta}$ as follows.

(A4) $\boldsymbol{\eta} = n^{-1/2}\boldsymbol{\zeta}^*$, where $\boldsymbol{\zeta}^*$ a T-dimensional random vector with mean zero and $\sigma_2^2 = E\|\boldsymbol{\zeta}^*\|^2 < \infty$.

The case with $\sigma_2^2 = 0$ corresponds to the null hypothesis H_0^η, and random vector $\boldsymbol{\zeta}^*$ may be correlated with covariates and even the idiosyncratic error. Let $\Sigma_4 = Q'E[\mathbf{u}_i\mathbf{u}_i'P_{\iota_T}^{\perp}(\mathbf{X}_i - E\mathbf{X}_i)]\Sigma_1^{-1}(E\mathbf{X}_i')Q$, and $\Sigma_5 = \Sigma_4 + \Sigma_4' - Q'E(\mathbf{X}_i)\Sigma_1^{-1}\Sigma_2\Sigma_1^{-1}E(\mathbf{X}'_i)Q$.

Theorem 2. *Suppose that $E(u_{it}^2) < \infty, E\|\mathbf{X}_i\|^2 < \infty, |\Sigma_1| > 0$ and $cov(\mathbf{X}_i'\boldsymbol{\iota}_T, \mathbf{u}_i'\boldsymbol{\iota}_T) = 0$. If Assumption (A4) holds, then*

$$T_\eta \to_d \|\sigma_u^{-1}Q'\boldsymbol{\zeta}^* + \boldsymbol{\zeta}_2\|^2$$

as $n \to \infty$, where $\boldsymbol{\zeta}_2$ is a $(T-1)$-dimensional normal random vector with mean zero and variance matrix $I_{T-1} - \sigma_u^{-2}\Sigma_5$.

If $\{\mathbf{u}_i\}$ is further assumed to be independent of $\{\mathbf{X}_i\}$, then the random vector $\boldsymbol{\zeta}_2$ is independent of the random variable ζ_1, which is defined as in Theorem 1.

The value of Q is involved in the asymptotic distribution of test T_η. When $Q'E(\mathbf{X}_i) = 0$, i.e. $E(X_{it})$ is independent of t, it holds that $\Sigma_5 = 0$, and the asymptotic distribution of T_η under the null hypothesis H_0^η is just the chi-square distribution with $T-1$ degrees of freedom, χ_{T-1}^2. In real applications, we may first center the covariates $\mathbf{X}_i$, resulting in $\widetilde{\mathbf{X}}_i$ of (3), and then perform the test T_η with p-values or critical values calculated from χ_{T-1}^2.

Similar to the case of T_μ^* in the previous section, we may consider another test statistic for the time effect,

$$T_\eta^* = \frac{T-1}{\widehat{\sigma}_{0u}^2} \cdot n(\widetilde{\sigma}_{2u}^2 - \widehat{\sigma}_{0u}^2) + (T-1) = \frac{T-1}{(n-1)(T-1)-p}F_\eta,$$

where

$$\widetilde{\sigma}_{2u}^2 = \frac{1}{T-1}\cdot\frac{1}{n}\sum_{i=1}^n \|Q'(\mathbf{y}_i - \mathbf{X}_i\widetilde{\beta}_2)\|^2,$$

$$\widetilde{\beta}_2 = \underset{\beta}{\text{argmin}} \sum_{i=1}^n \|Q'(\mathbf{y}_i - \mathbf{X}_i\beta)\|^2,$$

and

$$F_\eta = \frac{\sum_{i=1}^n\{\|Q'(\mathbf{y}_i - \mathbf{X}_i\widetilde{\beta}_2)\|^2 - \|Q'(\widetilde{\mathbf{y}}_i - \widetilde{X}_i\widehat{\beta})\|^2\}/(T-1)}{\sum_{i=1}^n \|Q'(\widetilde{\mathbf{y}}_i - \widetilde{\mathbf{X}}_i\widehat{\beta})\|^2/[(n-1)(T-1)-p]}.$$

Under H_0^η, the ANOVA F test statistic F_η follows the F distribution with $T-1$ and $(n-1)(T-1)-p$ degrees of freedom when the idiosyncratic error u_{it} is normally distributed, see Baltagi (2008).

Corollary 3. *Suppose that $E(u_{it}^2) < \infty, E\|\mathbf{X}_i\|^2 < \infty, |\Sigma_1| > 0$, $cov(\mathbf{X}_i'\boldsymbol{\iota}_T, \mathbf{u}_i'\boldsymbol{\iota}_T) = 0$ and $Q'E(\mathbf{X}_i) = 0$. If Assumption (A4) holds, then*

$$T_\eta^* \to_d \|\sigma_u^{-1}Q'\boldsymbol{\zeta}^* + \boldsymbol{\zeta}_2\|^2$$

as $n \to \infty$, where $\boldsymbol{\zeta}_2$ is defined as in Theorem 2.

From the above corollary, unlike the counterpart in the previous section, the test T_η is asymptotically not more powerful than, the ANOVA F test.

When the time effect is fixed, i.e. $\eta_1, \ldots, \eta_T$ are non-random with $\sum_{t=1}^T \eta_t = 0$, we assume that $\boldsymbol{\zeta}^*$ is a constant vector, and then Theorem 2 still holds. When the null hypothesis of H_0^η is accepted, we can consider the feasible GLS estimator of the coefficients β, and it is asymptotically efficient regardless of the presence of the individual effect, see Baltagi (2008).

4. Test jointly for both individual and time effects

Besides these tests in the previous two sections, we sometimes are interested in testing for the presence of individual and time effects jointly. The hypotheses can be formalized as follows,

$H_0^{\mu\eta}: \sigma_\mu^2 = \text{var}(\eta_1) = \cdots = \text{var}(\eta_T) = 0 \quad \text{vs}$

$H_1^{\mu\eta}$: at least one of them is nonzero.

Under $H_0^{\mu\eta}$, model (2) reduces to $\mathbf{y}_i = \alpha\boldsymbol{\iota}_T + \mathbf{X}_i\beta + \mathbf{u}_i$, and we may consider an estimator of σ_u^2 as follows,

$$\widehat{\sigma}_{3u}^2 = (nT)^{-1}\sum_{i=1}^n \|\mathbf{y}_i - \widehat{\alpha}\boldsymbol{\iota}_T - \mathbf{X}_i\widehat{\beta}\|^2,$$

where $E\|\mathbf{u}_i\|^2 = T\sigma_u^2$, $\widehat{\beta}$ is given as in (6), and $\widehat{\alpha} = (nT)^{-1}\sum_{i=1}^n \boldsymbol{\iota}_T'(\mathbf{y}_i - \mathbf{X}_i\widehat{\beta})$. Note that $\widehat{\alpha}$ and $\widehat{\sigma}_{3u}^2$ are consistent only under the null hypothesis $H_0^{\mu\eta}$. Similarly, we can construct a Hausman-type test by comparing $\widehat{\sigma}_{3u}^2$ with $\widehat{\sigma}_{0u}^2$ in (7), and it results in

$$T_{\mu\eta 1} = [0.5T(T-1)]^{1/2}\sqrt{n}\left(\frac{\widehat{\sigma}_{3u}^2}{\widehat{\sigma}_{0u}^2} - 1\right).$$

Let $(\widetilde{\alpha}_3, \widetilde{\beta}_3) = \text{argmin}_{\alpha,\beta}\sum_{i=1}^n \|\mathbf{y}_i - \alpha\boldsymbol{\iota}_T - \mathbf{X}_i\beta\|^2$, and $\widetilde{\sigma}_{3u}^2 = (nT)^{-1}\sum_{i=1}^n \|\mathbf{y}_i - \widetilde{\alpha}_3\boldsymbol{\iota}_T - \mathbf{X}_i\widetilde{\beta}_3\|^2$. As in the previous two sections, we may consider another test statistic

$$T_{\mu\eta 1}^* = [0.5T(T-1)]^{1/2}\sqrt{n}\left(\frac{\widetilde{\sigma}_{3u}^2}{\widehat{\sigma}_{0u}^2} - 1\right) = c_{2n}\cdot F_{\mu\eta} - d_{2n},$$

where $c_{2n} = [0.5nT(T-1)]^{1/2}(1-T^{-1})(1-n^{-1})(n+T-2)/[(n-1)(T-1)-p]$, $d_{2n} = [0.5nT(T-1)]^{1/2}(T+n-1)/(Tn)$ and

$$F_{\mu\eta} = \frac{\sum_{i=1}^n\{\|\mathbf{y}_i - \widetilde{\alpha}\boldsymbol{\iota}_T - \mathbf{X}_i\widetilde{\beta}\|^2 - \|Q'(\widetilde{\mathbf{y}}_i - \widetilde{X}_i\widehat{\beta})\|^2\}/(n+T-2)}{\sum_{i=1}^n \|Q'(\widetilde{\mathbf{y}}_i - \widetilde{\mathbf{X}}_i\widehat{\beta})\|^2/[(n-1)(T-1)-p]}.$$

When the idiosyncratic error u_{it} is normally distributed, the ANOVA F test statistic $F_{\mu\eta}$ follows the F distribution with $n+T-2$ and $(n-1)(T-1)-p$ degrees of freedom under $H_0^{\mu\eta}$, see Baltagi (2008). Let $\Sigma_6 = \Sigma_3 + E(\mathbf{X}_i')P_{\iota_T}^{\perp}E(\mathbf{X}_i)$ and

$$\delta = T^{-2}[\boldsymbol{\iota}_T'E(\mathbf{X}_i)\Sigma_6^{-1}\Omega'\boldsymbol{\iota}_T]^2 + T^{-1}\boldsymbol{\iota}_T'\Omega\Sigma_6^{-1}[2\Sigma_6 - E(\mathbf{X}_i'\mathbf{X}_i)]\Sigma_6^{-1}\Omega'\boldsymbol{\iota}_T.$$

Theorem 3. *Suppose that $E(u_{it}^4) < \infty, E\|\mathbf{X}_i\|^4 < \infty, |\Sigma_1| > 0$, $|\Sigma_6| > 0$ and $cov(\mathbf{X}_i'\boldsymbol{\iota}_T, \mathbf{u}_i'\boldsymbol{\iota}_T) = 0$. If Assumptions (A1), (A2) and (A4) hold, then*

$$T_{\mu\eta 1} \to_d \sigma_1^2[0.5T(T-1)]^{1/2}/\sigma_u^2 + \zeta_1$$

and

$$T_{\mu\eta 1}^* \to_d (\sigma_1^2 - \delta)[0.5T(T-1)]^{1/2}/\sigma_u^2 + \zeta_1$$

as $n \to \infty$, where Ω and ζ_1 are defined as in Theorem 1.

For both tests $T_{\mu\eta 1}$ and $T^*_{\mu\eta 1}$, we can refer to the standard normal distribution for critical values and p-values since $\Omega = 0$ and $\delta = 0$ as $\sigma_1^2 = 0$. It holds that $0 \le \delta \le \sigma_1^2$, and quantity δ equals to zero if $\text{cov}(\mu_i, \boldsymbol{\iota}_T'\mathbf{X}_i) = n^{-1/4}\boldsymbol{\iota}_T'\Omega = 0$.

Moreover, these two tests in Theorem 3 fail asymptotically to detect the presence of the time effect in Assumption (A4), i.e. they suffer asymptotic loss in the power comparing with the test for the time effect T_η in the previous section. Note that, from Theorems 1 and 2, $\sqrt{n}(\widehat{\sigma}_{1u}^2 - \widehat{\sigma}_{0u}^2) = O_p(1)$ and $n(\widehat{\sigma}_{2u}^2 - \widehat{\sigma}_{0u}^2) = O_p(1)$. We may argue that the test for the time effect can detect smaller departure from the null hypothesis than that for the individual effect, and such sensitivity will be masked in the testing for the null hypothesis $H_0^{\mu\eta}$.

From Theorem 2, test statistics T_μ and T_η are asymptotically independent if $\{\mathbf{u}_i\}$ is independent of $\{\mathbf{X}_i\}$, and then we may achieve a more powerful test for $H_0^{\mu\eta}$ by combining these two tests. A direct way is to employ the Bonferroni method to give the rejection region,

$$\{T_\mu > z_{\alpha/2} \text{ or } T_\eta > \chi^2_{\alpha/2,T-1}\},$$

where α is a predetermined significance level, and $z_{\alpha/2}$ and $\chi^2_{\alpha/2,T-1}$ are respectively the $100(1-\alpha/2)$th percentiles of the standard normal distribution and the χ^2_{T-1} distribution.

We next introduce a test with test statistic

$$T_{\mu\eta 2} = \omega T_\mu^2 + (1-\omega)T_\eta,$$

where the weight $\omega \in [0, 1]$ can be specified by practitioners. From Theorems 1 and 2, we have that

$$T_{\mu\eta 2} \to_d \omega|\sigma_1^2[0.5T(T-1)]^{1/2}/\sigma_u^2 + \zeta_1|^2 + (1-\omega)\|\sigma_u^{-1}Q'\boldsymbol{\zeta}^* + \boldsymbol{\zeta}_2\|^2$$

as $n \to \infty$. Under $H_0^{\mu\eta}$, if $Q'E(\mathbf{X}_i) = 0$ and $\{\mathbf{u}_i\}$ is independent of $\{\mathbf{X}_i\}$, then

$$T_{\mu\eta 2} \to_d \omega\chi_1^2 + (1-\omega)\chi^2_{T-1} \tag{9}$$

as $n \to \infty$. Sometimes we may have no preference about the weight in $T_{\mu\eta 2}$. It can be simply set to $\omega = 0.5$, and the asymptotic distribution of $T_{\mu\eta 2}$ under $H_0^{\mu\eta}$ is just $0.5\chi_T^2$. When $\omega = 1$, the test statistic $T_{\mu\eta 2}$ will reduce to T_μ^2, which corresponds to the two-sided test of T_μ and is then less powerful. Actually, there are still many other choices besides $T_{\mu\eta 2}$, however, their asymptotic distribution under the null hypothesis may not be as simple as the mixed chi-square distribution in (9).

When the null hypothesis $H_0^{\mu\eta}$ is accepted, the commonly used OLS estimator of the coefficients β will be efficient. Otherwise, we have to consider its estimation based on the general two-way error component model (2), and one can be referred to Chapter 3 of Baltagi (2008).

5. Simulation studies

We conduct four simulation experiments in this section to study the finite-sample performance of the proposed tests and the estimators for higher order moments of the idiosyncratic error and the individual effect. All simulation results are based on 1000 replications and the significance level is set to 0.05.

The first experiment is for the test T_μ. Six currently used tests for the individual effect are also calculated for the sake of comparison, and they are a test in Breusch and Pagan (1980) (hence BP$_\mu$ test), a test in Honda (1985) (H$_\mu$), a modified two-sided test (BSY1) and a modified one-sided test (BSY2) in Bera et al. (2001), a test robust to the time effect in Baltagi et al. (1992) (BCL$_\mu$) and the ANOVA F test (F_μ). The abbreviations are similar for the other two types of tests in this section. The data generating process is

$$y_{it} = 0.5 + X_{it}^{(1)} + 2X_{it}^{(2)} + \mu_i + \eta_t + u_{it}, \tag{10}$$

where $X_{it}^{(1)}, X_{it}^{(2)}, \mu_i$ and η_t follow the normal distributions with mean zero, $\text{var}(X_{it}^{(1)}) = \text{var}(X_{it}^{(2)}) = 1$, $\text{var}(\mu_i) = \sigma_\mu^2$, $\text{var}(\eta_t) = \sigma_\eta^2$, $\text{corr}(X_{it}^{(1)}, \mu_i) = \rho$, and u_{it} follows the standard normal distribution, $N(0, 1)$, or $\sqrt{0.5}(\chi_1^2 - 1)$. When $\rho \ne 0$, $X_{it}^{(1)}$ and $X_{is}^{(1)}$ are also correlated for $t \ne s$, and then we further set $\text{corr}(X_{it}^{(1)}, X_{is}^{(1)}) = \rho^2$. Note that $\sigma_\mu = 0$ or >0 corresponds respectively to the size or the power, and $\sigma_\eta = 0$ or >0 to the absence or the presence of the time effect. Let $\mathbf{X}_i^{(1)} = (X_{i1}^{(1)}, \ldots, X_{iT}^{(1)})'$. Sequences $\{\mathbf{X}_i^{(1)}\}$, $\{X_{it}^{(2)}\}$, $\{\mu_i\}$, $\{\eta_t\}$ and $\{u_{it}\}$ are set to be *i.i.d.*, and are independent of each other except for $\{\mathbf{X}_i^{(1)}\}$ and $\{\mu_i\}$ with $\rho \ne 0$. We check the performance of the test T_μ under three situations: (i) the standard setting, i.e. $\rho = 0$ and u_{it} follows $N(0, 1)$, (ii) the non-normal setting, i.e. $\rho = 0$ and u_{it} follows $\sqrt{0.5}(\chi_1^2 - 1)$, and (iii) the correlated setting, i.e. $\rho > 0$ and u_{it} follows $N(0, 1)$.

For the standard setting, we set the sample sizes $(n, T) = (100, 5), (100, 10)$ or $(200, 10)$, and Table 1 lists the empirical sizes and powers of the test T_μ and other six tests for the individual effect. When the time effect is present, the sizes of four tests BP$_\mu$, BSY1, H$_\mu$ and BSY2 are all distorted, where the first two are too sensitive while the last two are too conservative. This is consistent with our expectation since these four tests are all based on the one-way error component model. The test BCL$_\mu$ still has acceptable sizes even when the number of time points is as small as $T = 5$. Roughly speaking, tests BCL$_\mu$, F_μ and T_μ have comparable powers. Specifically, our test T_μ outperforms BCL$_\mu$ and F_μ. The test F_μ is more powerful than BCL$_\mu$ when the time effect is absent, however, it is less powerful when the time effect is presented. Table 2 gives the empirical sizes and powers under the non-normal setting, where u_{it} follows the non-normal distribution. It can be seen that our test T_μ is most powerful as in the standard setting, and both BCL$_\mu$ and F_μ are robust to the non-normality. We also tried the Student's t distribution with five degrees of freedom, t_5, for u_{it}, and the results are similar. For the correlated setting, we consider three values for the correlation coefficient of $X_{it}^{(1)}$ and μ_i, $\rho = 0.25, 0.5$ and 0.75, and the sample sizes are $(n, T) = (200, 10)$. Table 3 gives the empirical powers. When the value of ρ increases, the powers of tests BCL$_\mu$ and F_μ decrease substantially, however, those of our test T_μ are not affected. Note that $\rho = \text{corr}(X_{it}^{(1)}, \mu_i) = 0$ as $\text{var}(\mu_i) = \sigma_\mu^2 = 0$, i.e. the correlated setting will reduce to the standard setting in evaluating the sizes.

The second experiment is to compare the proposed test T_η with four currently used tests for the time effect, BP$_\eta$, H$_\eta$, BCL$_\eta$ and F_η. We employ the data generating process (10) with $\rho = 0$, and the sample sizes are set to $(n, T) = (100, 5)$. It is noteworthy to point out that $\sigma_\mu = 0$ or >0 corresponds to the absence or the presence of the individual effect, and $\sigma_\eta = 0$ or >0 to the size or the power. Three distributions are considered for the idiosyncratic error u_{it}: $N(0, 1)$, $\sqrt{0.5}(\chi_1^2 - 1)$ and $\sqrt{0.6}t_5$. Table 4 lists the empirical sizes and powers of these five tests for the time effect. Note that BP$_\eta$ and H$_\eta$ are based on the one-way error component model, and the distortion of their sizes is observed again. The three tests BCL$_\eta$, F_η and T_η have comparable powers, and the powers of our test T_η are all greater than those of F_η although the difference is slight. We also considered the case with the time effect and covariates correlated, and the findings are similar to the correlated setting in the first experiment.

The third experiment is to compare the proposed tests $T_{\mu\eta 1}$ and $T_{\mu\eta 2}$ with two currently used joint tests for both individual and time effects, BP$_{\mu\eta}$ and $F_{\mu\eta}$. The data generating process (10) is used again, and we consider the standard case, $\rho = 0$, and the case

Table 1
Empirical sizes and powers of test T_μ and other six tests for the individual effect under the standard setting. The nominal rate is 5%.

σ_η	σ_μ	BP_μ	BSY1	H_μ	BSY2	BCL_μ	F_μ	T_μ
		$(n, T) = (100, 5)$						
0.0	0.0	0.040	0.040	0.041	0.037	0.033	0.044	0.066
	0.1	0.050	0.045	0.070	0.065	0.052	0.070	0.111
	0.2	0.241	0.173	0.315	0.243	0.281	0.322	0.379
0.5	0.0	0.273	0.257	0.003	0.020	0.040	0.039	0.056
	0.1	0.237	0.218	0.011	0.038	0.091	0.086	0.123
	0.2	0.130	0.187	0.072	0.096	0.332	0.318	0.388
1.0	0.0	0.804	0.552	0.000	0.047	0.035	0.034	0.046
	0.1	0.760	0.503	0.001	0.062	0.107	0.096	0.140
	0.2	0.629	0.497	0.010	0.099	0.354	0.332	0.399
		$(n, T) = (100, 10)$						
0.0	0.0	0.046	0.052	0.045	0.039	0.042	0.046	0.060
	0.1	0.116	0.106	0.175	0.165	0.168	0.181	0.212
	0.2	0.665	0.588	0.743	0.683	0.726	0.747	0.784
0.5	0.0	0.285	0.276	0.002	0.008	0.047	0.044	0.056
	0.1	0.143	0.197	0.017	0.038	0.172	0.160	0.191
	0.2	0.185	0.220	0.233	0.250	0.751	0.735	0.774
1.0	0.0	0.917	0.663	0.000	0.006	0.045	0.041	0.057
	0.1	0.829	0.595	0.000	0.011	0.193	0.175	0.209
	0.2	0.513	0.474	0.012	0.067	0.752	0.737	0.770
		$(n, T) = (200, 10)$						
0.0	0.0	0.048	0.041	0.054	0.053	0.055	0.055	0.058
	0.1	0.175	0.154	0.245	0.223	0.236	0.248	0.287
	0.2	0.913	0.854	0.946	0.907	0.945	0.947	0.958
0.5	0.0	0.500	0.399	0.001	0.009	0.049	0.044	0.053
	0.1	0.262	0.284	0.014	0.036	0.241	0.229	0.261
	0.2	0.286	0.383	0.365	0.393	0.937	0.931	0.943
1.0	0.0	0.976	0.750	0.000	0.012	0.051	0.047	0.058
	0.1	0.927	0.704	0.000	0.025	0.254	0.240	0.270
	0.2	0.695	0.597	0.010	0.112	0.942	0.938	0.947

Table 2
Empirical sizes and powers of test T_μ and other six tests for the individual effect under the non-normal setting. The nominal rate is 5%.

σ_η	σ_μ	BP_μ	BSY1	H_μ	BSY2	BCL_μ	F_μ	T_μ
		$(n, T) = (100, 5)$						
0.0	0.0	0.044	0.037	0.045	0.051	0.036	0.046	0.064
	0.1	0.061	0.062	0.090	0.085	0.068	0.098	0.122
	0.2	0.206	0.150	0.296	0.216	0.253	0.304	0.373
0.5	0.0	0.316	0.247	0.009	0.025	0.048	0.045	0.061
	0.1	0.215	0.217	0.016	0.044	0.106	0.100	0.134
	0.2	0.120	0.208	0.054	0.110	0.299	0.282	0.365
1.0	0.0	0.795	0.579	0.002	0.043	0.051	0.048	0.063
	0.1	0.754	0.538	0.001	0.062	0.103	0.091	0.130
	0.2	0.630	0.506	0.009	0.084	0.347	0.334	0.390
		$(n, T) = (100, 10)$						
0.0	0.0	0.045	0.042	0.046	0.044	0.045	0.048	0.061
	0.1	0.106	0.104	0.158	0.155	0.144	0.158	0.192
	0.2	0.622	0.545	0.721	0.648	0.706	0.726	0.764
0.5	0.0	0.278	0.260	0.003	0.011	0.054	0.045	0.062
	0.1	0.152	0.197	0.022	0.034	0.164	0.159	0.186
	0.2	0.176	0.226	0.227	0.242	0.768	0.744	0.783
1.0	0.0	0.912	0.664	0.000	0.007	0.045	0.040	0.054
	0.1	0.831	0.631	0.000	0.021	0.161	0.146	0.176
	0.2	0.516	0.474	0.015	0.075	0.785	0.769	0.808
		$(n, T) = (200, 10)$						
0.0	0.0	0.044	0.047	0.033	0.032	0.031	0.035	0.044
	0.1	0.161	0.142	0.235	0.215	0.226	0.232	0.264
	0.2	0.904	0.834	0.933	0.900	0.931	0.934	0.946
0.5	0.0	0.489	0.420	0.003	0.012	0.054	0.050	0.059
	0.1	0.274	0.294	0.013	0.038	0.243	0.230	0.257
	0.2	0.304	0.403	0.393	0.427	0.955	0.952	0.956
1.0	0.0	0.972	0.758	0.000	0.012	0.049	0.045	0.055
	0.1	0.920	0.729	0.000	0.021	0.264	0.248	0.283
	0.2	0.685	0.578	0.013	0.084	0.940	0.938	0.949

Table 3
Empirical powers of test T_μ and other six tests for the individual effect under the correlated setting. The nominal rate is 5%, and $(n, T) = (200, 10)$.

σ_μ	ρ	BP_μ	BSY1	H_μ	BSY2	BCL_μ	F_μ	T_μ
		$\sigma_\eta = 0.0$						
0.1	0.25	0.119	0.099	0.180	0.164	0.180	0.202	0.256
	0.50	0.070	0.062	0.124	0.102	0.120	0.175	0.275
	0.75	0.054	0.047	0.066	0.063	0.064	0.111	0.291
0.2	0.25	0.825	0.749	0.882	0.832	0.876	0.910	0.940
	0.50	0.526	0.447	0.618	0.557	0.606	0.802	0.931
	0.75	0.104	0.088	0.170	0.134	0.155	0.464	0.917
		$\sigma_\eta = 0.5$						
0.1	0.25	0.272	0.282	0.007	0.024	0.215	0.227	0.276
	0.50	0.343	0.320	0.004	0.018	0.133	0.167	0.265
	0.75	0.432	0.372	0.001	0.010	0.070	0.100	0.284
0.2	0.25	0.203	0.323	0.242	0.304	0.900	0.920	0.949
	0.50	0.132	0.219	0.085	0.138	0.654	0.816	0.944
	0.75	0.294	0.293	0.008	0.021	0.188	0.471	0.918
		$\sigma_\eta = 1.0$						
0.1	0.25	0.938	0.689	0.000	0.024	0.221	0.231	0.280
	0.50	0.948	0.745	0.000	0.020	0.133	0.169	0.281
	0.75	0.970	0.753	0.000	0.013	0.082	0.118	0.288
0.2	0.25	0.709	0.586	0.004	0.094	0.891	0.912	0.950
	0.50	0.825	0.675	0.000	0.048	0.625	0.783	0.940
	0.75	0.937	0.719	0.000	0.019	0.171	0.466	0.904

Table 4
Empirical sizes and powers of test T_η and other four tests for the time effect. The nominal rate is 5%, and $(n, T) = (100, 5)$.

σ_μ	σ_η	BP_η	H_η	BCL_η	F_η	T_η
		$u_{it} \sim N(0, 1)$				
0.0	0.0	0.033	0.041	0.041	0.049	0.052
	0.1	0.225	0.275	0.270	0.312	0.330
	0.2	0.680	0.716	0.723	0.745	0.757
0.5	0.0	0.012	0.014	0.052	0.046	0.048
	0.1	0.146	0.179	0.324	0.303	0.313
	0.2	0.577	0.621	0.763	0.756	0.759
1.0	0.0	0.000	0.000	0.056	0.044	0.046
	0.1	0.029	0.041	0.329	0.299	0.303
	0.2	0.365	0.410	0.753	0.732	0.736
		$u_{it} \sim \sqrt{0.5}(\chi_1^2 - 1)$				
0.0	0.0	0.018	0.026	0.028	0.039	0.043
	0.1	0.230	0.277	0.276	0.320	0.332
	0.2	0.669	0.708	0.697	0.732	0.741
0.5	0.0	0.007	0.013	0.058	0.051	0.053
	0.1	0.137	0.171	0.339	0.316	0.331
	0.2	0.589	0.629	0.762	0.751	0.756
1.0	0.0	0.001	0.001	0.063	0.053	0.056
	0.1	0.025	0.046	0.366	0.338	0.342
	0.2	0.359	0.398	0.785	0.768	0.771
		$u_{it} \sim \sqrt{0.6}t_5$				
0.0	0.0	0.025	0.039	0.034	0.052	0.054
	0.1	0.241	0.285	0.287	0.315	0.324
	0.2	0.677	0.720	0.709	0.736	0.742
0.5	0.0	0.005	0.012	0.054	0.052	0.053
	0.1	0.148	0.189	0.333	0.314	0.330
	0.2	0.577	0.636	0.765	0.755	0.762
1.0	0.0	0.000	0.000	0.062	0.046	0.048
	0.1	0.026	0.046	0.356	0.324	0.332
	0.2	0.351	0.414	0.798	0.770	0.778

with correlation, $\rho = 0.5$. The sample sizes are set to $(n, T) = (100, 5)$, $(100, 10)$ or $(200, 10)$, and we fix $\omega = 0.5$ for simplicity. Table 5 gives the empirical sizes and powers of these four tests. Note that the correlation happens only between the individual effect and the covariates, and then there is no correlation when $\sigma_\mu = 0$. The sizes of all four tests are close to the nominal value 0.05, and our tests $T_{\mu\eta 1}$ and $T_{\mu\eta 2}$ are both robust to the possible presence of correlation. Test $T_{\mu\eta 1}$ is more powerful than $T_{\mu\eta 2}$ as

Table 5
Empirical sizes and powers of tests $T_{\mu\eta 1}$, $T_{\mu\eta 2}$ and other two joint tests for both effects. The nominal rate is 5%.

σ_η	σ_μ	$\rho = 0.0$				$\rho = 0.5$			
		$BP_{\mu\eta}$	$F_{\mu\eta}$	$T_{\mu\eta 1}$	$T_{\mu\eta 2}$	$BP_{\mu\eta}$	$F_{\mu\eta}$	$T_{\mu\eta 1}$	$T_{\mu\eta 2}$
		$(n, T) = (100, 5)$							
0.0	0.0	0.032	0.052	0.063	0.049				
	0.1	0.049	0.102	0.131	0.081	0.039	0.080	0.129	0.078
	0.2	0.171	0.306	0.375	0.211	0.074	0.226	0.376	0.220
0.1	0.0	0.191	0.072	0.097	0.282				
	0.1	0.180	0.123	0.169	0.247	0.220	0.132	0.197	0.325
	0.2	0.292	0.390	0.446	0.401	0.219	0.318	0.452	0.442
0.2	0.0	0.657	0.265	0.299	0.729				
	0.1	0.632	0.334	0.398	0.713	0.676	0.350	0.427	0.746
	0.2	0.676	0.600	0.649	0.796	0.643	0.505	0.646	0.780
		$(n, T) = (100, 10)$							
0.0	0.0	0.026	0.034	0.048	0.049				
	0.1	0.089	0.160	0.188	0.084	0.049	0.119	0.194	0.102
	0.2	0.561	0.706	0.742	0.423	0.226	0.539	0.750	0.431
0.1	0.0	0.334	0.141	0.163	0.465				
	0.1	0.372	0.307	0.345	0.462	0.376	0.267	0.358	0.522
	0.2	0.705	0.815	0.843	0.760	0.459	0.694	0.841	0.730
0.2	0.0	0.906	0.619	0.659	0.935				
	0.1	0.911	0.741	0.771	0.953	0.918	0.709	0.764	0.939
	0.2	0.958	0.965	0.972	0.974	0.930	0.930	0.970	0.978
		$(n, T) = (200, 10)$							
0.0	0.0	0.037	0.047	0.055	0.044				
	0.1	0.126	0.239	0.269	0.094	0.066	0.159	0.267	0.105
	0.2	0.866	0.940	0.953	0.723	0.407	0.778	0.917	0.678
0.1	0.0	0.672	0.213	0.234	0.775				
	0.1	0.713	0.535	0.564	0.816	0.677	0.419	0.533	0.816
	0.2	0.956	0.970	0.977	0.970	0.808	0.910	0.972	0.950
0.2	0.0	0.991	0.786	0.806	0.996				
	0.1	0.982	0.882	0.891	0.985	0.992	0.891	0.920	0.996
	0.2	0.999	1.000	1.000	0.998	0.990	0.992	0.999	0.999

$\sigma_\eta = 0$, however, is less powerful as $\sigma_\eta > 0$. Note that, when $\sigma_\eta = 0$, only the term T_μ^2 in the test statistic $T_{\mu\eta 2}$ can provide the power. It is equivalent to a two-sided test of T_μ, and is less powerful.

The fourth experiment is to study the estimators for higher order moments of u_{it} and μ_i in the Appendix. The data generating process is given as in (10) with $\rho = 0$. We consider four pairs of different distributions for u_{it} and μ_i: (i) $N(0, 1)$ and $N(0, 1)$, (ii) $\sqrt{0.75}t_8$ and $N(0, 1)$, (iii) $\sqrt{0.75}t_8$ and $\sqrt{0.75}t_8$, and (iv) $N(0, 1)$ and $\sqrt{0.5}(\chi_1^2 - 1)$, where t_8 is the Student's t distribution with eight degrees of freedom. The sample sizes are set to $(n, T) = (50, 10)$ or $(100, 10)$, and the corresponding estimators in Wu and Su (2010) are also calculated for the sake of comparison. The empirical means and empirical standard deviations are presented in Tables 6 and 7 respectively for the cases with $n = 50$ and 100. Note that the method in Wu and Su (2010) cannot be used to construct the estimators of third order moments. It can be seen that our estimators are slightly better.

Table 6
Empirical means and empirical standard deviations (in the parenthesis) of the estimators of higher order moments by the methods in the Appendix (WL) and in Wu and Su (2010) (WS) for four pairs of different distributions of u_{it} and μ_i with $(n, T) = (50, 10)$.

	True value	$\sigma_\eta = 0$		$\sigma_\eta = 1$	
		WL	WS	WL	WS
		Pair (i)			
γ_2^u	1.0000	1.0003	1.0012	1.0032	1.0030
		(0.0674)	(0.0989)	(0.0679)	(0.0964)
γ_3^u	0.0000	0.0005		−0.0077	
		(0.1902)		(0.1916)	
γ_4^u	3.0000	3.0008	2.9291	3.0112	2.9787
		(0.5383)	(1.1328)	(0.5394)	(1.1322)
γ_2^μ	1.0000	0.9978	0.9958	0.9922	0.9836
		(0.2182)	(0.3124)	(0.2318)	(0.3100)
γ_3^μ	0.0000	−0.0154		−0.0232	
		(0.4240)		(0.4350)	
γ_4^μ	3.0000	2.9951	2.7931	3.0049	2.5643
		(1.6088)	(2.9550)	(1.7287)	(2.6596)
		Pair (ii)			
γ_2^u	1.0000	1.0009	0.9993	0.9970	0.9960
		(0.0848)	(0.1066)	(0.0847)	(0.1054)
γ_3^u	0.0000	−0.0144		0.0059	
		(0.3949)		(0.3881)	
γ_4^u	4.5000	4.4068	4.2846	4.3926	4.2791
		(2.3077)	(2.8391)	(2.2102)	(2.5877)
γ_2^μ	1.0000	1.0066	1.0039	0.9820	0.9816
		(0.2258)	(0.3119)	(0.2247)	(0.2998)
γ_3^μ	0.0000	−0.0012		−0.0077	
		(0.4121)		(0.4115)	
γ_4^μ	3.0000	3.0268	2.7744	2.8639	2.5948
		(1.5644)	(2.7649)	(1.5430)	(2.5709)
		Pair (iii)			
γ_2^u	1.0000	0.9968	0.9980	0.9987	0.9941
		(0.0865)	(0.1112)	(0.0848)	(0.1117)
γ_3^u	0.0000	−0.0099		0.0046	
		(0.4726)		(0.3570)	
γ_4^u	4.5000	4.5442	4.5991	4.3503	4.2739
		(4.1050)	(7.8376)	(1.6972)	(2.5514)
γ_2^μ	1.0000	0.9806	0.9890	0.9949	0.9948
		(0.2729)	(0.3440)	(0.2580)	(0.3380)
γ_3^μ	0.0000	0.0150		−0.0061	
		(0.8831)		(0.7924)	
γ_4^μ	4.5000	4.1871	3.8972	4.1150	3.5968
		(5.6770)	(5.9752)	(4.2207)	(4.8422)
		Pair (iv)			
γ_2^u	1.0000	1.0025	0.9968	0.9998	0.9987
		(0.0681)	(0.0955)	(0.0661)	(0.0943)
γ_3^u	0.0000	−0.0110		0.0009	
		(0.1928)		(0.1874)	
γ_4^u	3.0000	3.0089	2.9613	2.9904	2.9375
		(0.5169)	(1.1209)	(0.5537)	(1.0983)
γ_2^μ	1.0000	0.9735	0.9817	0.9846	0.9888
		(0.5352)	(0.5723)	(0.5147)	(0.5500)
γ_3^μ	2.8284	2.7682		2.7114	
		(3.7871)		(3.2512)	
γ_4^μ	15.0000	14.7120	13.9410	13.7476	12.8299
		(41.9907)	(42.1844)	(28.1626)	(26.0643)

6. A real example

Cornwell and Trumbull (1994) considered an economic model of crime for panel data on 90 counties in North Carolina from 1981 to 1987,

$$R_{it} = \alpha + P_{it}'\gamma + X_{it}'\beta + \mu_i + \eta_t + u_{it},$$
$$i = 1, 2, \ldots, n, \ t = 1, 2, \ldots, T,$$

where R_{it} is the crime rate, which is an FBI index measuring the number of crimes divided by the county population, P_{it} includes deterrent variables and X_{it} contains variables measuring returns to legal opportunities, see also Baltagi (2006). There are four deterrent variables: the probability of arrest (P_A), which is measured by the ratio of arrests to offenses, the probability of conviction given arrest (P_C), which is measured by the ratio of convictions to arrests, the probability of a prison sentence given a conviction (P_P), measured by the proportion of total convictions resulting in prison sentences, and the average prison sentence in days (S) as a proxy for sanction severity. The variables in X include the number of police per capita (Police), the population density (Density), which is the county population divided by county land area, the percent young male (Male), which is the proportion of the county's population that is male and aged between 15 and 24, and the average weekly wages in the county from nine industries (Wage1–Wage9). There are sixteen covariates in total, and all variables are in logs.

Table 7
Empirical means and empirical standard deviations (in the parenthesis) of the estimators of higher order moments by the methods in the Appendix (WL) and in Wu and Su (2010) (WS) for four pairs of different distributions of u_{it} and μ_i with $(n, T) = (100, 10)$.

	True value	$\sigma_\eta = 0$		$\sigma_\eta = 1$	
		WL	WS	WL	WS
		Pair (i)			
γ_2^u	1.0000	1.0007 (0.0479)	1.0013 (0.0680)	1.0028 (0.0481)	1.0025 (0.0668)
γ_3^u	0.0000	0.0033 (0.1318)		0.0045 (0.1402)	
γ_4^u	3.0000	3.0028 (0.3783)	2.9931 (0.8336)	3.0138 (0.3884)	2.9736 (0.7779)
γ_2^μ	1.0000	0.9870 (0.1557)	0.9878 (0.2188)	0.9929 (0.1573)	0.9897 (0.2162)
γ_3^μ	0.0000	0.0061 (0.2965)		0.0001 (0.2795)	
γ_4^μ	3.0000	2.9337 (1.1136)	2.7650 (2.0021)	2.9462 (1.1144)	2.7955 (2.0447)
		Pair (ii)			
γ_2^u	1.0000	1.0025 (0.0621)	1.0035 (0.0772)	0.9977 (0.0620)	0.9959 (0.0776)
γ_3^u	0.0000	−0.0073 (0.2718)		0.0056 (0.2913)	
γ_4^u	4.5000	4.4709 (1.5476)	4.4185 (1.8632)	4.4522 (2.1921)	4.4373 (2.6783)
γ_2^μ	1.0000	0.9939 (0.1640)	0.9920 (0.2278)	0.9975 (0.1516)	0.9993 (0.2109)
γ_3^μ	0.0000	0.0061 (0.2858)		−0.0076 (0.2829)	
γ_4^μ	3.0000	2.9882 (1.1581)	2.8336 (2.1707)	2.9666 (1.0746)	2.7896 (1.8683)
		Pair (iii)			
γ_2^u	1.0000	1.0002 (0.0599)	0.9968 (0.0757)	0.9989 (0.0582)	0.9988 (0.0738)
γ_3^u	0.0000	0.0117 (0.2694)		−0.0025 (0.2564)	
γ_4^u	4.5000	4.4540 (1.3509)	4.3594 (1.7810)	4.4246 (1.3491)	4.3382 (1.7784)
γ_2^μ	1.0000	1.0014 (0.1950)	0.9994 (0.2466)	1.0072 (0.2037)	1.0012 (0.2481)
γ_3^μ	0.0000	0.0088 (0.6124)		−0.0239 (0.7479)	
γ_4^μ	4.5000	4.4286 (3.6203)	4.2291 (4.7358)	4.6572 (5.5419)	4.3361 (6.3122)
		Pair (iv)			
γ_2^u	1.0000	1.0019 (0.0474)	0.9999 (0.0654)	1.0005 (0.0462)	1.0004 (0.0663)
γ_3^u	0.0000	−0.0038 (0.1313)		0.0029 (0.1340)	
γ_4^u	3.0000	3.0108 (0.3744)	2.9604 (0.7585)	3.0021 (0.3746)	2.9542 (0.8156)
γ_2^μ	1.0000	1.0025 (0.3903)	0.9987 (0.4167)	0.9870 (0.3675)	0.9894 (0.3970)
γ_3^μ	2.8284	2.8708 (2.6188)		2.7569 (2.4520)	
γ_4^μ	15.0000	15.3352 (23.8481)	14.8863 (25.5117)	14.4445 (24.0418)	14.1470 (24.3741)

Cornwell and Trumbull (1994) emphasized the existence of the county-specific effect μ_i as well as the time effect η_t, and they were also verified by some F tests in Baltagi (2006). This section analyzes this panel data again to demonstrate the performance of the three proposed tests.

We concentrate on a subset of the above panel data, which consists of 21 counties in western North Carolina, since the p-values of tests on the whole data are all tiny. The test T_μ and other two tests for the county-specific effect, BCL_μ and F_μ, are first considered for the data in the first T years with $T = 3, \ldots, 7$. The p-values of BCL_μ with $T = 3$ and 4 are both around 0.8, and others are all less than 0.0001. We may conclude that the test BCL_μ fails to provide a correct decision due to the small values of T, and the county-specific effect is confirmed for this subset. Note that the test F_μ is equivalent to T_μ^*, and the three test statistics, BCL_μ, T_μ^* and T_μ, are comparable since they have the same asymptotic distribution. Table 8 lists the values of these three test statistics. It can be seen that the values of T_μ are significantly greater than those of other two test statistics. This partially verifies the powerfulness of our test T_μ comparing with BCL_μ and F_μ. Note that the county-specific characteristics were argued by Cornwell and Trumbull (1994) to be correlated with the covariates $(P_{it}', X_{it}')'$, see also Baltagi (2006).

Table 8
Values of three test statistics for the county-specific effect, BCL_μ, T_μ^* and T_μ.

	Length of time points (T)				
	3	4	5	6	7
BCL_μ	−0.87*	−0.82**	4.79	5.30	5.05
T_μ^*	31.76	33.83	28.41	28.79	20.49
T_μ	5527.00	1936.60	1578.50	280.81	462.66

Note: the p-values of BCL_μ at $T = 3$ and 4 are respectively 0.81 (*) and 0.79 (**), and all others are less than 0.0001.

Table 9
The p-values of the three tests for the time effect, BCL_η, F_η and T_η.

Tests	Length of time points (T)				
	3	4	5	6	7
BCL_η	0.5920	0.0898	0.1520	0.1376	0.0240
F_η	0.4226	0.1037	0.1364	0.1474	0.0457
T_η	0.2228	0.0317	0.0591	0.0759	0.0162

We next test for the presence of the time effect. The covariates P_{it} and X_{it} are first centered by $\bar{P}_{.t} = n^{-1}\sum_{i=1}^n P_{it}$ and $\bar{X}_{.t} = n^{-1}\sum_{i=1}^n X_{it}$, and then the test T_η as well as other two tests for the time effect, BCL_η and F_η, are conducted. Their p-values are given in Table 9, and those for the test T_η are calculated from a χ^2_{T-1}. At the 5% significance level, the time effect is confirmed by these three tests when $T = 7$, and is shown to be absent by BCL_η and F_η when $T < 7$. We may conclude the existence of the time effect based on our test while tests BCL_η and F_η may require a large sample size to be effective. From Table 9, the p-values of T_η are all less than those of BCL_η and F_η, and we may argue that our test T_η is more powerful. Finally, we perform the test $T_{\mu\eta 2}$ with $\omega = 0.5$ as well as $T_{\mu\eta 1}$, $\mathrm{BP}_{\mu\eta}$ and $F_{\mu\eta}$ to jointly check the presence of both effects, and all calculated p-values are less than 0.0001. Note that these four test statistics have different null distributions, and we may not be able to compare their powerfulness by values of the test statistics.

7. Discussions

An important application of tests for individual and time effects is to construct an efficient estimator of the coefficients β. For example, if the null hypothesis H_0^η for the time effect is accepted, then we may employ the feasible GLS to estimate β, and it is asymptotically efficient regardless of the absence of the individual effect. However, when H_0^η is wrongly accepted, the resulting feasible GLS estimator will be inefficient and even inconsistent as η_t is a fixed effect. Suppose that $EX_i = c\boldsymbol{\iota}_T$ for some scalar c, $\{\eta_t\}$ are *i.i.d.* with $N(0, \sigma_\eta^2)$, and are independent of $\{u_{it}\}$ and $\mathbf{X}_i$, and the significance level is set to 5%. When $\sigma_\eta^2 = \sigma_2^2/n > 0$ and $n \to \infty$, by Theorem 2, we can calculate the probability that H_0^η is falsely accepted,

$$P(T_\eta < \chi^2_{0.05,T-1}) = F\left(\frac{\chi^2_{0.05,T-1}}{(\sigma_2^2/\sigma_u^2) + 1}\right),$$

where $F(\cdot)$ is the distribution function of χ^2_{T-1}, and $\chi^2_{0.05,T-1}$ is the 95th percentile of χ^2_{T-1} with $F(\chi^2_{0.05,T-1}) = 0.95$. This probability will tend to 95% when $\sigma_2^2 \to 0$, and it may be even higher for a finite sample size n. Leeb and Potscher (2005) carefully studied all kinds of situations for the parameter estimations following a predetermined model selection procedure, and these are called "post-model-selection estimations". It is interesting to similarly discuss the "post-testing estimators" of the coefficients β under the proposed tests in this paper.

The second possible extension is to apply the methodology in this paper to the two-way error component model with autoregressive (AR) idiosyncratic errors, see Lillard and Willis (1978) and Chapter 5 of Baltagi (2008). For example, we can first apply the Prais–Winsten transformation matrix to transform the AR errors into serially uncorrelated errors, and then obtain the estimators of the coefficients in the AR part iteratively. Based on these fitted parameters, we can transform the serially correlated panel data model into the standard two-way error component model (1), and then the method in this paper can be applied.

Another application is to the spatial panel data models, which have attracted more and more attentions, see Baltagi et al. (2007), Yu et al. (2008), Lee and Yu (2010a,b), etc. The maximum likelihood estimation (MLE) is usually considered in this literature, and then the corresponding tests are based on the assumption of Gaussian. However, Baltagi (2008, Chapter 10) argued that the MLE may involve substantial problems in computation when the cross section dimension (n) is large. Kapoor et al. (2007) suggested a method of generalized moment (GM) estimation for the spatial parameter in the spatial panel data model with no assumption on the distributions. Although Kapoor et al. (2007) did not consider the time effect in the disturbances, it is not difficult to modify their moment conditions to obtain a new GM estimation when the time effect is presented. Based on this estimate, the spatial panel data model can be rewritten into the classical static panel data model (1). Then tests for individual and time effects and the estimation of higher order moments can be derived similarly.

Finally, after some proper modifications, the methodology in this paper can be extended to handle the dynamic panel models, see Wu and Zhu (2012). In addition, Li and Zhu (2010) considered a test for semi-parametric mixed models for longitudinal data without the time effect. We believe that our method can be further extended to more general settings such as semi-parametric or non-parametric models.

Acknowledgments

We thank Professor C. Hsiao, an associate editor and two anonymous referees for valuable comments that led to the substantial improvement of this paper, Professor W.K. Li for useful discussions, and the National Nature Science Foundation of China (Grant No. 11001238), the Specialized Research Fund for the Doctoral Program of Higher Education (Grant No. 20103326120002) and the Hong Kong RGC grant HKU703710P for partial support. Research was partially completed while G. Li was visiting the Institute for Mathematical Sciences, National University of Singapore in 2011.

Appendix. Technical details

We first present the proofs of Theorem 1, Corollaries 1 and 2, Theorem 2, Corollary 3 and Theorem 3, and then discuss the case of higher order moments of the idiosyncratic error and the individual effect.

Proof of Theorem 1. Since $\{\mathbf{X}_i\}$ and $\{\mathbf{u}_i\}$ are both *i.i.d.* sequences, we have that

$$n^{-1/2}\sum_{i=1}^{n}[\mathbf{X}_i - E(\mathbf{X}_i)] = O_p(1), \qquad n^{-1/2}\sum_{i=1}^{n}\mathbf{u}_i = O_p(1),$$

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\widetilde{\mathbf{X}}_i' P_{\iota_T}^{\perp}\widetilde{\mathbf{X}}_i &= \frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i' P_{\iota_T}^{\perp}\mathbf{X}_i \\ &\quad - \left(n^{-1}\sum_{j=1}^{n}\mathbf{X}_j\right)' P_{\iota_T}^{\perp}\left(n^{-1}\sum_{j=1}^{n}\mathbf{X}_j\right) \\ &= \Sigma_1 + o_p(1),\end{aligned} \tag{11}$$

and

$$\begin{aligned}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\widetilde{\mathbf{X}}_i' P_{\iota_T}^{\perp}\widetilde{\mathbf{u}}_i &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}(\mathbf{X}_i - E\mathbf{X}_i)' P_{\iota_T}^{\perp}\mathbf{u}_i \\ &\quad + \left[n^{-1}\sum_{i=1}^{n}(\mathbf{X}_i - E\mathbf{X}_i)\right]' P_{\iota_T}^{\perp}\left[n^{-1/2}\sum_{j=1}^{n}\mathbf{u}_j\right] \\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}(\mathbf{X}_i - E\mathbf{X}_i)' P_{\iota_T}^{\perp}\mathbf{u}_i + o_p(1).\end{aligned} \tag{12}$$

Note that $E[(\mathbf{X}_i - E\mathbf{X}_i)' P_{\iota_T}^{\perp}\mathbf{u}_i] = E(\mathbf{X}_i'\mathbf{u}_i) - T^{-1}\mathrm{cov}(\mathbf{X}_i'\iota_T, \mathbf{u}_i'\iota_T) = 0$. Together with (11) and (12), it implies that

$$\begin{aligned}\sqrt{n}(\widehat{\beta} - \beta) &= \Sigma_1^{-1}\cdot\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(\mathbf{X}_i - E\mathbf{X}_i)' P_{\iota_T}^{\perp}\mathbf{u}_i \\ &\quad + o_p(1) \to_d N(0, \Sigma_1^{-1}\Sigma_2\Sigma_1^{-1}),\end{aligned} \tag{13}$$

where $\Sigma_1 = E(\mathbf{X}_i' P_{\iota_T}^{\perp}\mathbf{X}_i) - E(\mathbf{X}_i') P_{\iota_T}^{\perp} E(\mathbf{X}_i)$, and $\Sigma_2 = E[(\mathbf{X}_i - E\mathbf{X}_i)' P_{\iota_T}^{\perp}\mathbf{u}_i\mathbf{u}_i' P_{\iota_T}^{\perp}(\mathbf{X}_i - E\mathbf{X}_i)]$. From (11)–(13), we can show that

$$\begin{aligned}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\|Q'(\widetilde{\mathbf{y}}_i - \widetilde{\mathbf{X}}_i\widehat{\beta})\|^2 &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\|Q'\widetilde{\mathbf{u}}_i + Q'\widetilde{\mathbf{X}}_i(\beta - \widehat{\beta})\|^2 \\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\widetilde{\mathbf{u}}_i' P_{\iota_T}^{\perp}\widetilde{\mathbf{u}}_i + o_p(1) \\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{u}_i' P_{\iota_T}^{\perp}\mathbf{u}_i + \left[n^{-1}\sum_{i=1}^{n}\mathbf{u}_i\right]' \\ &\quad \times P_{\iota_T}^{\perp}\left[n^{-1/2}\sum_{j=1}^{n}\mathbf{u}_j\right] + o_p(1) \\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\|Q'\mathbf{u}_i\|^2 + o_p(1),\end{aligned}$$

which implies that

$$\widehat{\sigma}_{0u}^2 = \frac{1}{n}\sum_{i=1}^{n}\|Q'\mathbf{u}_i\|^2/(T-1) + o_p(n^{-1/2}). \tag{14}$$

From Assumption (A2), we can show that

$$\begin{aligned}&n^{-1/4}\sum_{i=1}^{n}\mu_i = O_p(1), \qquad n^{-1/4}\sum_{i=1}^{n}\mu_i\mathbf{u}_i = O_p(1) \quad \text{and} \\ &n^{-3/4}\sum_{i=1}^{n}\mu_i\mathbf{X}_i = \Omega + o_p(1).\end{aligned} \tag{15}$$

It is implied that

$$\begin{aligned}&n^{-1/4}\sum_{i=1}^{n}\widetilde{\mu}_i\widetilde{\mathbf{u}}_i = O_p(1) \quad \text{and} \\ &n^{-3/4}\sum_{i=1}^{n}\widetilde{\mu}_i\widetilde{\mathbf{X}}_i = \Omega + o_p(1).\end{aligned} \tag{16}$$

Furthermore, by using methods respectively similar to (11) and (12), we have that

$$n^{-1}\sum_{i=1}^{n}\widetilde{\mathbf{X}}_i'\widetilde{\mathbf{X}}_i = O_p(1) \quad \text{and} \quad n^{-1/2}\sum_{i=1}^{n}\widetilde{\mathbf{X}}_i'\widetilde{\mathbf{u}}_i = O_p(1). \tag{17}$$

Thus, by (16), (17) and Assumption (A1),

$$\begin{aligned}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\|\widetilde{\mathbf{y}}_i - \widetilde{\mathbf{X}}_i\widehat{\beta}\|^2 &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\|\widetilde{\mathbf{u}}_i + \widetilde{\mu}_i\boldsymbol{\iota}_T + \widetilde{\mathbf{X}}_i(\beta-\widehat{\beta})\|^2 \\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\|\widetilde{\mathbf{u}}_i\|^2 + \|\widetilde{\mu}_i\boldsymbol{\iota}_T\|^2 + o_p(1) \\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\|\mathbf{u}_i\|^2 + T\sigma_1^2 + o_p(1)\end{aligned}$$

and

$$\widehat{\sigma}_{1u}^2 = \frac{1}{n}\sum_{i=1}^{n}\|\mathbf{u}_i\|^2/T + \sigma_1^2 + o_p(1). \tag{18}$$

Let $\xi_i = \|\mathbf{u}_i\|^2/T - \|Q'\mathbf{u}_i\|^2/(T-1)$. It holds that $E(\xi_i) = 0$ and $E(\xi_i^2) = 2(\sigma_u^2)^2/[T(T-1)]$. It is noteworthy that $\{\xi_i\}$ is an *i.i.d.* sequence. By (14), (18) and the central limit theorem, we can show that

$$\begin{aligned}\sqrt{n}(\widehat{\sigma}_{1u}^2 - \widehat{\sigma}_{0u}^2) &= \sigma_1^2 + \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\xi_i + o_p(1) \\ &\to_d \sigma_1^2 + N(0, 2(\sigma_u^2)^2/[T(T-1)]).\end{aligned}$$

Note that $\widehat{\sigma}_{0u}^2$ is a consistent estimator of σ_u^2 under both H_0^μ and H_1^μ. Following Slutsky's theorem, we can show that $T_\mu \to_d \sigma_1^2[0.5T(T-1)]^{1/2}/\sigma_u^2 + N(0,1)$. □

Proof of Corollary 1. Notice that $\widetilde{\mathbf{y}}_i = \widetilde{\mathbf{X}}_i\beta + \widetilde{\mu}_i\boldsymbol{\iota}_T + \widetilde{\mathbf{u}}_i$, and then

$$\widetilde{\beta}_1 = \left(\sum_{i=1}^{n}\widetilde{\mathbf{X}}_i'\widetilde{\mathbf{X}}_i\right)^{-1}\sum_{i=1}^{n}\widetilde{\mathbf{X}}_i'\widetilde{\mathbf{y}}_i = \beta + \nu_{1n} + \nu_{2n},$$

where $\nu_{1n} = \left(\sum_{i=1}^{n}\widetilde{\mathbf{X}}_i'\widetilde{\mathbf{X}}_i\right)^{-1}\sum_{i=1}^{n}\widetilde{\mathbf{X}}_i'\widetilde{\mu}_i\boldsymbol{\iota}_T$ and $\nu_{2n} = \left(\sum_{i=1}^{n}\widetilde{\mathbf{X}}_i'\widetilde{\mathbf{X}}_i\right)^{-1}\sum_{i=1}^{n}\widetilde{\mathbf{X}}_i'\widetilde{\mathbf{u}}_i$. From (16) and (17), we can show that $n^{1/4}\nu_{1n} = \Sigma_3^{-1}\Omega'\boldsymbol{\iota}_T + o_p(1)$, $\nu_{2n} = O_p(n^{-1/2})$, and

$$\begin{aligned}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\|\widetilde{\mathbf{y}}_i - \widetilde{\mathbf{X}}_i\widetilde{\beta}_1\|^2 &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\|\widetilde{\mathbf{u}}_i + \widetilde{\mu}_i\boldsymbol{\iota}_T - \widetilde{\mathbf{X}}_i\nu_{1n} - \widetilde{\mathbf{X}}_i\nu_{2n}\|^2 \\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}(\|\widetilde{\mathbf{u}}_i\|^2 + \|\widetilde{\mu}_i\boldsymbol{\iota}_T - \widetilde{\mathbf{X}}_i\nu_{1n}\|^2) + o_p(1) \\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\|\mathbf{u}_i\|^2 + T\sigma_1^2 - \boldsymbol{\iota}_T'\Omega\Sigma_3^{-1}\Omega'\boldsymbol{\iota}_T + o_p(1).\end{aligned}$$

Together with (14), it holds that

$$\begin{aligned}\sqrt{n}(\widetilde{\sigma}_{1u}^2 - \widehat{\sigma}_{0u}^2) &= \sigma_1^2 - \frac{1}{T}\boldsymbol{\iota}_T'\Omega\Sigma_3^{-1}\Omega'\boldsymbol{\iota}_T + \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\xi_i + o_p(1) \\ &\to_d \sigma_1^2 - T^{-1}\boldsymbol{\iota}_T'\Omega\Sigma_3^{-1}\Omega'\boldsymbol{\iota}_T \\ &\quad + N(0, 2(\sigma_u^2)^2/[T(T-1)]),\end{aligned}$$

where $\{\xi_i\}$ are defined as in the proof of Theorem 1. Following Slutsky's theorem again, we finish the proof. □

Proof of Corollary 2. By Assumption (A3), it can be shown that $n^{-1/4}\sum_{i=1}^{n}\mu_i = O_p(1)$, $n^{-1/4}\sum_{i=1}^{n}\mu_i\mathbf{u}_i = O_p(1)$ and $n^{-3/4}\sum_{i=1}^{n}\mu_i\mathbf{X}_i = O_p(1)$, and they imply that

$$n^{-1/4}\sum_{i=1}^{n}\widetilde{\mu}_i\widetilde{\mathbf{u}}_i = O_p(1) \quad \text{and} \quad n^{-3/4}\sum_{i=1}^{n}\widetilde{\mu}_i\widetilde{\mathbf{X}}_i = O_p(1).$$

Together with (17), by a method similar to (18), we can show that

$$\widehat{\sigma}_{1u}^2 = \frac{1}{n}\sum_{i=1}^{n}\|\mathbf{u}_i\|^2/T + \sigma_1^2 + o_p(1).$$

Following the proof of Theorem 1, we finish the proof. □

Proof of Theorem 2. It can be shown that, by (13),

$$\begin{aligned}&\frac{1}{\sqrt{n}}\sum_{j=1}^{n}Q'(\mathbf{y}_j - \mathbf{X}_j\widehat{\beta}) \\ &= \sqrt{n}Q'\boldsymbol{\eta} + \frac{1}{\sqrt{n}}\sum_{j=1}^{n}Q'\mathbf{u}_j - \frac{1}{n}\sum_{j=1}^{n}Q'\mathbf{X}_j\cdot\sqrt{n}(\widehat{\beta}-\beta) \\ &= \sqrt{n}Q'\boldsymbol{\eta} + \frac{1}{\sqrt{n}}\sum_{j=1}^{n}Q'\{I_T - E(\mathbf{X}_i)\Sigma_1^{-1}[\mathbf{X}_i - E(\mathbf{X}_i)]'P_{\iota_T}^{\perp}\}\mathbf{u}_j \\ &\quad + o_p(1) \\ &\to_d Q'\boldsymbol{\zeta}^* + \sigma_u\boldsymbol{\zeta}_2,\end{aligned} \tag{19}$$

where $\boldsymbol{\zeta}_2$ is a $(T-1)$-dimensional normal random vector with mean zero and variance matrix $I_{T-1} - \sigma_u^{-2}\Sigma_5$. It holds that

$$E(\xi_i\mathbf{u}_i) = \frac{1}{T}E(\|\mathbf{u}_i\|^2\mathbf{u}_i) - \frac{1}{T-1}E(\|Q'\mathbf{u}_i\|^2\mathbf{u}_i) = 0,$$

where $\xi_i = \|\mathbf{u}_i\|^2/T - \|Q'\mathbf{u}_i\|^2/(T-1)$ is defined as in the proof of Theorem 1, and $n^{-1/2}\sum_{i=1}^{n}\xi_i \to_d \sqrt{E\xi_i^2}\cdot\zeta_1$. Hence, when $\{\mathbf{u}_i\}$ is independent of $\{\mathbf{X}_i\}$, ζ_1 and $\boldsymbol{\zeta}_2$ are independent.

We further have that

$$\begin{aligned}&\sum_{i=1}^{n}\|Q'(\widetilde{\mathbf{y}}_i - \widetilde{\mathbf{X}}_i\widehat{\beta})\|^2 \\ &= \sum_{i=1}^{n}\left\|Q'(\mathbf{y}_i - \mathbf{X}_i\widehat{\beta}) - \frac{1}{n}\sum_{j=1}^{n}Q'(\mathbf{y}_j - \mathbf{X}_j\widehat{\beta})\right\|^2 \\ &= \sum_{i=1}^{n}\|Q'(\mathbf{y}_i - \mathbf{X}_i\widehat{\beta})\|^2 - \left\|\frac{1}{\sqrt{n}}\sum_{j=1}^{n}Q'(\mathbf{y}_j - \mathbf{X}_j\widehat{\beta})\right\|^2,\end{aligned} \tag{20}$$

and

$$\widehat{\sigma}_{2u}^2 = \frac{n-1}{n}\widehat{\sigma}_{0u}^2 + \frac{1}{n(T-1)}\left\|\frac{1}{\sqrt{n}}\sum_{j=1}^{n}Q'(\mathbf{y}_j - \mathbf{X}_j\widehat{\beta})\right\|^2.$$

Then, by (19)

$$\begin{aligned}T_\eta &= \frac{1}{\widehat{\sigma}_{0u}^2}\cdot\left\|\frac{1}{\sqrt{n}}\sum_{j=1}^{n}Q'(\mathbf{y}_j - \mathbf{X}_j\widehat{\beta})\right\|^2 \\ &= \frac{1}{\sigma_u^2}\cdot\left\|\frac{1}{\sqrt{n}}\sum_{j=1}^{n}Q'(\mathbf{y}_j - \mathbf{X}_j\widehat{\beta})\right\|^2 \\ &\quad + o_p(1) \to_d \|\sigma_u^{-1}Q'\boldsymbol{\zeta}^* + \boldsymbol{\zeta}_2\|^2.\end{aligned}$$

Note that the null hypothesis H_0^η corresponds to the case with $\sigma_2 = 0$. Hence, we finish the proof. □

Proof of Corollary 3. Denote $\Sigma_0 = E(\mathbf{X}_i'QQ'\mathbf{X}_i) = E(\mathbf{X}_i'P_{\iota_T}^{\perp}\mathbf{X}_i)$. By (13) and the condition that $Q'E(\mathbf{X}_i) = 0$, we have that $\Sigma_1 = \Sigma_0$ and

$$\sqrt{n}(\widehat{\beta} - \beta) = \Sigma_0^{-1} \cdot \left[\frac{1}{\sqrt{n}} \sum_{i=1}^n \mathbf{X}_i'QQ'\mathbf{u}_i - \left(\frac{1}{n}\sum_{i=1}^n \mathbf{X}_i'Q\right)\left(\frac{1}{\sqrt{n}}\sum_{i=1}^n Q'\mathbf{u}_i\right)\right] + o_p(n^{-1/2}) = O_p(1). \tag{21}$$

It holds that

$$\widetilde{\beta}_2 = \left(\sum_{i=1}^n \mathbf{X}_i'QQ'\mathbf{X}_i\right)^{-1} \sum_{i=1}^n \mathbf{X}_i'QQ'\mathbf{y}_i = \beta + \left(\sum_{i=1}^n \mathbf{X}_i'QQ'\mathbf{X}_i\right)^{-1} \sum_{i=1}^n \mathbf{X}_i'QQ'\mathbf{u}_i + \left(\sum_{i=1}^n \mathbf{X}_i'QQ'\mathbf{X}_i\right)^{-1} \sum_{i=1}^n \mathbf{X}_i'QQ'\boldsymbol{\eta}.$$

Notice that $\boldsymbol{\eta} = O_p(n^{-1/2})$. Thus,

$$\sqrt{n}(\widetilde{\beta}_2 - \beta) = \Sigma_0^{-1} \cdot \frac{1}{\sqrt{n}} \sum_{i=1}^n \mathbf{X}_i'QQ'\mathbf{u}_i + o_p(1) = O_p(1), \tag{22}$$

and

$$n(\widetilde{\beta}_2 - \widehat{\beta}) = \Sigma_0^{-1} \cdot \left[\left(\frac{1}{\sqrt{n}}\sum_{i=1}^n \mathbf{X}_i'Q\right) \cdot \sqrt{n}Q'\boldsymbol{\eta} + \left(\frac{1}{\sqrt{n}}\sum_{i=1}^n \mathbf{X}_i'Q\right)\left(\frac{1}{\sqrt{n}}\sum_{i=1}^n Q'\mathbf{u}_i\right)\right] + o_p(1) = O_p(1). \tag{23}$$

By (21)–(23), we can show that

$$\sum_{i=1}^n \|Q'(\mathbf{y}_i - \mathbf{X}_i\widetilde{\beta}_2)\|^2 = \sum_{i=1}^n \|Q'(\mathbf{y}_i - \mathbf{X}_i\widehat{\beta}) - Q'\mathbf{X}_i(\widetilde{\beta}_2 - \widehat{\beta})\|^2 = \sum_{i=1}^n \|Q'(\mathbf{y}_i - \mathbf{X}_i\widehat{\beta})\|^2 + o_p(1).$$

Together with (20) and following the proof of Theorem 2, we finish the proof. □

Proof of Theorem 3. We first show the result of test statistic $T_{\mu\eta 1}$. Note that $\mathbf{y}_i = \alpha\boldsymbol{\iota}_T + \mathbf{X}_i\beta + \mu_i\boldsymbol{\iota}_T + \boldsymbol{\eta} + \mathbf{u}_i$. By (13), (15), Assumptions (A1), (A2) and (A4), it holds that

$$\sqrt{n}(\widehat{\alpha} - \alpha) = \frac{1}{Tn}\sum_{i=1}^n \boldsymbol{\iota}_T'\mathbf{X}_i \cdot \sqrt{n}(\beta - \widehat{\beta}) + \frac{1}{\sqrt{n}}\sum_{i=1}^n \mu_i + \frac{\sqrt{n}}{T}\boldsymbol{\iota}_T'\boldsymbol{\eta} + \frac{1}{T\sqrt{n}}\sum_{i=1}^n \boldsymbol{\iota}_T'\mathbf{u}_i = O_p(1), \tag{24}$$

and

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n \|\mathbf{y}_i - \widehat{\alpha}\boldsymbol{\iota}_T - \mathbf{X}_i\widehat{\beta}\|^2 = \frac{1}{\sqrt{n}}\sum_{i=1}^n \|(\alpha - \widehat{\alpha})\boldsymbol{\iota}_T + \mathbf{X}_i(\beta - \widehat{\beta}) + \mu_i\boldsymbol{\iota}_T + \boldsymbol{\eta} + \mathbf{u}_i\|^2 = \frac{1}{\sqrt{n}}\sum_{i=1}^n \|\mathbf{u}_i\|^2 + T\sigma_1^2 + o_p(1).$$

Thus,

$$\widehat{\sigma}_{3u}^2 = \frac{1}{n}\sum_{i=1}^n \|\mathbf{u}_i\|^2/T + \sigma_1^2 + o_p(1).$$

Together with (14) and following the proof of Theorem 1, we derive the asymptotic distribution of $T_{\mu\eta 1}$.

Let $\bar{\mathbf{X}} = n^{-1}\sum_{i=1}^n \mathbf{X}_i$, $\bar{\mathbf{y}} = n^{-1}\sum_{i=1}^n \mathbf{y}_i$, and $\bar{\mathbf{u}} = n^{-1}\sum_{i=1}^n \mathbf{u}_i$. For quantity $\widetilde{\beta}_3$ in test statistic $T_{\mu\eta 1}^*$, we have that

$$\widetilde{\beta}_3 = \left(n^{-1}\sum_{i=1}^n \widetilde{\mathbf{X}}_i'\widetilde{\mathbf{X}}_i + \bar{\mathbf{X}}'QQ'\bar{\mathbf{X}}\right)^{-1}\left(n^{-1}\sum_{i=1}^n \widetilde{\mathbf{X}}_i'\widetilde{\mathbf{y}}_i + \bar{\mathbf{X}}'QQ'\bar{\mathbf{y}}\right) = \beta + \nu_{1n}^* + \nu_{2n}^*,$$

where

$$n^{1/4}\nu_{1n}^* = \left(n^{-1}\sum_{i=1}^n \widetilde{\mathbf{X}}_i'\widetilde{\mathbf{X}}_i + \bar{\mathbf{X}}'QQ'\bar{\mathbf{X}}\right)^{-1} n^{-3/4}\sum_{i=1}^n \widetilde{\mathbf{X}}_i'\widetilde{\mu}_i\boldsymbol{\iota}_T = \Sigma_6^{-1}\Omega'\boldsymbol{\iota}_T + o_p(1),$$

and

$$\nu_{2n}^* = \left(n^{-1}\sum_{i=1}^n \widetilde{\mathbf{X}}_i'\widetilde{\mathbf{X}}_i + \bar{\mathbf{X}}'QQ'\bar{\mathbf{X}}\right)^{-1} \times \left[n^{-1}\sum_{i=1}^n \widetilde{\mathbf{X}}_i'\widetilde{\mathbf{u}}_i + \bar{\mathbf{X}}QQ'(\bar{\mathbf{u}} + \boldsymbol{\eta})\right] = O_p(n^{-1/2}),$$

see also the proof of Corollary 1. By a method similar to (24), it can be shown that

$$\widetilde{\alpha}_3 = (nT)^{-1}\sum_{i=1}^n \boldsymbol{\iota}_T'(\mathbf{y}_i - \mathbf{X}_i\widetilde{\beta}_3) = \alpha + \tau_{1n} + \tau_{2n},$$

where

$$n^{1/4}\tau_{1n} = -\frac{1}{Tn}\sum_{i=1}^n \boldsymbol{\iota}_T'\mathbf{X}_i \cdot n^{1/4}\nu_{1n}^* = -T^{-1}\boldsymbol{\iota}_T'E(\mathbf{X}_i)\Sigma_6^{-1}\Omega'\boldsymbol{\iota}_T + o_p(1),$$

and

$$\tau_{2n} = -\frac{1}{Tn}\sum_{i=1}^n \boldsymbol{\iota}_T'\mathbf{X}_i \cdot \nu_{2n}^* + \frac{1}{n}\sum_{i=1}^n \mu_i + \frac{1}{T}\boldsymbol{\iota}_T'\boldsymbol{\eta} + \frac{1}{Tn}\sum_{i=1}^n \boldsymbol{\iota}_T'\mathbf{u}_i = O_p(n^{-1/2}).$$

Thus,

$$\begin{aligned}\frac{1}{\sqrt{n}}\sum_{i=1}^n \|\mathbf{y}_i - \widetilde{\alpha}_3\boldsymbol{\iota}_T - \mathbf{X}_i\widetilde{\beta}_3\|^2 &= \frac{1}{\sqrt{n}}\sum_{i=1}^n \|(\alpha - \widetilde{\alpha}_3)\boldsymbol{\iota}_T + \mathbf{X}_i(\beta - \widetilde{\beta}_3) + \mu_i\boldsymbol{\iota}_T + \boldsymbol{\eta} + \mathbf{u}_i\|^2 \\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^n (\|\mathbf{u}_i\|^2 + \|\mu_i\boldsymbol{\iota}_T - \tau_{1n}\boldsymbol{\iota}_T - \mathbf{X}_i\nu_{1n}^*\|^2) + o_p(1) \\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^n \|\mathbf{u}_i\|^2 + T\sigma_1^2 - T\delta + o_p(1).\end{aligned}$$

Together with (14) and following the proof of Theorem 1, we derive the asymptotic distribution of $T_{\mu\eta 1}^*$. □

Higher order moments of the idiosyncratic error and the individual effect. For $k \geq 2$, denote by γ_k^u the kth order moment of the idiosyncratic error u_{it}, and by γ_k^μ the kth order moment of the individual effect μ_i. We next attempt to construct the estimators of γ_k^u, which are robust to the presence of both individual and time effects, and the estimators of γ_k^μ, which are robust to the presence of the time effect. Note that $\gamma_2^u = \sigma_u^2$.

The higher order moments of the idiosyncratic error γ_k^u are estimated based on the residuals from model (5). For $k = 2$, we can use $\widehat{\sigma}_{0u}^2$ in (7) to estimate γ_2^u, i.e. $\widehat{\gamma}_2^u = \widehat{\sigma}_{0u}^2$.

Denote $q_1 = T^{-1/2}\boldsymbol{\iota}_T$, $Q = (q_2, q_3, \ldots, q_T)$, $q_l = (q_{1l}, q_{2l}, \ldots, q_{Tl})'$ for $1 \leq l \leq T$, $a_{1n} = \left(\sum_{l=2}^T \sum_{t=1}^T q_{tl}^3\right)(n^2 - 3n + 2)/n^2$, $a_{2n} = \left(\sum_{l=2}^T \sum_{t=1}^T q_{tl}^4\right)(n-1)(n^2 - 3n + 3)/n^3$, and $a_{3n} = 3a_{2n}^{-1}(n-1)^2(T-1)/n^2 - 3$. It holds that

$$E\left[\sum_{l=2}^T (q_l'\widetilde{\mathbf{u}}_i)^3\right] = E\left[\left(\sum_{l=2}^T q_l^{\otimes 3}\right)'(\widetilde{\mathbf{u}}_i^{\otimes 3})\right] = a_{1n}\gamma_3^u,$$

and

$$E\left[\sum_{l=2}^T (q_l'\widetilde{\mathbf{u}}_i)^4\right] = E\left[\left(\sum_{l=2}^T q_l^{\otimes 4}\right)'(\widetilde{\mathbf{u}}_i^{\otimes 4})\right] = a_{2n}\gamma_4^u + a_{2n}a_{3n}(\gamma_2^u)^2,$$

where $a^{\otimes k} = \underbrace{a \otimes \cdots \otimes a}_{k}$ with positive integer k, and $\otimes$ is the Kronecker product. We then can estimate the third- and fourth-order moments of u_{it} by

$$\widehat{\gamma}_3^u = a_{1n}^{-1} \cdot \frac{1}{n}\sum_{i=1}^n \left(\sum_{l=2}^T q_l^{\otimes 3}\right)'(\widetilde{\mathbf{y}}_i - \widetilde{\mathbf{X}}_i\widehat{\beta})^{\otimes 3},$$

and

$$\widehat{\gamma}_4^u = a_{2n}^{-1} \cdot \frac{1}{n}\sum_{i=1}^n \left(\sum_{l=2}^T q_l^{\otimes 4}\right)'(\widetilde{\mathbf{y}}_i - \widetilde{\mathbf{X}}_i\widehat{\beta})^{\otimes 4} - a_{3n}(\widehat{\gamma}_2^u)^2,$$

where $\widehat{\beta}$ is defined as in (6). Note that the quantities, $\sum_{l=2}^T \sum_{t=1}^T q_{tl}^3$, $\sum_{l=2}^T \sum_{t=1}^T q_{tl}^4$ and $\sum_{l=2}^T q_l^{\otimes k}$ with $k = 2, 3$, and 4, all depend on the matrix Q.

Note that, by the orthogonal transformation, the information of μ_i in model (3) is concentrated in model (4). We then can make use of model (4) only to estimate the higher order moments of the individual effect γ_k^μ. Suppose $\{\mu_i\}$ is independent of $\{u_{it}\}$, and then we can separate the moments of μ_i from those of $T\widetilde{\mu}_i + \boldsymbol{\iota}_T\widetilde{\mathbf{u}}_i$, i.e.

$$E(T\widetilde{\mu}_i)^2 = E(T\widetilde{\mu}_i + \boldsymbol{\iota}_T\widetilde{\mathbf{u}}_i)^2 - E(\boldsymbol{\iota}_T\widetilde{\mathbf{u}}_i)^2,$$
$$E(T\widetilde{\mu}_i)^3 = E(T\widetilde{\mu}_i + \boldsymbol{\iota}_T\widetilde{\mathbf{u}}_i)^3 - E(\boldsymbol{\iota}_T\widetilde{\mathbf{u}}_i)^3$$

and

$$E(T\widetilde{\mu}_i)^4 = E(T\widetilde{\mu}_i + \boldsymbol{\iota}_T\widetilde{\mathbf{u}}_i)^4 - E(\boldsymbol{\iota}_T\widetilde{\mathbf{u}}_i)^4 - 6E(T\widetilde{\mu}_i)^2E(\boldsymbol{\iota}_T\widetilde{\mathbf{u}}_i)^2.$$

Similar to the counterpart of u_{it}, we can estimate the higher order moments of the individual effect as follows,

$$\widehat{\gamma}_2^\mu = b_{1n}^{-1} \cdot \frac{1}{n}\sum_{i=1}^n (\boldsymbol{\iota}_T'\widetilde{\mathbf{y}}_i - \boldsymbol{\iota}_T'\widetilde{\mathbf{X}}_i\widehat{\beta})^2 - T^{-1}\widehat{\gamma}_2^u,$$

$$\widehat{\gamma}_3^\mu = b_{2n}^{-1} \cdot \frac{1}{n}\sum_{i=1}^n (\boldsymbol{\iota}_T'\widetilde{\mathbf{y}}_i - \boldsymbol{\iota}_T'\widetilde{\mathbf{X}}_i\widehat{\beta})^3 - T^{-2}\widehat{\gamma}_3^u,$$

$$\widehat{\gamma}_4^\mu = b_{3n}^{-1} \cdot \frac{1}{n}\sum_{i=1}^n (\boldsymbol{\iota}_T'\widetilde{\mathbf{y}}_i - \boldsymbol{\iota}_T'\widetilde{\mathbf{X}}_i\widehat{\beta})^4 - T^{-3}\widehat{\gamma}_4^u - b_{4n}\widehat{\gamma}_2^u\widehat{\gamma}_2^\mu - b_{5n}(\widehat{\gamma}_2^u)^2 - b_{6n}(\widehat{\gamma}_2^\mu)^2,$$

where $b_{1n} = T^2(n-1)/n$, $b_{2n} = T^3(n^2 - 3n + 2)/n^2$, $b_{3n} = T^4(n-1)(n^2 - 3n + 3)/n^3$, $b_{4n} = 6T^{-1}n(n-1)/(n^2 - 3n + 3)$, $b_{5n} = T^{-3}[3(2n-3) + 3n(n-1)(T-1)]/(n^2 - 3n + 3)$, and $b_{6n} = 3(2n-3)/(n^2 - 3n + 3)$.

Denote

$$\kappa_0 = \frac{1}{T-1}\left(\sum_{l=2}^T q_l^{\otimes 2}\right)' \operatorname{cov}(\mathbf{u}_i^{\otimes 2}, \mathbf{u}_i^{\otimes 4})\left(\sum_{l=2}^T q_l^{\otimes 4}\right),$$

$$\kappa_1 = \frac{1}{\sum_{l=2}^T \sum_{t=1}^T q_{tl}^4},$$

$$\kappa_2 = -6[\kappa_1(T-1) - 1]\gamma_2^u,$$

$$\kappa_3 = \frac{1}{T^3(T-1)}[6T^2\gamma_2^\mu + 6(T-1)\gamma_2^u],$$

$$\kappa_4 = \gamma_4^\mu + \frac{1}{T^3}[\gamma_4^u + 3(T-1)(\gamma_2^u)^2] + \frac{6}{T}\gamma_2^\mu\gamma_2^u,$$

$\kappa = (\kappa_1, \kappa_2)'$ and $W_k = \boldsymbol{\iota}_{T^k} - T\sum_{l=2}^T \sum_{t=1}^T q_{tl}^k \cdot \sum_{l=2}^T q_l^{\otimes k}$, where $\boldsymbol{\iota}_{T^k}$ is a T^k-dimensional vector with all elements equal to one, and $k = 2, 3$ and 4. Let

$$\Upsilon_2 = \frac{1}{(T-1)^2}\left(\sum_{l=2}^T q_l^{\otimes 2}\right)' \operatorname{var}(\mathbf{u}_i^{\otimes 2})\left(\sum_{l=2}^T q_l^{\otimes 2}\right),$$

$$\Upsilon_3 = \frac{1}{\left(\sum_{l=2}^T \sum_{t=1}^T q_{tl}^3\right)^2}\left(\sum_{l=2}^T q_l^{\otimes 3}\right)' \operatorname{var}(\mathbf{u}_i^{\otimes 3})\left(\sum_{l=2}^T q_l^{\otimes 3}\right),$$

$$\Upsilon_4 = \kappa'\begin{pmatrix} \left(\sum_{l=2}^T q_l^{\otimes 4}\right)' \operatorname{var}(\mathbf{u}_i^{\otimes 4})\left(\sum_{l=2}^T q_l^{\otimes 4}\right) & \kappa_0 \\ \kappa_0 & \Upsilon_2 \end{pmatrix}\kappa,$$

$$\Psi_2 = \frac{1}{T^4}W_2'E[(\mu_i\boldsymbol{\iota}_T + \mathbf{u}_i)(\mu_i\boldsymbol{\iota}_T + \mathbf{u}_i)']^{\otimes 2}W_2 - (\gamma_2^\mu)^2,$$

$$\Psi_3 = \frac{1}{T^6}W_3'E[(\mu_i\boldsymbol{\iota}_T + \mathbf{u}_i)(\mu_i\boldsymbol{\iota}_T + \mathbf{u}_i)']^{\otimes 3}W_3 - (\gamma_3^\mu)^2,$$

and

$$\begin{aligned}\Psi_4 = {} & \frac{1}{T^8}W_4'E[(\mu_i\boldsymbol{\iota}_T + \mathbf{u}_i)(\mu_i\boldsymbol{\iota}_T + \mathbf{u}_i)']^{\otimes 4}W_4 \\ & + \kappa_3^2\left(\sum_{l=2}^T q_l^{\otimes 2}\right)' E[(\mu_i\boldsymbol{\iota}_T + \mathbf{u}_i)(\mu_i\boldsymbol{\iota}_T + \mathbf{u}_i)']^{\otimes 2}\left(\sum_{l=2}^T q_l^{\otimes 2}\right) \\ & - \frac{2\kappa_3}{T^4}W_4'E[(\mu_i\boldsymbol{\iota}_T + \mathbf{u}_i)^{\otimes 4} \otimes ((\mu_i\boldsymbol{\iota}_T + \mathbf{u}_i)^{\otimes 2})']\left(\sum_{l=2}^T q_l^{\otimes 2}\right) \\ & - (\kappa_4)^2.\end{aligned}$$

By a method similar to Wu and Zhu (2010), for $k = 2, 3$ and 4, we can show that, if $E(u_{it}^{2k}) < \infty$ and $E\|\mathbf{X}_i\|^{2k} < \infty$, then

$$\sqrt{n}(\widehat{\gamma}_k^u - \gamma_k^u) \to_d N(0, \Upsilon_k)$$

as $n \to \infty$ and, if $\{\mu_i\}$ is further independent of $\{u_{it}\}$ and $\{\mathbf{X}_i\}$ with $E(\mu_i^{2k}) < \infty$, then

$$\sqrt{n}(\widehat{\gamma}_k^\mu - \gamma_k^\mu) \to_d N(0, \Psi_k)$$

as $n \to \infty$. The proof is omitted to save the space, and is available upon request.

Note that the asymptotic results are based on large n and fixed T, and we do not have enough information to obtain a consistent estimation of the higher order moments of η_t.

We sometimes may be interested in the skewness and kurtosis, instead of the third and the fourth order moments, for asymmetry and heavy tails. Note that the skewness is $\gamma_3/(\gamma_2)^{3/2}$ and the kurtosis is $\gamma_4/(\gamma_2)^2 - 3$. By the Delta method (van der Vaart, 1998, Chapter 3), it is easy to derive the asymptotic normalities of the skewness and kurtosis estimators based on $\widehat{\gamma}_k$.

For the matrix Q in the above estimators and asymptotic variances, we suggest to use $Q = (q_2, q_3, \ldots, q_T)$ and $q_l = \{(l-1)I_T(l) - \sum_{k=1}^{l-1} I_T(k)\}/\sqrt{l(l-1)}$, where $2 \le l \le T$ and $I_T(k)$ stands for the kth column vector of the identity matrix I_T. It holds that, under this value of Q, $\sum_{l=2}^{T}\sum_{t=1}^{T} q_{tl}^3 = \sum_{l=2}^{T}(l-2)/\sqrt{l(l-1)}$ and $\sum_{l=2}^{T}\sum_{t=1}^{T} q_{tl}^4 = T - 1 + \sum_{l=2}^{T}(3-2l)/(l^2-l)$. We have tried some different values of Q, and the results are similar. □

References

Baltagi, B.H., 2006. Estimating an economic model of crime using panel data from north carolina. Journal of Applied Econometrics 21, 543–547.
Baltagi, B.H., 2008. Econometric Analysis of Panel Data. John Wiley and Sons, Ltd.
Baltagi, B.H., Bresson, G., Pirotte, A., 2006. Joint LM test for heteroskedasticity in a one-way error component model. Journal of Econometrics 134, 401–417.
Baltagi, B.H., Chang, Y., Li, Q., 1992. Monte Carlo results on several new and existing tests for the error component model. Journal of Econometrics 54, 95–120.
Baltagi, B.H., Jung, B.C., Song, S.H., 2010. Testing for heteroskedasticity and serial correlation in a random effects panel data model. Journal of Econometrics 154, 122–124.
Baltagi, B.H., Song, S.H., Jung, B.C., Koh, W., 2007. Testing for serial correlation, spatial autocorrelation and random effects using panel data. Journal of Econometrics 140, 5–51.
Bera, A., Sosa Escudero, W., Yoon, M., 2001. Tests for the error-component model in the presence of local misspecification. Journal of Econometrics 101, 1–23.
Breusch, T.S., Pagan, A.R., 1980. The Lagrange multiplier test and its applications to model specification in econometrics. Reviews of Economic Studies 47, 239–253.
Cornwell, C., Trumbull, W.N., 1994. Estimating the economic model of crime with panel data. Review of Economics and Statistics 76, 360–366.
Hausman, J.A., 1978. Specification tests in econometrics. Econometrica 46, 1251–1271.
Honda, Y., 1985. Testing the error components model with non-normal disturbances. Review of Economic Studies 52, 681–690.
Hsiao, C., 2003. Analysis of Panel Data. Cambridge University Press, Cambridge.
Kapoor, M., Kelejian, H.H., Prucha, I.R., 2007. Panel data models with spatially correlated error components. Journal of Econometrics 140, 97–130.
Lee, L.-F., Yu, J., 2010a. Estimation of spatial autoregressive panel data models with fixed effects. Journal of Econometrics 154, 165–185.
Lee, L.-F., Yu, J., 2010b. A spatial dynamic panel data model with both time and individual fixed effects. Econometric Theory 26, 564–597.
Leeb, H., Potscher, B.M., 2005. Model selection and inference: facts and fiction. Econometric Theory 21, 21–59.
Li, Z., Zhu, L.-X., 2010. On variance components in semiparametric mixed models for longitudinal data. Scandinavian Journal of Statistics 37, 442–457.
Lillard, L.A., Willis, R.J., 1978. Dynamic aspects of earning mobility. Econometrica 46, 985–1012.
Montes-Roja, G., Sosa-Escudero, W., 2011. Robust tests for heteroskedasticity in the one-way error components model. Journal of Econometrics 160, 300–310.
van der Vaart, A.W., 1998. Asymptotic Statistics. Cambridge University Press, Cambridge.
Wooldridge, J.M., 2002. Econometric Analysis of Cross-Section and Panel Data. MIT Press, Massachusetts.
Wu, J., Su, W., 2010. Estimation of moment for linear panel data models with potential existence of time effects. Statistics & Probability Letters 80, 1933–1939.
Wu, P., Zhu, L.-X., 2010. An orthogonality-based estimation of moments for linear mixed models. Scandinavian Journal of Statistics 37, 253–263.
Wu, J., Zhu, L.-X., 2012. Estimation of and testing for random effects in dynamic panel data models. Test 21, 477–497.
Yu, J., de Jong, R., Lee, L.-F., 2008. Quasi-maximum likelihood estimators for spatial dynamic panel data with fixed effects when n and T are large. Journal of Econometrics 146, 118–134.