JOURNAL OF APPLIED STATISTICS
2020, VOL. 47, NO. 1, 91–116
https://doi.org/10.1080/02664763.2019.1634682

Taylor & Francis
Taylor & Francis Group

Testing for heteroskedasticity in two-way fixed effects panel data models

Sanying Feng[a], Gaorong Li[b], Tiejun Tong[c] and Shuanghua Luo[d]

[a]School of Mathematics and Statistics, Zhengzhou University, Zhengzhou, People's Republic of China; [b]School of Statistics, Beijing Normal University, Beijing, People's Republic of China; [c]Department of Mathematics, Hong Kong Baptist University, Hong Kong, People's Republic of China; [d]School of Science, Xi'an Polytechnic University, Xi'an, People's Republic of China

ABSTRACT

In this paper, we propose a new method for testing heteroskedasticity in two-way fixed effects panel data models under two important scenarios where the cross-sectional dimension is large and the temporal dimension is either large or fixed. Specifically, we will develop test statistics for both cases under the conditional moment framework, and derive their asymptotic distributions under both the null and alternative hypotheses. The proposed tests are distribution free and can easily be implemented using the simple auxiliary regressions. Simulation studies and two real data analyses demonstrate that our proposed tests perform well in practice, and may have the potential for wide application in econometric models with panel data.

ARTICLE HISTORY
Received 3 December 2018
Accepted 10 June 2019

KEYWORDS
Asymptotic distribution; heteroskedasticity; panel data; testing; two-way fixed effects model

1. Introduction

Econometric models with panel data play an important role in empirical microeconomic and macroeconomic studies. They have received increasing attention in the recent literature, mainly because of their capacity and flexibility in simultaneously modeling the individual heterogeneity and time effects. For an overview on panel data analysis, one may refer to, for example, Arellano [1], Baltagi [4] and Hsiao [15].

In this paper, we consider the following two-way panel data model:

$$Y_{it} = X_{it}^{\top}\beta + \mu_i + \xi_t + \varepsilon_{it}, \quad i = 1, \ldots, N, \quad t = 1, \ldots, T, \tag{1}$$

where Y_{it} is the observed dependent variable for the ith individual at time t, $X_{it} = (X_{it,1}, \ldots, X_{it,p})^{\top}$ is a p-dimensional vector of strictly exogenous covariates with '$\top$' denoting the transpose of a vector or a matrix, and $\beta = (\beta_1, \ldots, \beta_p)^{\top}$ is a p-dimensional vector of unknown regression coefficients. We also assume that the unobserved individual-specific effect μ_i is time-invariant and it accounts for the individual's unobserved ability, and the unobserved time effect ξ_t is individual-invariant and it accounts for the time-specific effect not included in the regression. Furthermore, ε_{it} is the error term with the

CONTACT Gaorong Li ligaorong@gmail.com School of Statistics, Beijing Normal University, 100875 Beijing, People's Republic of China

zero conditional mean and the conditional variance as

$$E(\varepsilon_{it}^2|X_i, Z_i) = \sigma^2 g(Z_{it}^\top \theta),$$

where σ^2 is a positive constant, $g(\cdot)$ is an arbitrary strictly positive and twice differentiable function satisfying $g(0) = 1$ and $g'(0) \neq 0$, and θ is a k-dimensional vector of unrestricted parameters. In addition, Z_{it} is a k-dimensional covariate which can be taken as a subset or all of X_{it}, or with other strictly exogenous variables that are not associated with X_{it}. Also for ease of notation, we let X_i and Z_i denote the $T \times p$ and $T \times k$ matrices that consist of the T observations of X_{it} and Z_{it}, respectively.

Model (1) is called a two-way fixed effects panel data model when μ_i and ξ_t are allowed to be correlated with $\{X_{it}\}$ in an unknown correlation structure [4,11,15,19,20]. On the other side, model (1) yields a two-way random effects panel data model [6,10,13,14,17,18,29,30,32]. More recently, panel data models with interactive fixed effects that include two-way fixed effects models as special cases have also emerged in the literature [3,21,24,27,31].

Testing for heteroskedasticity is a crucial step and an important topic in panel data analysis. In the literature, many panel data models have assumed that the disturbances have homoskedastic variances. In practice, however, this can be a rather restrictive assumption for panel data generated from the real world. For instance, cross-sectional units may vary in size so that the conditional variance may exhibit a conditional heteroskedasticity. It is also known that heteroskedasticity may lead to inefficient least squares estimates and inconsistent covariance matrix estimates, when the error terms are incorrectly specified as homoskedastic.

As a remedy, it is often suggested to carry out a test for heteroskedasticity before a specific model for panel data can be applied. For one-way random effects panel data models, Li and Stengos [22] considered the setting where heteroskedasticity is only in the remainder error term. Holly and Gardiol [12] studied the case where heteroskedasticity is present only in the individual effect component, and derived a test statistic in a Lagrange multiplier (LM) framework under normality. Baltagi *et al.* [5] assumed that heteroskedasticity exists in both the random individual effects and the remainder error term and derived a joint LM test for homoskedasticity under the Gaussian assumption. Montes-Rojas and Sosa-Escudero [23] considered the scenario when the normality assumption is violated and developed a robust test for heteroskedasticity in one-way random effects models. Baltagi *et al.* [8] proposed a conditional LM test for heteroskedasticity for panel data models with serial correlation. For one-way fixed effects panel data models, Juhl and Sosa-Escudero [16] conducted a robust test for heteroskedasticity after the fixed effects estimation.

The aforementioned tests are mainly focused on one-way panel data models. When it is unknown whether the time effects ξ_t are absent, the above tests will no longer be applicable. For two-way random effects panel data models, Kouassi *et al.* [18] proposed a joint score test for heteroskedasticity. Kouassi *et al.* [17] derived a conditional LM test and a joint LM test for heteroskedasticity for panel data models with spatial correlation. To summarize, there is a rich literature on the tests for heteroskedasticity in random effects panel data models, yet to our knowledge, little work has been done for fixed effects panel data models.

In this paper, we propose a new method for testing heteroskedasticity in two-way fixed effects panel data models. We assume that the regressor X_{it} in model (1) can be correlated with μ_i alone or with ξ_t alone, or can be correlated with μ_i and ξ_t simultaneously. We further consider two important scenarios when the cross-sectional dimension N is large and the temporal dimension T is either fixed or large. For the 'large N fixed T' scenario, we first eliminate the unknown fixed effects by the within transformation technique and obtain the estimates for both the regression coefficients and the residuals. We then propose a new test statistic based on an auxiliary regression model and show that its limiting distribution is a weighted sum of independent chi-squared variables with unknown weights under the null hypothesis. Note, however, that estimating the unknown weights may generate cumulative biases and may also decrease the power of the test. To avoid estimating the unknown weights, we propose an adjusted test statistic and prove that its asymptotic distribution follows a chi-squared distribution under the null hypothesis. Furthermore, to identify the source of heteroskedasticity, we also propose two new test statistics based on the modified auxiliary regression models and establish the asymptotic distributions of the proposed test statistics. Simulation results show that our proposed tests perform well, in which they can detect the local alternatives at a parametric rate and can identify the source of heteroskedasticity accurately.

For the 'large N large T' scenario, we can also propose a test statistic by the auxiliary regression model, and show that its asymptotic distribution follows a chi-squared distribution without estimating the unknown weights and the adjustment factor under the null hypothesis. Accordingly, we also propose two new test statistics based on the modified auxiliary regression models to identify the source of heteroskedasticity. Under both the null and alternative hypotheses, we establish the asymptotic distributions of the proposed test statistics under some regularity conditions. Noting that our new tests are derived under a conditional moment framework, they are distribution free by construction. Last but not least, simulation results show that our proposed tests are robust to the misspecification of the model on the existence of time fixed effects.

The remainder of the paper is organized as follows. In Section 2, we propose our new test statistics for the scenario when N is large but T is fixed, and derive their asymptotic properties under some regularity conditions. In Section 3, we further study the asymptotic properties of the proposed test statistics when N and T are both large. In Section 4, we conduct extensive simulation studies to evaluate the efficiency of the proposed tests. In Section 5, we analyze two real data examples to demonstrate the usefulness of our proposed tests in practice. We then conclude the paper in Section 6 with discussion and future work, and provide the technical proofs in Appendix.

2. Tests for heteroskedasticity when *T* is fixed

By the conditional variance $\mathrm{Var}(\varepsilon_{it}|X_i, Z_i) = \sigma^2 g(Z_{it}^\top \theta)$ for $i = 1, \ldots, N$ and $t = 1, \ldots, T$, it is easy to see that the error term ε_{it} is heteroskedastic if $\theta \neq 0$. Thus, we consider the null hypothesis $H_0 : \theta = 0$ to perform the test for heteroskedasticity in model (1) since the function $g(\cdot)$ satisfies $g(0) = 1$ and $g'(0) \neq 0$.

Let $\eta_{it} = \varepsilon_{it} - \bar{\varepsilon}_{i.} - \bar{\varepsilon}_{.t} + \bar{\varepsilon}_{..}$, where $\bar{\varepsilon}_{i.} = (1/T)\sum_{t=1}^{T}\varepsilon_{it}$, $\bar{\varepsilon}_{.t} = (1/N)\sum_{i=1}^{N}\varepsilon_{it}$ and $\bar{\varepsilon}_{..} = (1/NT)\sum_{i=1}^{N}\sum_{t=1}^{T}\varepsilon_{it}$. By a simple calculation, we have

$$E(\eta_{it}^2|X_i, Z_i) = \sigma^2\left\{\left(1 - \frac{2}{T} - \frac{2}{N} + \frac{4}{NT}\right)g(Z_{it}^\top\theta) + \frac{1}{T^2}\left(1 - \frac{2}{N}\right)\sum_{s=1}^{T}g(Z_{is}^\top\theta) + \frac{1}{N^2}\left(1 - \frac{2}{T}\right)\sum_{j=1}^{N}g(Z_{jt}^\top\theta) + \frac{1}{N^2T^2}\sum_{j=1}^{N}\sum_{s=1}^{T}g(Z_{js}^\top\theta)\right\}. \quad (2)$$

This moment condition can be used to construct the test statistic for the null hypothesis $H_0 : \theta = 0$. For fixed T, by (2), a simple calculation yields

$$E(\eta_{it}^2|X_i, Z_i) = \sigma^2\left(1 - \frac{2}{T}\right)g(Z_{it}^\top\theta) + \frac{\sigma^2}{T^2}\sum_{s=1}^{T}g(Z_{is}^\top\theta) + o(1). \quad (3)$$

For $T \to \infty$, we have

$$E(\eta_{it}^2|X_i, Z_i) = \sigma^2 g(Z_{it}^\top\theta) + o(1). \quad (4)$$

Hence, there are differences between a fixed T and a large T. We will consider to test heteroscedasticity in model (1) with 'fixed T' in Section 2 and 'large T' in Section 3, respectively.

2.1. The test statistics

Using the Taylor expansion for the variance function $g(Z_{it}^\top\theta)$ around zero, we have

$$g(Z_{it}^\top\theta) = g(0) + g'(0)Z_{it}^\top\theta + \tfrac{1}{2}g''(c^*)(Z_{it}^\top\theta)^2, \quad (5)$$

where $c^* \in [0, Z_{it}^\top\theta]$. We first consider the case with the fixed temporal dimension T. For fixed T, by (2) and (5), a simple calculation yields

$$\begin{aligned}
E(\eta_{it}^2|X_i, Z_i) &= \sigma^2\left(1 - \frac{2}{T}\right)g(Z_{it}^\top\theta) + \frac{\sigma^2}{T^2}\sum_{s=1}^{T}g(Z_{is}^\top\theta) + o(1) \\
&= \left(1 - \frac{1}{T}\right)\sigma^2 + \sigma^2 g'(0)\left(1 - \frac{2}{T}\right)Z_{it}^\top\theta \\
&\quad + \sigma^2 g'(0)\frac{1}{T}\bar{Z}_{i.}^\top\theta + o(\|\theta^*\|) + o(1) \\
&= \left(1 - \frac{1}{T}\right)\sigma^2 + \sigma^2 g'(0)Z_{it}^{*\top}\theta + o(\|\theta^*\|) + o(1), \quad (6)
\end{aligned}$$

where θ^* is between 0 and θ, and $Z_{it}^* = (1 - (2/T))Z_{it} + (1/T)\bar{Z}_{i.}$. This motivates us to consider the following auxiliary regression:

$$\eta_{it}^2 = \alpha_1 + Z_{it}^{*\top}\alpha_2 + e_{it}, \quad (7)$$

where $E(e_{it}) = 0$ and $\mathrm{Var}(e_{it}) = \sigma_e^2 < \infty$. Since η_{it}^2 is unknown, we replace it by its consistent estimator $\hat{\eta}_{it}^2$ and obtain the following estimated auxiliary regression:

$$\hat{\eta}_{it}^2 = \alpha_1 + Z_{it}^{*\top}\alpha_2 + e_{it}, \tag{8}$$

where $\hat{\eta}_{it} = (Y_{it} - \bar{Y}_{i.} - \bar{Y}_{.t} + \bar{Y}_{..}) - (X_{it} - \bar{X}_{i.} - \bar{X}_{.t} + \bar{X}_{..})^\top \hat{\beta}$. Noting that we use $\hat{\eta}_{it}^2$ instead of the estimator of ε_{it}^2 in the auxiliary regression (8), it thus avoids the estimation of the $N+T$ redundant parameters μ_i and ξ_t. In order to construct the test statistic, we need to first estimate the vector of regression coefficients β consistently.

For the sake of descriptive convenience, we denote $Y = (Y_1^\top, \ldots, Y_N^\top)^\top$, $Y_i = (Y_{i1}, \ldots, Y_{iT})^\top$, $X = (X_1^\top, \ldots, X_N^\top)^\top$, $X_i = (X_{i1}, \ldots, X_{iT})^\top$,

$$\bar{Y}_{i.} = \frac{1}{T}\sum_{t=1}^{T} Y_{it}, \quad \bar{Y}_{.t} = \frac{1}{N}\sum_{i=1}^{N} Y_{it}, \quad \bar{Y}_{..} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} Y_{it},$$
$$\bar{X}_{i.} = \frac{1}{T}\sum_{t=1}^{T} X_{it}, \quad \bar{X}_{.t} = \frac{1}{N}\sum_{i=1}^{N} X_{it}, \quad \bar{X}_{..} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} X_{it}.$$

Then, model (1) can be transformed to the following form:

$$Y_{it} - \bar{Y}_{i.} - \bar{Y}_{.t} + \bar{Y}_{..} = (X_{it} - \bar{X}_{i.} - \bar{X}_{.t} + \bar{X}_{..})^\top \beta + \varepsilon_{it} - \bar{\varepsilon}_{i.} - \bar{\varepsilon}_{.t} + \bar{\varepsilon}_{..}. \tag{9}$$

Let $Q = I_{NT} - I_N \otimes J_T - I_T \otimes J_N + J_{NT}$, where I_m is an $m \times m$ identity matrix, $J_m = (1/m)1_m 1_m^\top$ with 1_m being an m-dimensional vector with all elements equal to one. Thus, in matrix form, model (9) can be written as

$$QY = QX^\top \beta + Q\varepsilon, \tag{10}$$

where $\varepsilon = (\varepsilon_1^\top, \ldots, \varepsilon_N^\top)^\top$ and $\varepsilon_i = (\varepsilon_{i1}, \ldots, \varepsilon_{iT})^\top$. If we consider the ordinary least squares estimation for the above model, then by Baltagi [4], the within estimator of β is

$$\hat{\beta} = (X^\top QX)^{-1} X^\top QY.$$

Once we get the estimator $\hat{\beta}$ of the regression coefficient vector β, we can obtain the estimated auxiliary regression (8). By (6) and (8), testing the null hypothesis $H_0 : \theta = 0$ is equivalent to testing $H_0 : \alpha_2 = 0$. Let R_1^2 be the determination coefficient of the auxiliary regression (8), and define the test statistic as

$$\tilde{L}_1 = NTR_1^2. \tag{11}$$

Intuitively, if $\alpha_2 \neq 0$, the test statistic $\tilde{L}_1$ will tend to take large positive values. This suggests to reject the null hypothesis H_0 when $\tilde{L}_1$ is larger than a certain critical value. Letting t_0 be the observed value of $\tilde{L}_1$, the p-value of the test is defined as

$$p_0 = P(\tilde{L}_1 > t_0 | H_0),$$

which denotes the probability of the event $\{\tilde{L}_1 > t_0\}$. For a given significance level α_0, the null hypothesis H_0 will be rejected if $p_0 \leq \alpha_0$.

2.2. Asymptotic properties

For simplicity, let C denote a positive constant that may be different at each appearance throughout this paper. Let also $\sigma^2_{\eta^2} = \mathrm{Var}(\eta^2_{it})$, $Z = (Z_1^\top, \ldots, Z_n^\top)^\top$, $M = I_{NT} - (1/NT)1_{NT}1_{NT}^\top$, $M_G = I_N \otimes (I_T - (1/T)1_T 1_T^\top)$, $Z_0 = MZ$ and $Z_G = M_G Z$. To derive the asymptotic properties of our proposed tests, we need some regularity conditions.

- (C1) $\{(X_i, \varepsilon_i), i \geq 1\}$ is a sequence of independent and identically distributed random variables. Furthermore, we assume that $E(\|X_{it}\|^{2(2+\gamma)}) \leq C < \infty$ for some $\gamma > 0$.
- (C2) $E(\varepsilon_{it}|X_i, Z_i, \varepsilon_{i,t-1}, \varepsilon_{i,t-2}, \ldots) = 0$, almost surely, for all $1 \leq i \leq N$ and $1 \leq t \leq T$.
- (C3) $E(\varepsilon^2_{it}|X_i, Z_i, \varepsilon_{i,t-1}, \varepsilon_{i,t-2}, \ldots) = E(\varepsilon^2_{it}|X_i, Z_i)$, almost surely, for all $1 \leq i \leq N$ and $1 \leq t \leq T$.
- (C4) For each $i = 1, \ldots, N$ and $t = 1, \ldots, T$, we have $E[(\varepsilon^2_{it} - \sigma^2 g(Z_{it}^\top \theta))^2|X_i, Z_i] = \varpi < \infty$.
- (C5) Assume that the following matrices are all full rank, where $\Omega_i = \mathrm{Var}(\eta^2_i|X_i, Z_i)$ and

$$\Sigma_1 = \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^N E(Z_{0i}^{*\top}\Omega_i Z_{0i}^*), \qquad \Sigma_1^* = \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^N E(Z_{0i}^{*\top} Z_{0i}^*),$$

$$\Sigma_2 = \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^N E(Z_{Gi}^{*\top}\Omega_i Z_{Gi}^*), \qquad \Sigma_2^* = \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^N E(Z_{Gi}^{*\top} Z_{Gi}^*),$$

$$\Sigma_3 = \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^N E(Z_{Hi}^{*\top}\Omega_i Z_{Hi}^*), \qquad \Sigma_3^* = \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^N E(Z_{Hi}^{*\top} Z_{Hi}^*),$$

$$\Sigma_4 = \lim_{N,T\to\infty} \frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T E(Z_{0it}Z_{0it}^\top), \quad \Sigma_5 = \lim_{N,T\to\infty} \frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T E(Z_{Git}Z_{Git}^\top),$$

and

$$\Sigma_6 = \lim_{N,T\to\infty} \frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T E(Z_{Hit}Z_{Hit}^\top).$$

Theorem 2.1: *Suppose that regularity conditions* (C1)–(C5) *hold. Under* $H_0: \theta = 0$, *as* $N \to \infty$, *we have*

$$\tilde{L}_1 \xrightarrow{\mathcal{L}} \phi_1\chi^2_{1,1} + \cdots + \phi_k\chi^2_{1,k},$$

where $\chi^2_{1,j}$ $(j = 1, \ldots, k)$ *are independent* χ^2_1 *variables,* ϕ_j *are the eigenvalues of* $\Sigma_1^{*-1}\Sigma_1/\sigma^2_{\eta^2}$, *and* Σ_1 *and* Σ_1^* *are defined in condition* (C5).

To apply Theorem 2.1 to test heteroskedasticity, we need to first estimate the unknown weights ϕ_j consistently since the accuracy of this approximation depends on the estimated values of ϕ_j. Nevertheless, it is often difficult to obtain the accuracy estimates of the weights ϕ_j such that the power of test can be maintained. To solve this problem, we consider the following adjusted test statistic and show that the asymptotic distribution of the adjusted test statistic is exactly a chi-squared with k degrees of freedom. Let $\hat{\eta}^2 = (\hat{\eta}_1^{2\top}, \ldots, \hat{\eta}_N^{2\top})^\top$, $\hat{\eta}_i^2 = (\hat{\eta}_{i1}^2, \ldots, \hat{\eta}_{iT}^2)^\top$, $Z^* = (Z_1^{*\top}, \ldots, Z_N^{*\top})^\top$, $Z_i^* = (Z_{i1}^*, \ldots, Z_{iT}^*)^\top$, $Z_0^* = MZ^*$, $\hat{\eta}_0^2 = M\hat{\eta}^2$, and

$$\hat{\Sigma}_1 = \frac{1}{N}\sum_{i=1}^N Z_{0i}^{*\top}\hat{\eta}_{0i}^2\hat{\eta}_{0i}^{2\top}Z_{0i}^*, \quad \hat{\Sigma}_1^* = \frac{1}{N}Z^{*\top}MZ^*, \quad \hat{\sigma}_{\eta^2}^2 = \frac{1}{NT}\hat{\eta}^{2\top}M\hat{\eta}^2.$$

Following Rao and Scott [26], the distribution of $\rho\sum_{j=1}^k \phi_j\chi_{1,j}^2$ can be approximated by χ_k^2, where $\rho = k/\mathrm{tr}\{\Sigma_1^{*-1}\Sigma_1/\sigma_{\eta^2}^2\}$. This implies that the asymptotic distribution of $\tilde{\rho}\tilde{L}_1$ can be approximated by χ_k^2 by Theorem 2.1 and the consistency of $\hat{\Sigma}_1$, $\hat{\Sigma}_1^*$ and $\hat{\sigma}_{\eta^2}^2$, where $\tilde{\rho} = k/\mathrm{tr}\{\hat{\Sigma}_1^{*-1}\hat{\Sigma}_1/\hat{\sigma}_{\eta^2}^2\}$. However the accuracy of this approximation still depends on the values of ϕ_j. Next, we give an adjusted test statistic whose asymptotic distribution is exactly a chi-squared with k degrees of freedom. Note that $\tilde{\rho}$ can be written as

$$\tilde{\rho} = \frac{\mathrm{tr}\{\hat{\Sigma}_1^{-1}\hat{\Sigma}_1\}}{\mathrm{tr}\{\hat{\Sigma}_1^{*-1}\hat{\Sigma}_1/\hat{\sigma}_{\eta^2}^2\}}.$$

We replace $\hat{\Sigma}_1$ in $\tilde{\rho}$ by $\hat{B} = (1/N)(Z^{*\top}M\hat{\eta}^2)(Z^{*\top}M\hat{\eta}^2)^\top$ and get a new adjustment factor as

$$\hat{\rho} = \frac{\mathrm{tr}\{\hat{\Sigma}_1^{-1}\hat{B}\}}{\mathrm{tr}\{\hat{\Sigma}_1^{*-1}\hat{B}/\hat{\sigma}_{\eta^2}^2\}}.$$

Finally, the adjusted test statistic is defined as $L_1 = \hat{\rho}\tilde{L}_1 = NT\hat{\rho}R_1^2$.

Theorem 2.2: *Suppose that regularity conditions* (C1)–(C5) *hold. As* $N \to \infty$, *then*

(1) *under the null hypothesis* $H_0 : \theta = 0$, *we have* $L_1 \xrightarrow{\mathcal{L}} \chi_k^2$;
(2) *under the alternative hypothesis* $H_1 : \theta = N^{-1/2}\delta$, L_1 *follows asymptotically the noncentral* $\chi^2(k, \varsigma_1)$ *distribution with k degrees of freedom and the noncentral parameter* $\varsigma_1 = \sigma^4 g'(0)^2\delta^\top\Sigma_1^*\Sigma_1^{-1}\Sigma_1^*\delta$.

Note that, compared with the simple cross-sectional or time-series model, the conditional variances of panel data models may vary between the cross-sectional levels, within the cross-sectional levels, or both. To identify the source of heteroskedasticity, by Juhl and Sosa-Escudero [16] we construct the following auxiliary regressions:

$$\begin{aligned}\hat{\eta}_{it}^2 - \bar{\hat{\eta}}_{i.}^2 &= \alpha_1 + (Z_{it}^* - \bar{Z}_{i.}^*)^\top\alpha_2 + e_{it} \\ &= \alpha_1 + \left(1 - \frac{2}{T}\right)(Z_{it} - \bar{Z}_{i.})^\top\alpha_2 + e_{it} \qquad (12)\end{aligned}$$

and

$$\hat{\eta}_{it}^2 - \bar{\hat{\eta}}_{.t}^2 = \alpha_1 + (Z_{it}^* - \bar{Z}_{.t}^*)^\top\alpha_2 + e_{it}, \qquad (13)$$

where $\bar{\hat{\eta}}_{i.}^2 = (1/T)\sum_{s=1}^T \hat{\eta}_{is}^2$ and $\bar{\hat{\eta}}_{.t}^2 = (1/T)\sum_{j=1}^N \hat{\eta}_{jt}^2$.

Let R_2^2 and R_3^2 be the coefficients of determination of the auxiliary regression models (12) and (13) respectively, and we define two adjustment factors as

$$\hat{\varrho}_1 = \frac{\mathrm{tr}\{\hat{\Sigma}_2^{-1}\hat{B}_G\}}{\mathrm{tr}\{\hat{\Sigma}_2^{*-1}\hat{B}_G/\hat{\sigma}^2_{\eta^2}\}} \quad \text{and} \quad \hat{\varrho}_2 = \frac{\mathrm{tr}\{\hat{\Sigma}_3^{-1}\hat{B}_H\}}{\mathrm{tr}\{\hat{\Sigma}_3^{*-1}\hat{B}_H/\hat{\sigma}^2_{\eta^2}\}},$$

where $\hat{\Sigma}_2 = (1/N)\sum_{i=1}^N Z_{Gi}^{*\top}\hat{\eta}^2_{Gi}\hat{\eta}^{2\top}_{Gi}Z^*_{Gi}$, $\hat{\Sigma}_2^* = (1/N)Z^{*\top}M_GZ^*$, $Z_G^* = M_GZ^*$, $\hat{\eta}^2_G = M_G\hat{\eta}^2$, $\hat{B}_G = (1/N)(Z^{*\top}M_G\hat{\eta}^2)(Z^{*\top}M_G\hat{\eta}^2)^\top$, and $\hat{\Sigma}_3 = (1/N)\sum_{i=1}^N Z_{Hi}^{*\top}\hat{\eta}^2_{Hi}\hat{\eta}^{2\top}_{Hi} Z^*_{Hi}$, $\hat{\Sigma}_3^* = (1/N)\tilde{Z}^{*\top}M_H\tilde{Z}^*$, $M_H = I_T \otimes (I_N - (1/N)1_N1_N^\top)$, $Z_H^* = M_H\tilde{Z}^*$, $\hat{\eta}^2_H = M_H\tilde{\hat{\eta}}^2$, $\hat{B}_H = (1/N)(\tilde{Z}^{*\top}M_H\tilde{\hat{\eta}}^2)(\tilde{Z}^{*\top}M_H\tilde{\hat{\eta}}^2)^\top$, $\tilde{\hat{\eta}}^2 = (\tilde{\hat{\eta}}_1^{2\top},\ldots,\tilde{\hat{\eta}}_T^{2\top})^\top$, $\tilde{\hat{\eta}}_t^2 = (\hat{\eta}^2_{1t},\ldots,\hat{\eta}^2_{Nt})^\top$, $\tilde{Z}^* = (\tilde{Z}_1^*,\ldots,\tilde{Z}_T^*)^\top$ and $\tilde{Z}_t^* = (Z_{1t}^*,\ldots,Z_{Nt}^*)^\top$. We then define the test statistics, respectively, as

$$L_2 = NT\hat{\varrho}_1R_2^2 \quad \text{and} \quad L_3 = NT\hat{\varrho}_2R_3^2.$$

Theorem 2.3: *Suppose that regularity conditions* (C1)–(C5) *hold. As $N \to \infty$, then*

(1) *under the null hypothesis $H_0 : \theta = 0$, we have $L_2 \xrightarrow{\mathcal{L}} \chi_k^2$;*
(2) *under the alternative hypothesis $H_1 : \theta = N^{-1/2}\delta$, L_2 follows asymptotically the noncentral $\chi^2(k, \varsigma_2)$ distribution with k degrees of freedom and the noncentral parameter $\varsigma_2 = \sigma^4 g'(0)^2(1-(2/T))^4\delta^\top\Sigma_2^*\Sigma_2^{-1}\Sigma_2^*\delta$, where Σ_2 and Σ_2^* are defined in condition* (C5).

Theorem 2.4: *Suppose that regularity conditions* (C1)–(C5) *hold. As $N \to \infty$, then*

(1) *under the null hypothesis $H_0 : \theta = 0$, we have $L_3 \xrightarrow{\mathcal{L}} \chi_k^2$;*
(2) *under the alternative hypothesis $H_1 : \theta = N^{-1/2}\delta$, L_3 follows asymptotically the noncentral $\chi^2(k, \varsigma_3)$ distribution with k degrees of freedom and the noncentral parameter $\varsigma_3 = \sigma^4 g'(0)^2\delta^\top\Sigma_3^*\Sigma_3^{-1}\Sigma_3^*\delta$, where Σ_3 and Σ_3^* are defined in condition* (C5).

Unlike the L_1 test, L_2 will be rejected only if heteroskedasticity occurs beyond the between level, and L_3 will be rejected only if heteroskedasticity occurs beyond the within level. Thus in practice, we can use the test statistics L_1, L_2 and L_3 to distinguish the following four relevant cases: homoskedasticity at both levels, heteroskedasticity at the between level only, heteroskedasticity at the within level only, and heteroskedasticity at both levels. That is, if L_1 does not reject the null hypothesis, then homoskedasticity is concluded to be related to Z_{it}; if L_1 rejects the null hypothesis but L_2 does not, then we conclude that heteroskedasticity may exist at the between level only; if L_1 rejects the null hypothesis but L_3 does not, then heteroskedasticity may exist at the within level only; and finally, if all the three tests are rejected, then it implies that heteroskedasticity may potentially exists at both levels.

3. Tests for heteroskedasticity when *T* is large

Although Theorems 2.1–2.4 can be used to test for heteroscedasticity and identify the source of heteroscedasticity for model (1) with fixed T, there are some difficulties remaining to be solved. For example, we need to estimate the weights ϕ_j in Theorem 2.1 and the adjustment factors in Theorems 2.2–2.4 consistently. In order to show the standard asymptotic distribution of test statistic and avoid estimating the weights ϕ_j and the adjustment factors, we consider the scenario of large N and large T in this section.

If $T \to \infty$, invoking (2) and Taylor expansion for $g(\cdot)$, we have

$$\begin{aligned} E(\eta_{it}^2|X_i, Z_i) &= \sigma^2 g(Z_{it}^\top \theta) + o(1) \\ &= \sigma^2 + \sigma^2 g'(0) Z_{it}^\top \theta + o(\|\theta^*\|) + o(1), \end{aligned} \tag{14}$$

where θ^* is between 0 and θ. In order to derive the test, we consider the following auxiliary regression:

$$\hat{\eta}_{it}^2 = \alpha_1 + Z_{it}^\top \alpha_2 + e_{it}. \tag{15}$$

Following the same idea, we define the test statistic: $L_4 = NTR_4^2$, where R_4^2 is the determination coefficient of the auxiliary regression (15).

Theorem 3.1: *Suppose that regularity conditions* (C1)–(C5) *hold. As* $N \to \infty$ *and* $T \to \infty$, *then*

(1) *under the null hypothesis* $H_0 : \theta = 0$, *we have* $L_4 \xrightarrow{\mathcal{L}} \chi_k^2$;
(2) *under the alternative hypothesis* $H_1 : \theta = (NT)^{-1/2}\delta$, L_4 *follows asymptotically the noncentral* $\chi^2(k, \varsigma_4)$ *distribution with k degrees of freedom and the noncentral parameter* $\varsigma_4 = \sigma^4 g'(0)^2 \delta^\top \Sigma_4 \delta / \sigma_{\eta^2}^2$, *where* Σ_4 *is defined in condition* (C5).

Theorem 3.1 shows that the test statistic L_4 is able to detect the alternatives approaching the null at the parameter rate. Compare with Theorem 2.1 and Theorem 2.2, it is easy to see that Theorem 3.1 provides the standard asymptotic χ_k^2 distribution of L_4 under null hypothesis H_0 as $T \to \infty$. This avoids estimating the weights ϕ_j in Theorem 2.1 and the adjustment factor ρ in Theorem 2.2 such that it can improve the power of test.

In order to identify the source of heteroskedasticity for the scenario with large N and large T, we construct the following auxiliary regression models:

$$\hat{\eta}_{it}^2 - \bar{\hat{\eta}}_{i.}^2 = \alpha_1 + (Z_{it} - \bar{Z}_{i.})^\top \alpha_2 + e_{it} \tag{16}$$

and

$$\hat{\eta}_{it}^2 - \bar{\hat{\eta}}_{.t}^2 = \alpha_1 + (Z_{it} - \bar{Z}_{.t})^\top \alpha_2 + e_{it}. \tag{17}$$

Let R_5^2 and R_6^2 be the coefficient of determination of the auxiliary regression models (16) and (17) respectively, we then define two new test statistics:

$$L_5 = NTR_5^2 \quad \text{and} \quad L_6 = NTR_6^2.$$

Theorem 3.2: *Suppose that regularity conditions* (C1)–(C5) *hold. As* $N \to \infty$ *and* $T \to \infty$, *then*

(1) *under the null hypothesis* $H_0 : \theta = 0$, *we have* $L_5 \xrightarrow{\mathcal{L}} \chi^2_k$;
(2) *under the alternative hypothesis* $H_1 : \theta = (NT)^{-1/2}\delta$, L_5 *follows asymptotically the noncentral* $\chi^2(k, \varsigma_5)$ *distribution with k degrees of freedom and the noncentral parameter* $\varsigma_5 = \sigma^4 g'(0)^2\delta^\top \Sigma_5\delta/\sigma^2_{\eta^2}$, *where* Σ_5 *is defined in condition* (C5).

Theorem 3.3: *Suppose that regularity conditions* (C1)–(C5) *hold. As* $N \to \infty$ *and* $T \to \infty$, *then*

(1) *under the null hypothesis* $H_0 : \theta = 0$, *we have* $L_6 \xrightarrow{\mathcal{L}} \chi^2_k$;
(2) *under the alternative hypothesis* $H_1 : \theta = (NT)^{-1/2}\delta$, L_6 *follows asymptotically the noncentral* $\chi^2(k, \varsigma_6)$ *distribution with k degrees of freedom and the noncentral parameter* $\varsigma_6 = \sigma^4 g'(0)^2\delta^\top \Sigma_6\delta/\sigma^2_{\eta^2}$, *where* Σ_6 *is defined in condition* (C5).

Note that, since $\Sigma_4 - \Sigma_5$ and $\Sigma_4 - \Sigma_6$ are two positive semi-definite matrices, the local power of L_4 is larger than those of L_5 and L_6 when both sources of variability are present.

Due to the presence of the unobserved $N+T$ redundant parameters μ_i and ξ_t, the idiosyncratic error ε_{it} cannot be estimated accurately by the within transformation residuals $\hat{\eta}_{it} = \varepsilon_{it} - \bar{\varepsilon}_{i.} - \bar{\varepsilon}_{.t} + \bar{\varepsilon}_{..} + (X_{it} - \bar{X}_{i.} - \bar{X}_{.t} + \bar{X}_{..})^\top(\beta - \hat{\beta})$. The second term $\bar{\varepsilon}_{i.} = (1/T)\sum_{t=1}^{T}\varepsilon_{it}$ is created by the within transformation to wipe out the unobserved individual effects μ_i, and it has the order of $O_p((1/\sqrt{T}))$. Hence, the accuracy of the within transformation residuals depends on the temporal dimension T. On the other hand, in order to avoid estimating μ_i and ξ_t, we construct the auxiliary regression models with η^2_{it} instead of ε^2_{it}, but it brings a new problem, that is, the conditional correlations between η^2_{it} and η^2_{is} are not zero. By a simple calculation, we have $\mathrm{Cov}(\eta^2_{it}, \eta^2_{is}|X_i, Z_i) = O(T^{-2})$, $\forall t \neq s$. The correlations can be asymptotically ignored if $T \to \infty$. For fixed T, however, the correlations can not be ignored. Consequently, the test statistics have different asymptotic distributions for fixed T and large T, respectively.

One problem that needs to be solved in practice is when to use the fixed T test and the large T test. For this purpose, a simulation example is used to address this issue in Section 4. From Figure 1, it shows that the test statistic L_4 achieves correct size as T increases larger than 30 and has size distortions when T is small. Therefore, we recommend using the large T test when the temporal dimension T is larger than 30.

4. Simulation studies

In this section, we conduct simulation studies to assess the efficiency of our proposed tests. The data are generated from the following model:

$$Y_{it} = X_{it}\beta + \mu_i + \xi_t + \varepsilon_{it}, \quad i = 1, \ldots, N, \quad t = 1, \ldots, T, \tag{18}$$

where $\beta = 2$. The fixed effects μ_i and ξ_t are generated by

$$\mu_i \sim N(0, 2^2),\ i = 1, \ldots, N, \quad \text{and} \quad \xi_t \sim N(0, 5^2),\ t = 1, \ldots, T.$$

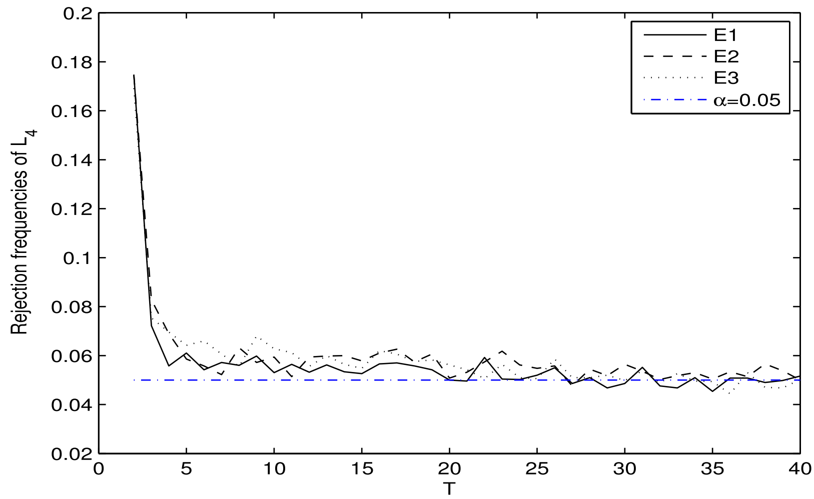

Figure 1. The empirical rejection frequencies of L_4 with $N = 40$ and T varying.

The regressor is correlated with μ_i and ξ_t, and is generated according to

$$X_{it} = 2 + \mu_i + \xi_t + 0.2 \times \mu_i \times \xi_t + e_{it}, \tag{19}$$

where $e_{it} \sim N(0, 1)$.

Throughout the simulations, the sample size is set to $N = 40$, 70 and 100. For fixed T, we set $T = 4$, 8 and 12, and for large T, we set $T = 30$, 50 and 70, respectively. For each setting, we repeat the simulation 1000 times. In our simulation, we consider the following six models:

(E1) $\varepsilon_{it} \sim N(0, 1)$;
(E2) $\varepsilon_{it} \sim t(2)$;
(E3) $\varepsilon_{it} \sim \chi^2_3$;
(E4) $\varepsilon_{it} = \nu_{it} \exp(5\delta X_{it})$;
(E5) $\varepsilon_{it} = \nu_{it} \exp(5\delta \mu_i)$;
(E6) $\varepsilon_{it} = \nu_{it} \exp(5\delta \xi_t)$;

where ν_{it} is standard normal and independent of X_{it}, μ_i and ξ_t.

Cases (E1)–(E3) correspond to the hypothesis of homoskedasticity, with (E1) for the standard normal error, and (E2) and (E3) for nonnormal errors. We take $Z_{it} = X_{it}$ and calculate the empirical rejection frequencies with the significance level $\alpha = 0.05$. The simulation results are presented in Tables 1 and 2, respectively.

From the results of the two tables, we have the following observations. (1) When the auxiliary regression models are correct specified, for all cases, the empirical rejection frequencies of our proposed tests are all close to the significance level 0.05, and the proposed method has the better performance with the sample size increasing. (2) The performance of the proposed tests are robust to the nonnormal distributions including the t-distribution and the chi-squared distribution. (3) For large T, when the auxiliary regression models

Table 1. Empirical rejection frequencies for fixed T in model (18).

		L_1			L_2			L_3			L_4		
N	T	$N(0,1)$	$t(2)$	χ^2_3	$N(0,1)$	$t(2)$	χ^2_3	$N(0,1)$	$t(2)$	χ^2_3	$N(0,1)$	$t(2)$	χ^2_3
40	4	0.0450	0.0490	0.0660	0.0770	0.0810	0.0840	0.0770	0.0780	0.0850	0.0580	0.0700	0.0790
	8	0.0610	0.0550	0.0640	0.0630	0.0660	0.0590	0.0690	0.0600	0.0710	0.0740	0.0640	0.0730
	12	0.0480	0.0540	0.0630	0.0550	0.0570	0.0580	0.0560	0.0590	0.0620	0.0620	0.0590	0.0660
70	4	0.0580	0.0430	0.0440	0.0680	0.0720	0.0750	0.0780	0.0810	0.0790	0.0720	0.0650	0.0670
	8	0.0470	0.0620	0.0480	0.0550	0.0560	0.0600	0.0560	0.0480	0.0630	0.0380	0.0700	0.0540
	12	0.0500	0.0610	0.0510	0.0530	0.0540	0.0570	0.0530	0.0580	0.0580	0.0580	0.0660	0.0680
100	4	0.0430	0.0440	0.0440	0.0660	0.0740	0.0730	0.0690	0.0760	0.0750	0.0680	0.0350	0.0710
	8	0.0490	0.0540	0.0560	0.0540	0.0600	0.0590	0.0560	0.0610	0.0610	0.0450	0.0680	0.0710
	12	0.0520	0.0470	0.0540	0.0520	0.0540	0.0580	0.0520	0.0470	0.0540	0.0570	0.0570	0.0580

Table 2. Empirical rejection frequencies for large T in model (18).

		L_4			L_5			L_6			L_1		
N	T	$N(0,1)$	$t(2)$	χ^2_3	$N(0,1)$	$t(2)$	χ^2_3	$N(0,1)$	$t(2)$	χ^2_3	$N(0,1)$	$t(2)$	χ^2_3
40	30	0.0480	0.0550	0.0530	0.0540	0.0570	0.0620	0.0570	0.0500	0.0600	0.0480	0.0550	0.0540
	50	0.0540	0.0560	0.0470	0.0480	0.0630	0.0480	0.0520	0.0510	0.0540	0.0540	0.0560	0.0470
	70	0.0500	0.0480	0.0510	0.0520	0.0440	0.0590	0.0530	0.0460	0.0490	0.0510	0.0480	0.0510
70	30	0.0510	0.0540	0.0530	0.0450	0.0520	0.0580	0.0490	0.0530	0.0650	0.0530	0.0540	0.0540
	50	0.0560	0.0580	0.0540	0.0570	0.0530	0.0460	0.0530	0.0470	0.0570	0.0560	0.0570	0.0540
	70	0.0520	0.0550	0.0510	0.0530	0.0510	0.0530	0.0520	0.0540	0.0470	0.0520	0.0550	0.0510
100	30	0.0450	0.0470	0.0560	0.0470	0.0710	0.0480	0.0550	0.0600	0.0620	0.0430	0.0470	0.0560
	50	0.0480	0.0570	0.0500	0.0540	0.0560	0.0470	0.0560	0.0460	0.0570	0.0480	0.0570	0.0510
	70	0.0490	0.0530	0.0470	0.0480	0.0580	0.0550	0.0520	0.0580	0.0560	0.0490	0.0530	0.0470

are misspecified, L_4 and L_1 have very similar rejection rates. This indicates that the difference can be neglected, especially when $T \to \infty$. (4) For fixed T, however, if the auxiliary regression models are misspecified, the difference between L_1 and L_4 cannot be neglected.

Further, in order to assess the sensitivity of the proposed test statistics to the size of temporal dimension T, we fix $N = 40$ and consider 5000 simulations for each $T \in \{2, 3, \ldots, 40\}$. Simulation results are presented graphically in Figure 1. It shows that the test statistic L_4 has size distortions for small T, and achieves the correct size as T gets larger than 30.

Case (E4) corresponds to heteroskedasticity at both within and between levels. Case (E5) corresponds to heteroskedasticity at the between level only. Case (E6) corresponds to heteroskedasticity at the within level only. The estimated power function curves of cases (E4)–(E6) with the significance level $\alpha = 0.05$ are displayed in Figures 2–4 respectively. If $\delta = 0$, then cases (E4)–(E6) reduce to the null hypothesis of full homoskedasticity.

From Figure 2 we can see that, for case (E4), when the null hypothesis holds ($\delta = 0$), the size of our tests is close to the significance level 0.05. This demonstrates that the proposed procedures give the right level of testing. When the alternative hypothesis is true ($\delta > 0$), the power functions increase rapidly as δ increases. For case (E5), since all of the heteroskedastic variation comes from the different cross-sectional units, from Figure 3 we can see that, the proposed L_1 and L_4 tests reject for this type of alternative, while the L_2 and L_5 tests have the correct size. For case (E6), since all of the heteroskedastic variation comes from the within levels, from Figure 4 we can see that, the proposed L_1 and L_4 tests reject for this type of alternative, while the L_3 and L_6 tests have the correct size. These results show that the proposed test statistics perform satisfactorily.

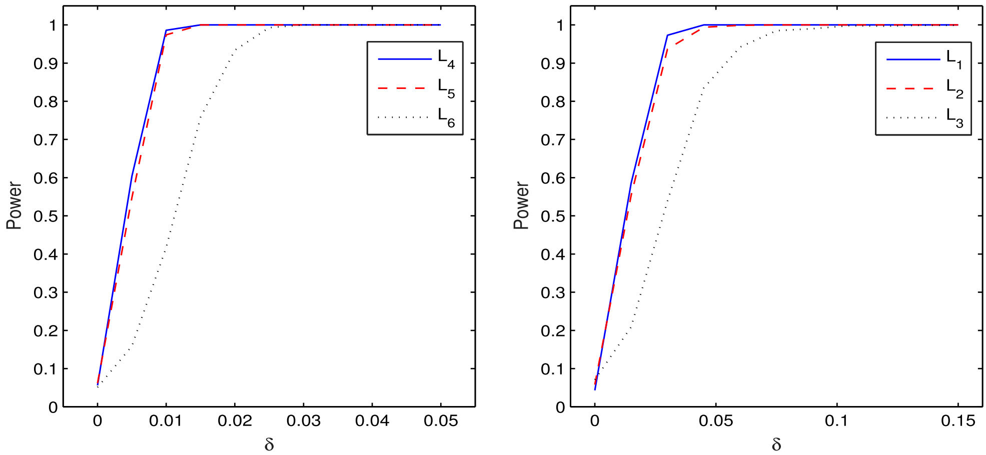

Figure 2. The power curve for case (E4), the left plot for $N = 70$ and $T = 50$, the right plot for $N = 70$ and $T = 8$.

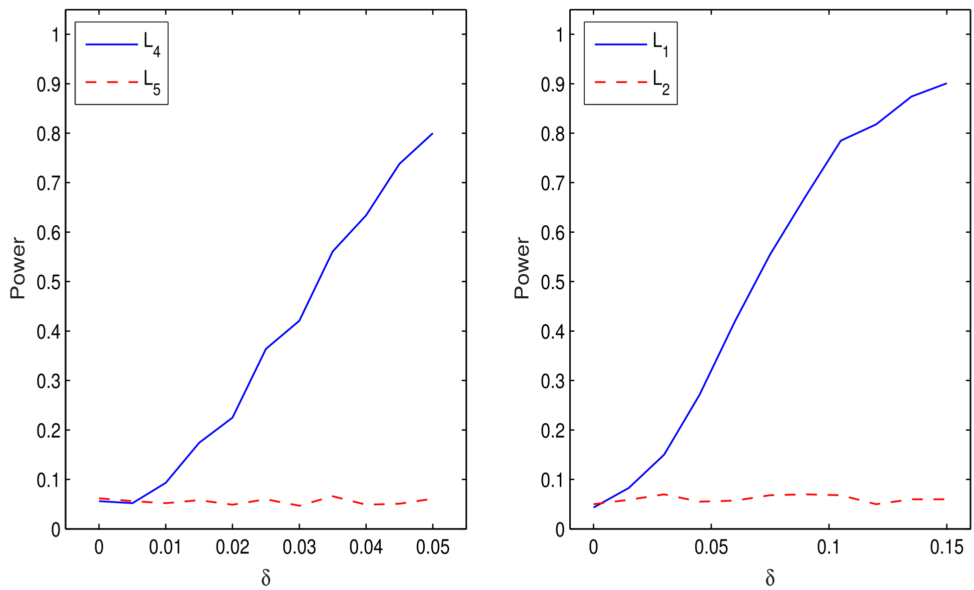

Figure 3. The power curve for case (E5), the left plot for $N = 70$ and $T = 50$, the right plot for $N = 70$ and $T = 8$.

Next, we show that our proposed tests are still feasible when the time fixed effects do not exist, while the tests of Juhl and Sosa-Escudero [16] for one-way fixed effects panel data models are invalid when the time fixed effects exist. For comparison, we further consider the following model that has been used by Juhl and Sosa-Escudero [16]:

$$Y_{it} = X_{it}\beta + \mu_i + \varepsilon_{it}, \quad i = 1, \ldots, N, \quad t = 1, \ldots, T, \tag{20}$$

where $\beta = 1$, $\mu_i \sim N(1, 1)$, and the regressor $X_{it} = \mu_i + e_{it}$ with $e_{it} \sim N(0, 1)$. We denote by LM the test statistic based on the artificial regression from Juhl and Sosa-Escudero [16].

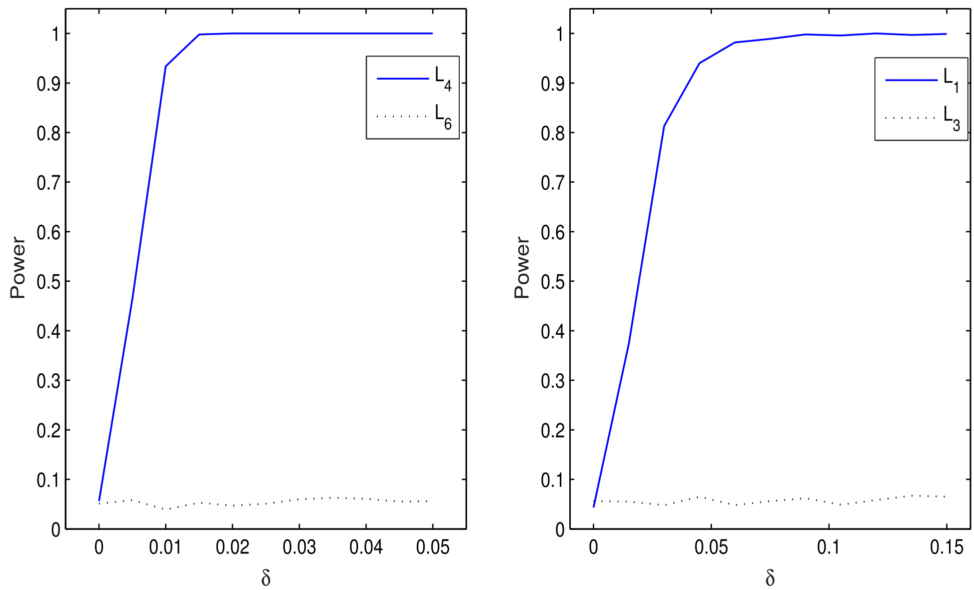

Figure 4. The power curve for case (E6), the left plot for $N = 70$ and $T = 50$, the right plot for $N = 70$ and $T = 8$.

Table 3. Empirical rejection frequencies for fixed T in model (20).

		L_1			LM		
N	T	$N(0,1)$	$t(2)$	χ_3^2	$N(0,1)$	$t(2)$	χ_3^2
40	4	0.0750	0.0910	0.0890	0.0690	0.0840	0.0720
	8	0.0570	0.0560	0.0710	0.0580	0.0540	0.0640
	12	0.0560	0.0490	0.0630	0.0550	0.0470	0.0580
70	4	0.0740	0.0870	0.0840	0.0700	0.0780	0.0830
	8	0.0580	0.0640	0.0700	0.0550	0.0600	0.0670
	12	0.0530	0.0630	0.0570	0.0470	0.0620	0.0530
100	4	0.0730	0.0870	0.0820	0.0640	0.0800	0.0760
	8	0.0490	0.0710	0.0540	0.0510	0.0650	0.0530
	12	0.0520	0.0600	0.0540	0.0520	0.0590	0.0540

The simulation results reported in Tables 3–6 show that the LM tests in Juhl and Sosa-Escudero [16] perform well only when we have prior information on the absence of the time fixed effects. When the time effects do exist, however, the empirical rejection frequencies of the tests become too large, that is, the tests cannot maintain the significance level. In contrast, our proposed tests are robust to these misspecifications and hold the significance level.

5. Real data examples

We apply our proposed methods to analyze two real data examples. One is the well-known Cigar data [2,4,7,33], and the other is the US Public Capital data.

Example 5.1 (Public capital data): The public capital data were considered in Baltagi [4], Baltagi and Pinnoi [9], Munnell [25], Su *et al.* [28] and Wu and Zhu [32]. To investigate the

Table 4. Empirical rejection frequencies for large T in model (20).

N	T	L_4 $N(0,1)$	L_4 $t(2)$	L_4 χ_3^2	LM $N(0,1)$	LM $t(2)$	LM χ_3^2
40	30	0.0530	0.0570	0.0420	0.0450	0.0580	0.0470
	50	0.0470	0.0550	0.0460	0.0480	0.0550	0.0480
	70	0.0530	0.0510	0.0440	0.0550	0.0510	0.0450
70	30	0.0420	0.0670	0.0510	0.0440	0.0670	0.0480
	50	0.0700	0.0420	0.0490	0.0680	0.0430	0.0500
	70	0.0470	0.0570	0.0520	0.0500	0.0560	0.0530
100	30	0.0440	0.0710	0.0590	0.0460	0.0690	0.0570
	50	0.0460	0.0560	0.0450	0.0490	0.0550	0.0470
	70	0.0540	0.0600	0.0480	0.0530	0.0600	0.0520

Table 5. Empirical rejection frequencies for fixed T in model (18).

N	T	L_1 $N(0,1)$	L_1 $t(2)$	L_1 χ_3^2	LM $N(0,1)$	LM $t(2)$	LM χ_3^2
40	4	0.0440	0.0490	0.0660	0.4130	0.1960	0.2240
	8	0.0610	0.0550	0.0640	0.4150	0.1690	0.2820
	12	0.0480	0.0540	0.0630	0.4940	0.1920	0.3370
70	4	0.0580	0.0430	0.0440	0.5250	0.2160	0.3210
	8	0.0470	0.0620	0.0480	0.5590	0.2420	0.4320
	12	0.0500	0.0610	0.0510	0.6300	0.1980	0.4310
100	4	0.0430	0.0440	0.0440	0.5820	0.2500	0.3690
	8	0.0550	0.0540	0.0560	0.6490	0.2520	0.4810
	12	0.0520	0.0470	0.0540	0.6760	0.2190	0.5290

Table 6. Empirical rejection frequencies for large T in model (18).

N	T	L_4 $N(0,1)$	L_4 $t(2)$	L_4 χ_3^2	LM $N(0,1)$	LM $t(2)$	LM χ_3^2
40	30	0.0480	0.0550	0.0530	0.4860	0.1660	0.6900
	50	0.0540	0.0560	0.0470	0.5860	0.1580	0.5940
	70	0.0500	0.0480	0.0510	0.6050	0.1390	0.7720
70	30	0.0510	0.0540	0.0530	0.5930	0.1890	0.5670
	50	0.0560	0.0580	0.0540	0.6400	0.1740	0.5810
	70	0.0520	0.0550	0.0510	0.7010	0.1570	0.6380
100	30	0.0450	0.0470	0.0560	0.6950	0.2160	0.6190
	50	0.0480	0.0570	0.0500	0.7100	0.1750	0.6220
	70	0.0490	0.0530	0.0470	0.7510	0.4940	0.6540

relationship between public capital and private sector output, the following two-way fixed effects panel data model was considered:

$$\log(Y_{it}) = \mu_i + \xi_t + \beta_1 \log(Pc_{it}) + \beta_2 \log(Ps_{it}) + \beta_3 \log(L_{it}) \\ + \beta_4 Unemp_{it} + \varepsilon_{it}, \quad i = 1, \ldots, 48, \ t = 1, \ldots, 17, \tag{21}$$

where Y_{it} denotes the gross state product of state i in year t, Pc_{it} denotes public capital including highways and streets, water and sewer facilities and other public buildings, Ps_{it} denotes the stock of private capital, L_{it} is the labor input measured as employment in nonagricultural payrolls, and $Unemp_{it}$ is the state unemployment rate included to capture

Table 7. The test results of model (21) with the significance level $\alpha = 0.05$.

Z_{it}	L_1	L_2	L_3
Pc_{it}	Reject	Accept	Reject
Ps_{it}	Reject	Accept	Reject
L_{it}	Reject	Accept	Reject
$Unemp_{it}$	Accept	—	—
Pc_{it}, Ps_{it}	Reject	Accept	Reject
Pc_{it}, L_{it}	Reject	Accept	Reject
$Pc_{it}, Unemp_{it}$	Reject	Accept	Reject
Ps_{it}, L_{it}	Reject	Reject	Reject
$Ps_{it}, Unemp_{it}$	Accept	—	—
$L_{it}, Unemp_{it}$	Reject	Accept	Reject
Pc_{it}, Ps_{it}, L_{it}	Reject	Reject	Reject
$Pc_{it}, Ps_{it}, Unemp_{it}$	Reject	Accept	Reject
$Pc_{it}, L_{it}, Unemp_{it}$	Reject	Accept	Reject
$Ps_{it}, L_{it}, Unemp_{it}$	Reject	Reject	Reject
$Pc_{it}, Ps_{it}, L_{it}, Unemp_{it}$	Reject	Reject	Reject

business cycle effects. This panel data set consists of annual observations for the US 48 contiguous states over the period 1970–1986.

To the best of our knowledge, most previous works are conducted with homoskedastic errors. To investigate whether this assumption is justified, we applied the proposed test procedures to the data set. The p-values of test statistics L_1, L_2 and L_3 with $Z_{it} = X_{it} = (Pc_{it}, Ps_{it}, L_{it}, Unemp_{it})^\top$ are 0.0000, 0.0002 and 0.0000, respectively. The tests suggest that the error variances are heteroskedastic, and the heteroskedasticity is present at both within and between levels.

The test results with different Z_{it} are displayed in Table 7. The results show that, in almost all cases, the null of absence of heteroskedasticity in the data is rejected with the significance level $\alpha = 0.05$. Furthermore, in many cases, the heteroskedastic variation only comes from the different cross-sectional units. The above results should be interpreted with care because of the specific form of Z_{it}, but it seems reasonable to consider a more efficient inference for the public capital data which the heteroskedasticity is taken into account.

Example 5.2 (Cigar data): These panel data contain per capita cigarette consumptions of $N = 46$ US states from 1963 to 1992 ($T = 30$) as well as data about the income per capita and cigarette prices. We fit the data using the following model:

$$\begin{aligned}\log(Y_{it}) &= \mu_i + \xi_t + \beta_1 \log(X_{it,1}) + \beta_2 \log(X_{it,2}) + \beta_3 \log(X_{it,3}) \\ &\quad + \beta_4 \log(X_{it,4}) + \varepsilon_{it}, \quad i = 1, \ldots, 46,\ t = 1, \ldots, 30, \qquad (22)\end{aligned}$$

where Y_{it} represents the sales of cigarettes (packs of cigarettes per capita), $X_{it,1}$ is the average retail price of a pack of cigarettes measured in real terms, $X_{it,2}$ is the consumer price index, $X_{it,3}$ is the real per capita disposable income, $X_{it,4}$ denotes the minimum real price of cigarettes in any neighboring state, μ_i is the unobserved state-specific fixed effect, and ξ_t is the unobserved year-specific fixed effect.

A crucial assumption in the application of previous studies is constant variance of the observations. To investigate whether this assumption is justified, we apply the proposed test procedure to the Cigar data. That is, our main interest lies in testing $H_0 : \theta = 0$ (homoskedasticity) against a two-sided alternative (heteroskedasticity). Rejection of the

Table 8. The test results of model (22) with the significance level $\alpha = 0.05$.

Z_{it}	L_4	L_5	L_6
$X_{it,1}$	Accept	—	—
$X_{it,2}$	Accept	—	—
$X_{it,3}$	Accept	—	—
$X_{it,4}$	Accept	—	—
$X_{it,1}, X_{it,2}$	Reject	Reject	Reject
$X_{it,1}, X_{it,3}$	Accept	—	—
$X_{it,1}, X_{it,4}$	Reject	Reject	Reject
$X_{it,2}, X_{it,3}$	Reject	Accept	Reject
$X_{it,2}, X_{it,4}$	Reject	Reject	Accept
$X_{it,3}, X_{it,4}$	Accept	—	—
$X_{it,1}, X_{it,2}, X_{it,3}$	Reject	Reject	Reject
$X_{it,1}, X_{it,2}, X_{it,4}$	Reject	Reject	Reject
$X_{it,1}, X_{it,3}, X_{it,4}$	Reject	Reject	Reject
$X_{it,2}, X_{it,3}, X_{it,4}$	Reject	Reject	Reject
$X_{it,1}, X_{it,2}, X_{it,3}, X_{it,4}$	Reject	Reject	Reject

null hypothesis would suggest that the nonconstant response variance should be modeled as well. The p-values of test statistics L_4, L_5 and L_6 with $Z_{it} = X_{it} = (X_{it,1}, \ldots, X_{it,4})^\top$ are 0.0000, 0.0000 and 0.0000, respectively, which implies that the error variances are nonconstant. This conclusion is roughly consistent with the finding in Kouassi *et al.* [17], in which the authors investigated the data with a two-way random effects panel data model, and they assumed that the heteroskedasticity function is polynomial with $\bar{X}_{i\cdot}$. However, our tests suggest that the heteroskedastic variation is both within and between groups. Table 8 presents the test results with different forms of Z_{it}. Although the above results should be interpreted with care because of the specific form of Z_{it}, the results call for further investigation regarding the appropriateness of the specifications used in previous studies on the demand for cigarettes.

6. Concluding remarks

In this paper we construct tests for heteroskedasticity in two-way fixed effects panel data models under two important scenarios. For each case, three types of tests are derived in order to distinguish whether heteroskedasticity is present at the between level only, the within level only or both between and within levels. The proposed tests are distribution free and can be easily implemented after the estimation of the unknown regression coefficients and residuals. With some regularity conditions, we establish the asymptotic distributions of the proposed test statistics under both the null and alternative hypotheses. Both simulation studies and real data analyses are conducted to assess the finite sample performance of the proposed tests and to verify the established theoretical results.

There are several directions for future research. This paper focuses mostly on the additive fixed effects models, whereas the panel data models with interactive fixed effects can easily be accommodated in our framework since our tests can be implemented through simple auxiliary regressions. In addition, the method presented in this paper may also be applied to test heteroskedasticity for nonparametric and semiparametric panel data models with two-way fixed effects, such as varying coefficient panel data models.

Acknowledgments

The authors would like to thank the editor, associate editor and two referees for insightful comments that led to an improvement of an earlier manuscript.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

Sanying Feng's research was supported by the National Natural Science Foundation of China [grant number 11501522]. Gaorong Li's research was supported by the National Natural Science Foundation of China [grant numbers 11871001 and 11471029] and the Beijing Natural Science Foundation [grant number 1182003]. Tiejun Tong's research was supported by the National Natural Science Foundation of China [grant number 11671338] and the Health and Medical Research Fund [grant number 04150476]. Shuanghua Luo's research was supported by the National Natural Science Foundation of China [grant number 11601409].

References

[1] M. Arellano, *Panel Data Econometrics*, Oxford University Press, New York, 2003.
[2] O. Bada and D. Liebl, *phtt: Panel data analysis with heterogeneous time trends in R*, J. Stat. Softw 59 (2014), p. 6.
[3] J.S. Bai, *Panel data models with interactive fixed effects*, Econometrica 77 (2009), pp. 1229–1279.
[4] B.H. Baltagi, *Econometrics Analysis of Panel Data*, 5th ed., Wiley, New York, 2013.
[5] B.H. Baltagi, G. Bresson, and A. Pirotte, *Joint LM test for homoskedasticity in a one-way error component model*, J. Econometrics 134 (2006), pp. 401–417.
[6] B.H. Baltagi, Y.J. Chang, and Q. Li, *Monte Carlo results on several new and existing tests for the error component model*, J. Econometrics 54 (1992), pp. 95–120.
[7] B.H. Baltagi and D. Li, *Prediction in the panel data model with spatial correlation: The case of liquor*, Spat. Econ. Anal. 1 (2006), pp. 175–185.
[8] B.H. Baltagi, B.C. Jung, and S.H. Song, *Testing for heteroskedasticity and serial correlation in a random effects panel data model*, J. Econometrics 154 (2010), pp. 122–124.
[9] B.H. Baltagi, and N. Pinnoi, *Public capital stock and state productivity growth: Further evidence from an error components model*, Empir. Econ. 20 (1995), pp. 351–359.
[10] I. Fernandezval and M. Weidner, *Individual and time effects in nonlinear panel models with large N, T*, J. Econometrics 192 (2016), pp. 291–312.
[11] M. He and K.P. Lin, *Locally adjusted LM test for spatial dependence in fixed effects panel data models*, Econom. Lett. 121 (2013), pp. 59–63.
[12] A. Holly and L. Gardiol, *A score test for individual heteroscedasticity in a one-way error components model*. in *Panel Data Econometrics: Future Directions*, J. Krishnakumar and E. Ronchetti, eds., Elsevier, Amsterdam, 2000 pp. 199–211. Chapter 10.
[13] Y. Honda, *A standardized test for the error components model with the two-way layout*, Econom. Lett. 37 (1991), pp. 125–128.
[14] Y.M. Hong and C.W. Kao, *Wavelet-based testing for serial correlation of unknown form in panel models*, Econometrica 72 (2004), pp. 1519–1563.
[15] C. Hsiao, *Analysis of Panel Data*, 2nd ed., Cambridge University Press, Cambridge, 2003.
[16] T. Juhl and W. Sosa-Escudero, *Testing for heteroskedasticity in fixed effects models*, J. Econometrics 178 (2014), pp. 484–494.
[17] E. Kouassi, M. Mougoue, J. Sango, J.M. Brou, C.M. Amba, and A.A. Salisu, *Testing for heteroskedasticity and spatial correlation in a two way random effects model*, Comput. Stat. Data Anal. 70 (2014), pp. 153–171.

[18] E. Kouassi, J. Sango, J.M. Brou, and M. Mougoue, *A joint score test for heteroscedasticity in the two way error components model*, Comm. Statist. Theory Methods 43 (2014), pp. 2734–2751.
[19] L.F. Lee and J. Yu, *Estimation of spatial autoregressive panel data models with fixed effects*, J. Econometrics 154 (2010), pp. 165–185.
[20] L.F. Lee and J. Yu, *Estimation of fixed effects panel regression models with separable and nonseparable space-time filters*, J. Econometrics 184 (2015), pp. 174–192.
[21] D.G. Li, J.H. Qian, and L.J. Su, *Panel data models with interactive fixed effects and multiple structural breaks*, J. Amer. Statist. Assoc. 111 (2016), pp. 1804–1819.
[22] Q. Li and T. Stengos, *Adaptive estimation in the panel data error component model with heteroskedasticity of unknown form*, Internat. Econom. Rev. 35 (1994), pp. 981–1000.
[23] G. Montes-Rojas and W. Sosa-Escudero, *Robust tests for heteroskedasticity in the one-way error components model*, J. Econometrics 160 (2011), pp. 300–310.
[24] H.R. Moon and M. Weidner, *Linear regression for panel with unknown number of factors as interactive fixed effects*, Econometrica 83 (2015), pp. 1543–1579.
[25] A.H. Munnell, *Why has productivity growth declined? Productivity and public investment*, New Engl. Econom. Rev. 30 (1990), pp. 3–22.
[26] J.N.K. Rao and A.J. Scott, *The analysis of categorical data from complex sample surveys: Chi-squared tests for goodness of fit and independence in two-way tables*, J. Amer. Statist. Assoc. 76 (1981), pp. 221–230.
[27] L.J. Su and Q.H. Chen, *Testing homogeneity in panel data models with interactive fixed effects*, Economet. Theor. 29 (2013), pp. 1079–1135.
[28] L.J. Su, A. Ullah, and Y. Wang, *Nonparametric regression estimation with general parametric error covariance: A more efficient two-step estimator*, Empir. Econ. 45 (2013), pp. 1009–1024.
[29] J.H. Wu, *Robust random effects tests for two-way error component models with panel data*, Econom. Modelling 59 (2016), pp. 1–8.
[30] J.H. Wu and G.D. Li, *Moment-based tests for individual and time effects in panel data models*, J. Econometrics 178 (2014), pp. 569–581.
[31] J.H. Wu and J.C. Li, *Testing for individual and time effects in panel data models with interactive effects*, Econom. Lett. 125 (2014), pp. 306–310.
[32] J.H. Wu and L.X. Zhu, *Testing for serial correlation and random effects in a two-way error component regression model*, Econom. Modelling 28 (2011), pp. 2377–2386.
[33] Z. Yang, C. Li, and Y.K. Tse, *Functional form and spatial dependence in dynamic panels*, Econom. Lett. 91 (2006), pp. 138–145.

Appendix: Proofs of the theorems

To prove the theorems, the following lemmas are required.

Lemma A.1: *Suppose that regularity conditions* (C1)–(C5) *hold. As* $N \to \infty$, *then*

$$\hat{\beta} - \beta = O_p(N^{-1/2}).$$

The proof of Lemma A.1 can be found in Baltagi [4], and we hence omit the details.

Lemma A.2: *Suppose that the regularity conditions* (C1)–(C5) *hold. As* $N \to \infty$ *and* $T \to \infty$, *then*

$$\hat{\beta} - \beta = O_p((NT)^{-1/2}).$$

The proof of Lemma A.2 can be found in Hsiao [15], and we hence omit the details.

Proof of Theorem 2.1.: By (8), $\tilde{L}_1$ can be expressed as

$$\tilde{L}_1 = NTR_1^2 = NT\frac{\hat{\eta}^{2\top}MZ^*(Z^{*\top}MZ^*)^{-1}Z^{*\top}M\hat{\eta}^2}{\hat{\eta}^{2\top}M\hat{\eta}^2}.$$

By Lemma A.1 and the fact

$$\begin{aligned}\hat{\eta}_{it} &= (Y_{it} - \bar{Y}_{i.} - \bar{Y}_{.t} + \bar{Y}_{..}) - (X_{it} - \bar{X}_{i.} - \bar{X}_{.t} + \bar{X}_{..})^\top \hat{\beta} \\ &= (\varepsilon_{it} - \bar{\varepsilon}_{i.} - \bar{\varepsilon}_{.t} + \bar{\varepsilon}_{..}) + (X_{it} - \bar{X}_{i.} - \bar{X}_{.t} + \bar{X}_{..})^\top(\beta - \hat{\beta}),\end{aligned}$$

we have

$$\begin{aligned}\frac{1}{\sqrt{N}}\hat{\eta}^{2\top}MZ^* &= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{\eta}_{it}^2(Z_{it}^* - \bar{Z}_{..}^*) \\ &= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}(\varepsilon_{it} - \bar{\varepsilon}_{i.} - \bar{\varepsilon}_{.t} + \bar{\varepsilon}_{..})^2(Z_{it}^* - \bar{Z}_{..}^*) \\ &\quad + \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}(\beta - \hat{\beta})^\top(X_{it} - \bar{X}_{i.} - \bar{X}_{.t} + \bar{X}_{..}) \\ &\quad \times (X_{it} - \bar{X}_{i.} - \bar{X}_{.t} + \bar{X}_{..})^\top(\beta - \hat{\beta})(Z_{it}^* - \bar{Z}_{..}^*) \\ &\quad + \frac{2}{\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}(\varepsilon_{it} - \bar{\varepsilon}_{i.} - \bar{\varepsilon}_{.t} + \bar{\varepsilon}_{..})(X_{it} - \bar{X}_{i.} - \bar{X}_{.t} + \bar{X}_{..})^\top(\beta - \hat{\beta})(Z_{it}^* - \bar{Z}_{..}^*) \\ &= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}(\varepsilon_{it} - \bar{\varepsilon}_{i.} - \bar{\varepsilon}_{.t} + \bar{\varepsilon}_{..})^2(Z_{it}^* - \bar{Z}_{..}^*) + o_p(1),\end{aligned}$$

where $\bar{Z}_{..}^* = (1/NT)\sum_{i=1}^{N}\sum_{t=1}^{T}Z_{it}^*$. This yields that

$$\frac{1}{\sqrt{N}}\hat{\eta}^{2\top}MZ^* = \frac{1}{\sqrt{NT}}\eta^{2\top}MZ^* + o_p(1),$$

and

$$\begin{aligned}NTR_1^2 &= NT\frac{\eta^{2\top}MZ^*(Z^{*\top}MZ^*)^{-1}Z^{*\top}M\eta^2}{\eta^{2\top}M\eta^2} + o_p(1) \\ &= T\left(\frac{Z_0^{*\top}\eta_0^2}{\sqrt{N}}\right)^\top\left[\left(\frac{\eta_0^{2\top}\eta_0^2}{N}\right)\frac{Z_0^{*\top}Z_0^*}{N}\right]^{-1}\left(\frac{Z_0^{*\top}\eta_0^2}{\sqrt{N}}\right) + o_p(1),\end{aligned}$$

where $Z_0^* = MZ^*$, $\eta_0^2 = M\eta^2$.

Next, we prove the asymptotic normality of $(1/\sqrt{N})Z_0^{*\top}\eta_0^2$. Note that,

$$\begin{aligned}\frac{1}{\sqrt{N}}c^\top Z_0^{*\top}\eta_0^2 &= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}c^\top Z_{0i}^{*\top}\eta_{0i}^2 \\ &= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}c^\top Z_{0i}^{*\top}\left[(\eta_i^2 - d\sigma^2 1_T) + (d\sigma^2 1_T - \bar{\eta}_{..}^2 1_T)\right]\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{\sqrt{N}} \sum_{i=1}^{N} c^\top Z_{0i}^{*\top} (\eta_i^2 - d\sigma^2 1_T) \\
&\equiv \frac{1}{\sqrt{N}} \sum_{i=1}^{N} m_i
\end{aligned}
$$

for any k-dimensional nonzero vector c, where $d = 1 - (1/T) - (1/N) + (1/NT)$. By conditions (C2) and (C3), it is easy to show that $E(m_i) = 0$. Note that, for fixed T, simple calculation yields $\mathrm{Cov}(\eta_{it}^2, \eta_{is}^2 | X_i, Z_i) = O((1/T^2))$, $t \neq s$. Thus,

$$
\begin{aligned}
\mathrm{Var}(m_i) = E(m_i^2) &= E\left[c^\top Z_{0i}^{*\top} (\eta_i^2 - d\sigma^2 1_T)(\eta_i^2 - d\sigma^2 1_T)^\top Z_{0i}^* c\right] \\
&= c^\top E\left\{Z_{0i}^{*\top} E[(\eta_i^2 - d\sigma^2 1_T)(\eta_i^2 - d\sigma^2 1_T)^\top | Z_i] Z_{0i}^*\right\} c \\
&= c^\top \Sigma_{1i} c,
\end{aligned}
$$

where $\Sigma_{1i} = E(Z_{0i}^{*\top} \Omega_i Z_{0i}^*)$ and $\Omega_i = \mathrm{Var}(\eta_i^2 | X_i, Z_i)$.

By Lindeberg-Feller central limit theorem, we can obtain that

$$
\frac{1}{\sqrt{N}} c^\top Z_0^{*\top} \eta_0^2 = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} m_i \xrightarrow{\mathcal{L}} N(0, c^\top \Sigma_1 c).
$$

Therefore, by Cramer–Wold Theorem, we have

$$
\frac{1}{\sqrt{N}} Z_0^{*\top} \eta_0^2 = \frac{1}{\sqrt{N}} Z^{*\top} M \eta^2 \xrightarrow{\mathcal{L}} N(0, \Sigma_1). \tag{A1}
$$

Note that, under H_0, invoking the law of large numbers and conditions (C2)–(C4), we can derive that

$$
\begin{aligned}
\frac{Z_0^{*\top} Z_0^*}{N} &= \frac{1}{N} Z^{*\top} M Z^* = \frac{1}{N} \sum_{i=1}^{N} Z_{0i}^{*\top} Z_{0i}^* \xrightarrow{P} \Sigma_1^*, \\
\frac{\eta_0^{2\top} \eta_0^2}{N} &= \frac{1}{N} \eta^{2\top} M \eta^2 = \frac{1}{N} \sum_{i=1}^{N} \sum_{t=1}^{T} (\eta_{it}^2 - \bar{\eta}_{..}^2)^2 \xrightarrow{P} T\sigma_{\eta^2}^2.
\end{aligned}
$$

Then, we have

$$
\begin{aligned}
\tilde{L}_1 &= NT \frac{\eta^{2\top} M Z^* (Z^{*\top} M Z^*)^{-1} Z^{*\top} M \eta^2}{\eta^{2\top} M \eta^2} + o_p(1) \\
&= \left(\Sigma_1^{-1/2} \frac{1}{\sqrt{N}} Z^{*\top} M \eta^2\right)^\top \tilde{\Sigma} \left(\Sigma_1^{-1/2} \frac{1}{\sqrt{N}} Z^{*\top} M \eta^2\right) + o_p(1),
\end{aligned} \tag{A2}
$$

where $\tilde{\Sigma} = \Sigma_1^{1/2} \Sigma_1^{*-1} \Sigma_1^{1/2} / \sigma_{\eta^2}^2$.

Let $\Lambda = \mathrm{diag}(\phi_1, \ldots, \phi_k)$, where ϕ_j are the eigenvalues of $\Sigma_1^{*-1} \Sigma_1 / \sigma_{\eta^2}^2$. Because $\tilde{\Sigma}$ and $\Sigma_1^{*-1} \Sigma_1 / \sigma_{\eta^2}^2$ have the same eigenvalues, an orthonormal matrix Q exists such that $Q^\top \Lambda Q = \tilde{\Sigma}$. By (A2), we have

$$
\tilde{L}_1 = \left(Q\Sigma_1^{-1/2} \frac{1}{\sqrt{N}} Z^{*\top} M \eta^2\right)^\top \Lambda \left(Q\Sigma_1^{-1/2} \frac{1}{\sqrt{N}} Z^{*\top} M \eta^2\right) + o_p(1). \tag{A3}
$$

Invoking (A1) and (A3), we can prove that

$$
\tilde{L}_1 \xrightarrow{\mathcal{L}} \phi_1 \chi_{1,1}^2 + \cdots + \phi_k \chi_{1,k}^2,
$$

where $\chi_{1,j}^2$ are independent χ_1^2 variables. ■

Proof of Theorem 2.2.: For part (1), by

$$\hat{\rho} = \frac{\mathrm{tr}\{\hat{\Sigma}_1^{-1}\hat{B}\}}{\mathrm{tr}\{\hat{\Sigma}_1^{*-1}\hat{B}/\hat{\sigma}^2_{\eta^2}\}}$$

and the proof of Theorem 2.1, it is easy to check that

$$L_1 = \left(\frac{1}{\sqrt{N}}Z^{*\top}M\hat{\eta}^2\right)^\top \hat{\Sigma}_1^{-1}\left(\frac{1}{\sqrt{N}}Z^{*\top}M\hat{\eta}^2\right) + o_p(1) \xrightarrow{\mathcal{L}} \chi^2_k.$$

The proof of part (2) is similar to the arguments above, and we omit the details. We now provide the detailed calculation of the noncentral parameter. By Taylor's expansion of $g(Z_{it}^\top\theta)$, we get

$$g(Z_{it}^\top\theta) = g(0) + g'(0)Z_{it}^\top\theta + \tfrac{1}{2}g''(c^*)(Z_{it}^\top\theta)^2,$$

where $c^* \in [0, Z_{it}^\top\theta]$. Under the alternative hypothesis $H_1 : \theta = \delta/\sqrt{N}$, where δ is a k-dimensional finite vector. Noting that $\sum_{i=1}^N\sum_{t=1}^T Z_{0it}^* = 0$, we have

$$\begin{aligned}
E\left(\frac{1}{\sqrt{N}}Z^{*\top}M\eta^2\right) &= \frac{1}{\sqrt{N}}\sum_{i=1}^N\sum_{t=1}^T E[Z_{0it}^*(\varepsilon_{it} - \bar{\varepsilon}_{i.} - \bar{\varepsilon}_{.t} + \bar{\varepsilon}_{..})^2] \\
&= \frac{\sigma^2}{\sqrt{N}}\sum_{i=1}^N\sum_{t=1}^T E\left\{Z_{0it}^*\left[\left(1 - \frac{2}{T} - \frac{2}{N} + \frac{4}{NT}\right)g(Z_{it}^\top\theta) + \frac{1}{T^2}\left(1 - \frac{2}{N}\right)\sum_{s=1}^T g(Z_{is}^\top\theta)\right.\right. \\
&\quad \left.\left. + \frac{1}{N^2}\left(1 - \frac{2}{T}\right)\sum_{j=1}^N g(Z_{jt}^\top\theta) + \frac{1}{N^2T^2}\sum_{j=1}^N\sum_{s=1}^T g(Z_{js}^\top\theta)\right]\right\} \\
&= \sigma^2 g'(0)\frac{1}{N}\sum_{i=1}^N\sum_{t=1}^T E\left\{Z_{0it}^*\left[\left(1 - \frac{2}{T}\right)Z_{it}^\top\delta + \frac{1}{T^2}\sum_{s=1}^T Z_{is}^\top\delta\right]\right\} + o(1) \\
&= \sigma^2 g'(0)\frac{1}{N}\sum_{i=1}^N E(Z_{0i}^{*\top}Z_{0i}^*)\delta + o(1).
\end{aligned}$$

Thus, the limiting noncentral parameter becomes $\varsigma_1 = \sigma^4 g'(0)^2\delta^\top\Sigma_1^*\Sigma_1^{-1}\Sigma_1^*\delta$. ■

Proof of Theorem 2.3.: Similar to the proof of Theorem 2.1, we can get

$$\begin{aligned}
\frac{1}{\sqrt{N}}\hat{\eta}^{2\top}M_G Z^* &= \frac{1}{\sqrt{N}}\eta^{2\top}M_G Z^* + o_p(1), \\
\frac{1}{\sqrt{N}}c^\top Z_G^{*\top}\eta_G^2 &= \frac{1}{\sqrt{N}}\sum_{i=1}^N c^\top Z_{Gi}^{*\top}\left(\eta_i^2 - d\sigma^2 1_T\right) \equiv \frac{1}{\sqrt{N}}\sum_{i=1}^N m_{Gi}.
\end{aligned}$$

By the Central Limit Theorems, we can obtain that

$$\frac{1}{\sqrt{N}}c^\top Z_G^{*\top}\eta_G^2 = \frac{1}{\sqrt{N}}\sum_{i=1}^N m_{Gi} \xrightarrow{\mathcal{L}} N(0, \sigma^2_{\eta^2}c^\top\Sigma_2 c),$$

and further

$$\frac{1}{\sqrt{N}}Z_G^{*\top}\eta_G^2 \xrightarrow{\mathcal{L}} N(0, \sigma^2_{\eta^2}\Sigma_2).$$

In addition, under H_0 we have

$$\frac{1}{N}\eta^{2\top}M_G\eta^2 = \frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T}(\eta_{it}^2 - \bar{\eta}_{i.}^2)^2 \xrightarrow{P} T\sigma_{\eta^2}^2,$$

$$\frac{1}{N}Z^{*\top}M_G Z^* = \frac{1}{N}Z_G^{*\top}Z_G^* = \frac{1}{N}\sum_{i=1}^{N}Z_{Gi}^{*\top}Z_{Gi}^* \xrightarrow{P} \Sigma_2^*.$$

Thus, by (12) and

$$\hat{\varrho}_1 = \frac{\mathrm{tr}\{\hat{\Sigma}_2^{-1}\hat{B}_G\}}{\mathrm{tr}\{\hat{\Sigma}_2^{*-1}\hat{B}_G/\hat{\sigma}_{\eta^2}^2\}},$$

it is easy to show that

$$L_2 = NT\hat{\varrho}_1 R_2^2 = \left(\frac{1}{\sqrt{N}}Z^{*\top}M_G\hat{\eta}^2\right)^\top \hat{\Sigma}_2^{-1}\left(\frac{1}{\sqrt{N}}Z^{*\top}M_G\hat{\eta}^2\right) + o_p(1) \xrightarrow{\mathcal{L}} \chi_k^2.$$

Next, we drive the asymptotic distribution of L_2 under the alternative hypothesis $H_1 : \theta = \delta/\sqrt{N}$. Invoking $\sum_{t=1}^{T}(Z_{it} - \bar{Z}_{i.}) = 0$ and the Taylor's expansion of $g(\cdot)$, we have

$$\begin{aligned}
&E\left(\frac{1}{\sqrt{N}}Z^{*\top}M_G\eta^2\right)\\
&= \sigma^2 g'(0)\frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T}E\left\{Z_{Git}^*\left[\left(1-\frac{2}{T}\right)Z_{it}^\top\delta + \frac{1}{T^2}\sum_{s=1}^{T}Z_{is}^\top\delta\right]\right\} + o(1)\\
&= \sigma^2 g'(0)\frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T}E\left\{\left[\left(1-\frac{2}{T}\right)(Z_{it}-\bar{Z}_{i.})\right]\left[\left(1-\frac{2}{T}\right)Z_{it}^\top\delta + \frac{1}{T}\bar{Z}_{i.}^\top\delta\right]\right\} + o(1)\\
&= \sigma^2 g'(0)\left(1-\frac{2}{T}\right)^2\frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T}E(Z_{it}-\bar{Z}_{i.})(Z_{it}-\bar{Z}_{i.})^\top\delta + o(1)\\
&= \sigma^2 g'(0)\left(1-\frac{2}{T}\right)^2\frac{1}{N}\sum_{i=1}^{N}E(Z_i^\top A_G Z_i)\delta + o(1),
\end{aligned}$$

where $A_G = I_T - (1/T)1_T 1_T^\top$. Hence, the noncentral parameter becomes $\varsigma_2 = \sigma^4 g'(0)^2(1-(2/T))^4\delta^\top\Sigma_2^*\Sigma_2^{-1}\Sigma_2^*\delta$. ■

Proof of Theorem 2.4.: By the same arguments as in the proof of Theorems 2.1–2.3, the part (1) of Theorem 2.4 can be readily proved so that we omit the details.

To drive the asymptotic distribution of L_3 under the alternative hypothesis $H_1 : \theta = \delta/\sqrt{N}$. By invoking $\sum_{i=1}^{N}(Z_{it} - \bar{Z}_{.t}) = 0$ and Taylor's expansion of $g(\cdot)$, we have

$$\begin{aligned}
&E\left(\frac{1}{\sqrt{N}}\tilde{Z}^{*\top}M_H\tilde{\eta}^2\right)\\
&= \sigma^2 g'(0)\frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T}E\left\{Z_{Hit}^*\left[\left(1-\frac{2}{T}\right)Z_{it}^\top\delta + \frac{1}{T}\bar{Z}_{i.}^\top\delta\right]\right\} + o(1)\\
&= \sigma^2 g'(0)\frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T}E\left\{Z_{Hit}^*\left[\left(1-\frac{2}{T}\right)(Z_{it}-\bar{Z}_{i.})^\top\delta + \frac{1}{T}(\bar{Z}_{i.}-\bar{Z}_{..})^\top\delta\right]\right\} + o(1)
\end{aligned}$$

$$= \sigma^2 g'(0) \frac{1}{N} \sum_{i=1}^{N} \sum_{t=1}^{T} E(Z_{Hit}^{*} Z_{Hit}^{*\top}) \delta + o(1)$$

$$= \sigma^2 g'(0) \frac{1}{N} \sum_{i=1}^{N} E(Z_{Hi}^{*\top} Z_{Hi}^{*}) \delta + o(1).$$

Hence, the noncentral parameter becomes $\varsigma_3 = \sigma^4 g'(0)^2 \delta^\top \Sigma_3^* \Sigma_3^{-1} \Sigma_3^* \delta$. ■

Proof of Theorem 3.1.: We first prove part (1). By (15), L_4 can be expressed as

$$L_4 = NTR_4^2 = NT \frac{\hat{\eta}^{2\top} MZ(Z^\top MZ)^{-1} Z^\top M \hat{\eta}^2}{\hat{\eta}^{2\top} M \hat{\eta}^2}.$$

Note that, for $T \to \infty$, a simple calculation yields $\mathrm{Cov}(\eta_{it}^2, \eta_{is}^2 | X_i, Z_i, \mu_i) = O(T^{-2}) = o(1), \forall t \neq s$. Thus, under the null hypothesis H_0, similar to the proof of Theorem 2.1 we can obtain that

$$\frac{1}{\sqrt{NT}} Z_0^\top \eta_0^2 \xrightarrow{\mathcal{L}} N(0, \sigma_{\eta^2}^2 \Sigma_4).$$

$$\frac{Z_0^\top Z_0}{NT} = \frac{1}{NT} Z^\top MZ = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} Z_{0it} Z_{0it}^\top \xrightarrow{P} \Sigma_4,$$

$$\frac{\eta_0^{2\top} \eta_0^2}{NT} = \frac{1}{NT} \eta^{2\top} M \eta^2 = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} (\eta_{it}^2 - \bar{\eta}_{..}^2)^2 \xrightarrow{P} \sigma_{\eta^2}^2,$$

where $Z_0 = MZ$ and $\eta_0^2 = M\eta^2$. Then, invoking Lemma A.2, we have

$$L_4 = NT \frac{\eta^{2\top} MZ(Z^\top MZ)^{-1} Z^\top M \eta^2}{\eta^{2\top} M \eta^2} + o_p(1)$$

$$= \left(\frac{Z_0^\top \eta_0^2}{\sqrt{NT}} \right)^\top \left[\left(\frac{\eta_0^{2\top} \eta_0^2}{NT} \right) \frac{Z_0^\top Z_0}{NT} \right]^{-1} \left(\frac{Z_0^\top \eta_0^2}{\sqrt{NT}} \right) + o_p(1) \xrightarrow{\mathcal{L}} \chi_k^2.$$

On the other hand, under the alternative hypothesis $H_1 : \theta = \delta/\sqrt{NT}$, invoking $\sum_{i=1}^{N} \sum_{t=1}^{T} Z_{0it} = 0$ and the Taylor expansion of $g(\cdot)$, we have

$$E\left(\frac{1}{\sqrt{NT}} Z^\top M \eta^2 \right) = \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \sum_{t=1}^{T} E[Z_{0it}(\varepsilon_{it} - \bar{\varepsilon}_{i.} - \bar{\varepsilon}_{.t} + \bar{\varepsilon}_{..})^2]$$

$$= \frac{\sigma^2}{\sqrt{NT}} \sum_{i=1}^{N} \sum_{t=1}^{T} E \left\{ Z_{0it} \left[\left(1 - \frac{2}{T} - \frac{2}{N} + \frac{4}{NT} \right) g(Z_{it}^\top \theta) + \frac{1}{T^2} \left(1 - \frac{2}{N} \right) \sum_{s=1}^{T} g(Z_{is}^\top \theta) \right.\right.$$

$$\left.\left. + \frac{1}{N^2} \left(1 - \frac{2}{T} \right) \sum_{j=1}^{N} g(Z_{jt}^\top \theta) + \frac{1}{N^2 T^2} \sum_{j=1}^{N} \sum_{s=1}^{T} g(Z_{js}^\top \theta) \right] \right\}$$

$$= \sigma^2 g'(0) \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} E(Z_{0it} Z_{it}^\top) \delta + o(1)$$

$$= \sigma^2 g'(0) \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} E[Z_{0it}(Z_{it} - \bar{Z}_{..})^\top]\delta + o(1)$$

$$= \sigma^2 g'(0) \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} E(Z_{0it} Z_{0it}^\top)\delta + o(1).$$

Thus, the limiting noncentral parameter becomes

$$\varsigma_4 = \frac{\sigma^4 g'(0)^2 \delta^\top \Sigma_4 \delta}{\sigma_{\eta^2}^2}. \qquad \blacksquare$$

Proof of Theorem 3.2.: By (16), L_5 can be expressed as

$$L_5 = NTR_5^2 = NT \frac{\hat{\eta}^{2\top} M_G Z (Z^\top M_G Z)^{-1} Z^\top M_G \hat{\eta}^2}{\hat{\eta}^{2\top} M_G \hat{\eta}^2},$$

where $M_G = I_N \otimes (I_T - 1_T 1_T^\top / T)$. Similar to the proof of Theorems 2.1 and 2.3, we can get

$$\frac{1}{\sqrt{NT}} \hat{\eta}^{2\top} M_G Z = \frac{1}{\sqrt{NT}} \eta^{2\top} M_G Z + o_p(1),$$

$$\frac{1}{\sqrt{NT}} Z_G^\top \eta_G^2 \xrightarrow{\mathcal{L}} N(0, \sigma_{\eta^2}^2 \Sigma_5).$$

In addition, under H_0 we have

$$\frac{1}{NT} \eta^{2\top} M_G \eta^2 = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} (\eta_{it}^2 - \bar{\eta}_{i.}^2)^2 \xrightarrow{P} \sigma_{\eta^2}^2,$$

$$\frac{1}{NT} Z^\top M_G Z = \frac{1}{NT} Z_G^\top Z_G = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} Z_{Git} Z_{Git}^\top \xrightarrow{P} \Sigma_5.$$

From the above discussion, we can show that $L_5 = NTR_5^2 \xrightarrow{\mathcal{L}} \chi_k^2$.

On the other hand, under the alternative hypothesis $H_1 : \theta = \delta/\sqrt{NT}$, invoking $\sum_{t=1}^{T} Z_{Git} = 0$ and Taylor's expansion of $g(\cdot)$, we have

$$E\left(\frac{1}{\sqrt{NT}} Z^\top M_G \eta^2\right) = \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \sum_{t=1}^{T} E[Z_{Git}(\varepsilon_{it} - \bar{\varepsilon}_{i.} - \bar{\varepsilon}_{.t} + \bar{\varepsilon}_{..})^2]$$

$$= \sigma^2 g'(0) \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} E(Z_{Git} Z_{it}^\top)\delta + o(1)$$

$$= \sigma^2 g'(0) \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} E[Z_{Git}(Z_{it} - \bar{Z}_{i.})^\top]\delta + o(1)$$

$$= \sigma^2 g'(0) \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} E(Z_{Git} Z_{Git}^\top)\delta + o(1).$$

Hence, the limiting noncentral parameter becomes

$$\varsigma_5 = \frac{\sigma^4 g'(0)^2 \delta^\top \Sigma_5 \delta}{\sigma_{\eta^2}^2}. \qquad \blacksquare$$

Proof of Theorem 3.3.: By (17), L_6 can be expressed as

$$L_6 = NTR_6^2 = NT\frac{\tilde{\hat{\eta}}^{2\top} M_H \tilde{Z}(\tilde{Z}^\top M_H \tilde{Z})^{-1}\tilde{Z}^\top M_H \tilde{\hat{\eta}}^2}{\tilde{\hat{\eta}}^{2\top} M_H \tilde{\hat{\eta}}^2},$$

where $M_H = I_T \otimes (I_N - 1_N 1_N^\top/N)$, $\tilde{\hat{\eta}}^2 = (\tilde{\hat{\eta}}_1^{2\top}, \ldots, \tilde{\hat{\eta}}_T^{2\top})^\top$, $\tilde{\hat{\eta}}_t^2 = (\hat{\eta}_{1t}^2, \ldots, \hat{\eta}_{Nt}^2)^\top$, $\tilde{Z} = (\tilde{Z}_1, \ldots, \tilde{Z}_T)^\top$, and $\tilde{Z}_t = (Z_{1t}, \ldots, Z_{Nt})^\top$.

Similar to the proof of Theorem 2.1, we can get

$$\frac{1}{\sqrt{NT}}\tilde{\hat{\eta}}^{2\top} M_H \tilde{Z} = \frac{1}{\sqrt{NT}}\tilde{\eta}^{2\top} M_H \tilde{Z} + o_p(1) = \frac{1}{\sqrt{NT}}\sum_{i=1}^N\sum_{t=1}^T \eta_{it}^2 (Z_{it} - \bar{Z}_{.t}) + o_p(1),$$

$$\frac{1}{\sqrt{NT}}\tilde{Z}^\top M_H \tilde{\hat{\eta}}^2 \xrightarrow{\mathcal{L}} N(0, \sigma_{\eta^2}^2 \Sigma_6).$$

In addition, under H_0, by the law of large numbers and conditions (C2)–(C4), we have $\tilde{\hat{\eta}}^{2\top} M_H \tilde{\hat{\eta}}^2 \xrightarrow{P} \sigma_{\eta^2}^2$ and

$$\frac{1}{NT}\tilde{Z}^\top M_H \tilde{Z} = \frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T Z_{Hit} Z_{Hit}^\top \xrightarrow{P} \Sigma_6,$$

where $Z_{Hit} = Z_{it} - \bar{Z}_{.t}$. From the above discussion, we can get $L_6 = NTR_6^2 \xrightarrow{\mathcal{L}} \chi_k^2$.

Next, we drive the asymptotic distribution of L_6 under the alternative hypothesis $H_1 : \theta = \delta/\sqrt{NT}$. Invoking $\sum_{i=1}^N Z_{Hit} = 0$ and Taylor's expansion of $g(\cdot)$, we have

$$\begin{aligned}
E\left(\frac{1}{\sqrt{NT}}\tilde{Z}^\top M_H \tilde{\eta}^2\right) &= \frac{1}{\sqrt{NT}}\sum_{i=1}^N\sum_{t=1}^T E[Z_{Hit}(\varepsilon_{it} - \bar{\varepsilon}_{i.} - \bar{\varepsilon}_{.t} + \bar{\varepsilon}_{..})^2] \\
&= \sigma^2 g'(0)\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T E(Z_{Hit} Z_{it}^\top)\delta + o(1) \\
&= \sigma^2 g'(0)\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T E(Z_{Hit} Z_{Hit}^\top)\delta + o(1).
\end{aligned}$$

Hence, the limiting noncentral parameter becomes

$$\varsigma_6 = \frac{\sigma^4 g'(0)^2 \delta^\top \Sigma_6 \delta}{\sigma_{\eta^2}^2}. \qquad \blacksquare$$