Communications in Mathematics and Statistics
https://doi.org/10.1007/s40304-023-00371-5

New Approaches for Testing Slope Homogeneity in Large Panel Data Models

Guanghui Wang[1] · Long Feng[2] · Ping Zhao[2]

Received: 15 June 2021 / Revised: 28 July 2022 / Accepted: 22 July 2023

Abstract
Testing slope homogeneity is important in panel data modeling. Existing approaches typically take the summation over a sequence of test statistics that measure the heterogeneity of individual panels; they are referred to as Sum tests. We propose two procedures for slope homogeneity testing in large panel data models. One is called a Max test that takes the maximum over these individual test statistics. The other is referred to as a Combo test, which combines a certain Sum test (i.e., that of Pesaran and Yamagata in J Econom 142:50-93, 2008) and the proposed Max test together. We derive the limiting null distributions of the two test statistics, respectively, when both the number of individuals and temporal observations jointly diverge to infinity, and demonstrate that the Max test is asymptotically independent of the Sum test. Numerical results show that the proposed approaches perform satisfactorily.

Keywords Asymptotic independence · Large panels · Panel data models · Slope homogeneity

Mathematics Subject Classification 62G10

✉ Ping Zhao
zhaoping603@163.com

Long Feng
flnankai@gmail.com

1 KLATASDS-MOE, School of Statistics, and Academy of Statistics and Interdisciplinary Sciences, East China Normal University, Shanghai, People's Republic of China

2 School of Statistics and Data Science, KLMDASR, LEBPS, and LPMC, Nankai University, Tianjin, People's Republic of China

Published online: 25 January 2024

1 Introduction

In classical panel data analysis, it is often assumed that the slope coefficients of interest in panel data models are homogeneous across individual units. However, in practice, they can be individually specific. Ignoring this form of heterogeneity may result in biased estimation and inference. Thus, a formal test for slope homogeneity is necessary. When the number of individuals or panels, N, is fixed, and the number of temporal observations, T, diverges, a simple method is to use the standard F-test, which assumes exogenous regressors and homoskedastic errors. To eliminate the effect of heteroscedasticity, Swamay [20] proposed a dispersion test based on generalized least squares estimators under a random coefficient model. Another type of tests is based on Hausman's test [11], where the standard fixed effects estimator is compared to the mean group estimator; see, for example, Pesaran et al. [17] and Phillips and Sul [19]. However, these methods are not applicable in the case of panel data models that contain only strictly exogenous regressors and/or in the case of pure autoregressive models [18]. An early work of [22] proposed the seemingly unrelated regression equation (SURE) approach to incorporate cross-sectional dependence. The above approaches assume that $N < T$, and would lose their efficiency or even fail when N is comparable to, or even larger than, T, such as in many micro-econometric applications; the latter situation is referred to as large or high-dimensional panel data models.

In a high-dimensional setup, the dispersion test proposed by [17] allows $N > T$. Pesaran et al. [18] investigated the asymptotic distribution of the test statistic proposed by [20] in a large N, T scenario, and proposed a modified Swamy-type statistic, based on different estimators of regression error variances. Under the paradigm of fixed T but diverging N, Juhl and Lugovskyy [13] proposed a conditional Lagrange multiplier test based on the conditional Gaussian likelihood function, and [4] proposed some Lagrange multiplier tests, generalizing the test proposed by [5] against random individual effects to all regression coefficients.

Most approaches mentioned above are based on the summation of a sequence of test statistics for individual units, which are referred to as Sum tests. Sum tests turn to be efficient under dense alternatives, in the sense that the number of individual units with heterogeneous slope coefficients is large. However, for sparse alternatives when there are only a few heterogeneous individual units, Sum tests would be inefficient. In the latter situation, a maximum-based strategy can be more suitable, as widely discussed in the statistical literature, such as [6] and [21]. Motivated by this, we first propose a Max test based on the maximum of these individual test statistics. We establish its asymptotic distribution under the null hypothesis when $N, T \to \infty$, and show that the Max test outperforms a certain Sum test [18] in terms of power under sparse alternatives.

In practice, we seldomly know whether the alternatives are dense or sparse. Thus, it is kind of risky to simply apply a single Sum or Max test, if we have no priors on the sparsity level. This motivates us to develop an adaptive test to different levels of sparsity. We propose a Combo test, which combines the Sum and Max tests together, by taking the minimum p-value of these two separate tests. The asymptotic independence of Sum- and Max-type test statistics has been widely studied in the literature, such as [7, 12, 15], and [10], to name a few. Under some mild conditions, we show that the

Sum test statistic is asymptotically independent of the Max test statistic under the null hypothesis when $N, T \to \infty$. Consequently, the Combo test statistic is asymptotically distributed as the minimum of two independent standard uniform random variables under the null hypothesis. Theoretical results and simulation studies show that the Combo test performs very robust to either dense or sparse alternatives.

The rest of this paper is organized as follows. In Sect. 2, we give a brief literature review of testing procedures for slope homogeneity. We introduce Max and Combo tests, and establish their theoretical properties in Sect. 3. In Sect. 4, some numerical studies including real-data examples are conducted to evaluate the performance of the proposed methods. Some discussions are given in Sect. 5, and all technical details are deferred to Appendix.

2 The Model and Existing Approaches

We consider the following panel data model with fixed effects and potential heterogeneous slopes

$$y_{it} = \alpha_i + \boldsymbol{x}_{it}^\top \boldsymbol{\beta}_i + u_{it}, \ i = 1, \ldots, N, \ t = 1, \ldots, T, \tag{2.1}$$

where $\boldsymbol{x}_{it}$ is a p-dimensional vector of strictly exogenous regressors, α_i and $\boldsymbol{\beta}_i$ are the scalar intercept and p-dimensional slopes, respectively, and u_{it} are random errors with mean 0 and variance σ_i^2. Suppose that α_i are bounded on a compact set and $\boldsymbol{\beta}_i$ are bounded in the sense that $\|\boldsymbol{\beta}_i\| < K$ for some constant $K > 0$, where $\|\cdot\|$ is the Euclidean norm. Write in a compact form

$$\boldsymbol{Y}_i = \alpha_i \mathbf{1}_T + \mathbf{X}_i \boldsymbol{\beta}_i + \boldsymbol{u}_i,$$

where $\boldsymbol{Y}_i = (y_{i1}, \ldots, y_{iT})^\top$, $\mathbf{1}_T$ is a T-dimensional vector with all elements being 1, $\mathbf{X}_i = (\boldsymbol{x}_{i1}, \ldots, \boldsymbol{x}_{iT})^\top$, and $\boldsymbol{u}_i = (u_{i1}, \ldots, u_{iT})^\top$. Of interest is to test the null hypothesis

$$H_0 : \boldsymbol{\beta}_i = \boldsymbol{\beta} \ \text{ for all } i = 1, \ldots, N, \tag{2.2}$$

against the alternative hypothesis

$$H_1 : \text{there exist some } 1 \le i \ne j \le N \text{such that } \boldsymbol{\beta}_i \ne \boldsymbol{\beta}_j.$$

A well-known test is the standard F-test, which is valid for fixed N and diverging T and when the error variances are homoskedastic, i.e., $\sigma_i^2 = \sigma^2$. For $N > T$, [17] proposed a Hausman-type test [11], by comparing the standard fixed effects estimator with the mean group estimator, that is,

$$\hat{\boldsymbol{\beta}}_{\mathrm{FE}} = \left(\sum_{i=1}^N \mathbf{X}_i^\top \mathbf{M} \mathbf{X}_i \right)^{-1} \sum_{i=1}^N \mathbf{X}_i^\top \mathbf{M} \boldsymbol{Y}_i \text{ and } \hat{\boldsymbol{\beta}}_{\mathrm{MG}} = \frac{1}{N} \sum_{i=1}^N \hat{\boldsymbol{\beta}}_i,$$

respectively, where $\hat{\boldsymbol{\beta}}_i = \left(\mathbf{X}_i^\top \mathbf{M}\mathbf{X}_i\right)^{-1} \mathbf{X}_i^\top \mathbf{M}\boldsymbol{Y}_i$, $\mathbf{M} = \mathbf{I}_T - \mathbf{1}_T(\mathbf{1}_T^\top \mathbf{1}_T)^{-1}\mathbf{1}_T^\top$, and $\mathbf{I}_T$ is a $T \times T$ identity matrix. However, this test would lack power under a random coefficient model such that $E(\hat{\boldsymbol{\beta}}_{\text{FE}} - \hat{\boldsymbol{\beta}}_{\text{MG}}) = 0$. Phillips and Sul [19] proposed a different Hausman-type test based on

$$\left(\hat{\boldsymbol{\beta}} - \mathbf{1}_N \otimes \hat{\boldsymbol{\beta}}_{\text{FE}}\right)^\top \hat{\boldsymbol{\Sigma}}^{-1} \left(\hat{\boldsymbol{\beta}} - \mathbf{1}_N \otimes \hat{\boldsymbol{\beta}}_{\text{FE}}\right),$$

where $\hat{\boldsymbol{\beta}} = (\hat{\boldsymbol{\beta}}_1^\top, \ldots, \hat{\boldsymbol{\beta}}_N^\top)^\top$, $\hat{\boldsymbol{\Sigma}}$ is a consistent estimator of the variance matrix of $\hat{\boldsymbol{\beta}} - \mathbf{1}_N \otimes \hat{\boldsymbol{\beta}}_{\text{FE}}$ under H_0. This test is likely to be more powerful than that proposed by [17], but is still limited for fixed N. In the case of fixed N, Swamay [20] proposed a test based on

$$\hat{S} = \sum_{i=1}^{N} \left(\hat{\boldsymbol{\beta}}_i - \hat{\boldsymbol{\beta}}_{\text{WFE}}\right)^\top \frac{\mathbf{X}_i^\top \mathbf{M}\mathbf{X}_i}{\hat{\sigma}_i^2} \left(\hat{\boldsymbol{\beta}}_i - \hat{\boldsymbol{\beta}}_{\text{WFE}}\right),$$

where

$$\hat{\boldsymbol{\beta}}_{\text{WFE}} = \left(\sum_{i=1}^{N} \frac{\mathbf{X}_i^\top \mathbf{M}\mathbf{X}_i}{\hat{\sigma}_i^2}\right)^{-1} \sum_{i=1}^{N} \frac{\mathbf{X}_i^\top \mathbf{M}\boldsymbol{Y}_i}{\hat{\sigma}_i^2}$$

and $\hat{\sigma}_i^2 = (T - p - 1)^{-1}\left(\boldsymbol{Y}_i - \mathbf{X}_i\hat{\boldsymbol{\beta}}_i\right)^\top \mathbf{M}\left(\boldsymbol{Y}_i - \mathbf{X}_i\hat{\boldsymbol{\beta}}_i\right)$. Based on $\hat{S}$, Pesaran and Yamagata [18] showed that, as $N, T \to \infty$,

$$\hat{\Delta} = \sqrt{N(T+1)}\left(\frac{\hat{S}/N - p}{\sqrt{2p(T-p-1)}}\right)$$

converges to the standard normal distribution in distribution, if $N/T^2 \to 0$. Moreover, they proposed an adjusted test statistic, that is,

$$\tilde{\Delta}_{\text{adj}} = \sqrt{N(T+1)}\left(\frac{\tilde{S}/N - p}{\sqrt{2p(T-p-1)}}\right), \tag{2.3}$$

to weaken the dimension restriction, where

$$\tilde{S} = \sum_{i=1}^{N} \tilde{S}_i,\ \ \tilde{S}_i = \left(\hat{\boldsymbol{\beta}}_i - \tilde{\boldsymbol{\beta}}_{\text{WFE}}\right)^\top \frac{\mathbf{X}_i^\top \mathbf{M}\mathbf{X}_i}{\tilde{\sigma}_i^2} \left(\hat{\boldsymbol{\beta}}_i - \tilde{\boldsymbol{\beta}}_{\text{WFE}}\right), \tag{2.4}$$

$$\tilde{\boldsymbol{\beta}}_{\text{WFE}} = \left(\sum_{i=1}^{N} \frac{\mathbf{X}_i^\top \mathbf{M}\mathbf{X}_i}{\tilde{\sigma}_i^2}\right)^{-1} \sum_{i=1}^{N} \frac{\mathbf{X}_i^\top \mathbf{M}\boldsymbol{Y}_i}{\tilde{\sigma}_i^2},$$

and $\tilde{\sigma}_i^2 = (T-1)^{-1}\left(\boldsymbol{Y}_i - \mathbf{X}_i\hat{\boldsymbol{\beta}}_{\mathrm{FE}}\right)^\top \mathbf{M}\left(\boldsymbol{Y}_i - \mathbf{X}_i\hat{\boldsymbol{\beta}}_{\mathrm{FE}}\right)$. In other words, it modifies the $\hat{S}$ test by replacing the variance estimators $\hat{\sigma}_i^2$ by $\tilde{\sigma}_i^2$. The authors investigated the asymptotic normality under H_0 for non-normal errors, provided that $N/T^4 \to \infty$. Notice that for normal errors, both $\hat{\Delta}$ and $\tilde{\Delta}_{\mathrm{adj}}$ are valid without any restrictions on N and T.

Under the asymptotic regime of diverging N but fixed T, [13] proposed a conditional Lagrange multiplier test based on

$$T_{\mathrm{CLM}} = \sum_{i=1}^{N} S_i^\top \left(\sum_{i=1}^{N} S_i S_i^\top\right)^{-1} \sum_{i=1}^{N} S_i, \tag{2.5}$$

where $S_i = \hat{\boldsymbol{u}}_i^\top \mathbf{M}\mathbf{X}_i\mathbf{X}_i^\top\mathbf{M}\hat{\boldsymbol{u}}_i - \hat{\sigma}_i^2 \mathrm{tr}(\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i)$ and $\hat{\boldsymbol{u}}_i = \mathbf{M}(\boldsymbol{Y}_i - \mathbf{X}_i\hat{\boldsymbol{\beta}}_{\mathrm{FE}})$. By the fact that

$$\hat{S} = \sum_{i=1}^{N} \hat{\boldsymbol{u}}_i^\top \mathbf{M}\mathbf{X}_i(\sigma^2\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i)^{-1}\mathbf{X}_i^\top\mathbf{M}\hat{\boldsymbol{u}}_i + o_p(1),$$

the main difference between T_{CLM} and $\hat{\Delta}$ is that the statistics S_i neglect such terms $(\sigma_i^2\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i)^{-1}$ in $\hat{S}$. In fact, both can be regarded as testing the independence of $\boldsymbol{u}_i$ and $\mathbf{M}\mathbf{X}_i$ by the moment conditions $E(\boldsymbol{u}_i^\top\mathbf{M}\mathbf{X}_i\boldsymbol{W}_i\mathbf{X}_i^\top\mathbf{M}\boldsymbol{u}_i) = \sigma_i^2 E(\mathrm{tr}(\mathbf{M}\mathbf{X}_i\boldsymbol{W}_i\mathbf{X}_i^\top\mathbf{M}))$ with properly defined $\boldsymbol{W}_i$; that is, $\boldsymbol{W}_i = \mathbf{I}_p$ for T_{CLM} and $\boldsymbol{W}_i = (\sigma_i^2\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i)^{-1}$ for $\hat{\Delta}$. [4] proposed a Lagrange multiplier test under the heteroskedastic errors based on

$$T_{\mathrm{LM}} = \left(\sum_{i=1}^{N}\sum_{t=2}^{T}\tilde{u}_{it}\tilde{z}_{it}\right)^\top \left(\sum_{i=1}^{N}\sum_{t=2}^{T}\tilde{u}_{it}^2\tilde{z}_{it}\tilde{z}_{it}^\top\right)^{-1} \left(\sum_{i=1}^{N}\sum_{t=2}^{T}\tilde{u}_{it}\tilde{z}_{it}\right), \tag{2.6}$$

where $\tilde{u}_{it}$ is the t-th component of $\hat{\boldsymbol{u}}_i$ and $\tilde{z}_{it} = \boldsymbol{x}_{it}\sum_{s=1}^{t-1}\tilde{u}_{is}\boldsymbol{x}_{is}$. They showed that $T_{\mathrm{LM}} \to \chi_p^2$ in distribution, as $N \to \infty$ but keeping T fixed.

3 Our Tests

3.1 Methodology

A large value of

$$\tilde{S}_i = \left(\hat{\boldsymbol{\beta}}_i - \tilde{\boldsymbol{\beta}}_{\mathrm{WFE}}\right)^\top \left(\mathbf{X}_i^\top\mathbf{M}\mathbf{X}_i/\tilde{\sigma}_i{}^2\right)\left(\hat{\boldsymbol{\beta}}_i - \tilde{\boldsymbol{\beta}}_{\mathrm{WFE}}\right)$$

(cf. (2.4)) indicates a heterogeneous individual slope, for $i = 1, \ldots, N$. Most existing procedures for slope homogeneity testing are based on the summation of all $\tilde{S}_i$ or some variants. When there are a large proportion of individual units are heterogeneous with different slope coefficients (referred to as dense signals), Sum tests can accumulate all departure information together, thus making a powerful test against H_0. In contrast, when the number of heterogeneous individuals is very small (i.e., sparse signals), the summation statistic brings with redundant noises, which greatly decrease the testing power. Motivated by this, we propose a maximum-based statistic

$$T_{\text{Max}} = \max_{1 \leq i \leq N} \tilde{S}_i, \tag{3.1}$$

and we refer to the associated testing procedure as the Max test. It can be expected that the Max test would be more powerful against sparse alternatives.

Sometimes, we have some knowledge of the sparsity level, and we can choose between a Sum or Max test. However, if such priors are unavailable, a new method that is adaptive to the sparsity is demanded. We propose combining the Sum and Max tests in the following way

$$T_{\text{Combo}} = \min\{p_S, p_M\}, \tag{3.2}$$

where p_M and p_S are the p-values of the Max and Sum tests, respectively. To be specific, $p_M = 1 - F\{T_{\text{Max}} - 2\log(N) - (p-2)\log(\log(N)) + 2\log(\Gamma(p/2))\}$ and $p_S = 1 - \Phi\left\{\tilde{\Delta}_{\text{adj}}\right\}$, where $F(y) = e^{-e^{-y/2}}$ is the type-I extreme distribution function (i.e., the Gumbel distribution function), and $\Phi(y)$ denotes the standard normal distribution function. Here we use $\tilde{\Delta}_{\text{adj}}$ [18] as the Sum test statistic. We refer to this new test as the Combo test, which is expected to perform well, regardless of whether the alternatives are sparse or dense.

We summarize some theoretical properties of the Max and Combo tests here; more details are revealed in the following sections. For the Max test, we show that under some mild conditions,

$$\tilde{T}_{\text{Max}} \equiv T_{\text{Max}} - 2\log(N) - (p-2)\log(\log(N)) + 2\log(\Gamma(p/2))$$

converges to the type-I extreme distribution in distribution under H_0, as $N, T \to \infty$. Hence given a significance level $\alpha \in (0, 1)$, we can reject H_0 if $\tilde{T}_{\text{Max}}$ is larger than the $(1-\alpha)$-quantile of $F(y)$, say $q_\alpha \equiv -2\log(\log(1-\alpha)^{-1})$. We also derive the limiting null distribution of the Combo test by demonstrating that the Max statistic is asymptotically independent of the Sum statistic under H_0, as $N, T \to \infty$. Consequently, an asymptotic level-α test is to reject H_0 if $T_{\text{Combo}} < 1 - \sqrt{1-\alpha}$.

3.2 Max Test

To establish theoretical properties of the Max test, we need the following conditions

(C1) $u_{it} \sim N(0, \sigma_i^2)$ and $\sigma_{max}^2 = \max_{1 \leq i \leq N} \sigma_i^2$ is bounded.
(C2) u_{it} and u_{js} are independently, for $i \neq j$ and/or $t \neq s$.
(C3) For $i = 1, \ldots, N$, $\boldsymbol{\Sigma}_{iT} \equiv T^{-1}\mathbf{X}_i^\top \mathbf{M}\mathbf{X}_i$ is positive definite and bounded, and converges to a non-stochastic positive definite and bounded matrix $\boldsymbol{\Sigma}_i$, as $T \to$

∞. $\boldsymbol{\Sigma}_A \equiv (NT)^{-1}\left(\sum_{i=1}^N \mathbf{X}_i^\top \mathbf{M}\mathbf{X}_i\right)$ is positive definite and converges to a non-stochastic positive definite matrix $\boldsymbol{\Sigma}$, as $N, T \to \infty$.

(C4) u_{it} is independent of x_{js}, for all i, j, t, s.

Condition (C1) is crucial to obtain the asymptotic distribution of the test statistic T_{Max} and the asymptotic independence between T_{Max} and $\tilde{\Delta}_{\text{adj}}$. An extension to non-normal errors deserves further studies, see discussions in Section 5. Condition (C2) assumes the cross-sectional independence, Condition (C3) is used for the consistency of the least square estimators of $\boldsymbol{\beta}_i$, and Condition (C4) means that $\boldsymbol{x}_{it}$ are strictly exogenous; these conditions are standard in the literature, see, for example, [18].

Theorem 3.1 *Suppose conditions* (*C*1)–(*C*4) *hold. Under* H_0, *if* $\log(N) = o(T^{1/3})$, *then*

$$P\left\{T_{\text{Max}} - 2\log(N) - (p-2)\log(\log(N)) + 2\log(\Gamma(p/2)) \le x\right\} \\ \to \exp(-\exp(-x/2)),$$

as $N, T \to \infty$.

According to the limiting null distribution, we can reject H_0 if

$$\tilde{T}_{\text{Max}} \equiv T_{\text{Max}} - 2\log(N) - (p-2)\log(\log(N)) + 2\log(\Gamma(p/2)) \ge q_\alpha,$$

where q_α is the $(1-\alpha)$-quantile of the type-I extreme value distribution with the cumulative distribution function $\exp\{-\exp(-x/2)\}$, namely, $q_\alpha = -2\log(\log(1-\alpha)^{-1})$.

Now, we turn to the power analysis of the Max test. Define

$$\mathcal{A}(c) = \left\{\boldsymbol{\delta} : \max_{1\le i\le N} T\sigma_i^{-2}\boldsymbol{\omega}_i^\top \boldsymbol{\Sigma}_{iT}\boldsymbol{\omega}_i \ge c\log(N)\right\},$$

where

$$\boldsymbol{\omega}_i = \boldsymbol{\delta}_i - \left(N^{-1}\sum_{i=1}^N \sigma_i^{-2}\boldsymbol{\Sigma}_{iT}\right)^{-1}\left(N^{-1}\sum_{i=1}^N \sigma_i^{-2}\boldsymbol{\Sigma}_{iT}\boldsymbol{\delta}_i\right) \text{ and } \boldsymbol{\delta}_i = \boldsymbol{\beta}_i - \boldsymbol{\beta}.$$

Notice that [18] considered the following local alternative hypotheses for the Sum test, that is, $\sum_{i=1}^N \sigma_i^{-2}\boldsymbol{\omega}_i^\top \boldsymbol{\Sigma}_{iT}\boldsymbol{\omega}_i = O(T^{-1}N^{1/2})$.

Theorem 3.2 *Suppose conditions* (*C*1)–(*C*4) *hold. If* $\log(N) = o(T^{1/3})$, *then for any* $\epsilon > 0$,

$$\inf_{\boldsymbol{\delta}\in\mathcal{A}(16+\epsilon)} P(\Psi_\alpha = 1) \to 1,$$

as $N, T \to \infty$, *where* $\Psi_\alpha = I(\tilde{T}_{\text{Max}} \ge q_\alpha)$ *is the power function.*

Theorem 3.2 shows that the proposed Max test is consistent if some $\sigma_i^{-2}\boldsymbol{\omega}_i^\top \boldsymbol{\Sigma}_{iT}\boldsymbol{\omega}_i$ is larger than the order $\log(N)/T$.

To make a comparison between the Max and Sum tests, define a class of sparse alternatives

$$\mathcal{S}(s_N, c_{T,N}) = \left\{ \boldsymbol{\delta} : \sum_{i=1}^{N} I(\boldsymbol{\delta}_i \neq 0) \leq s_N,\ (16+\epsilon)\log(N) \leq \max_{1\leq i\leq N} T\sigma_i^{-2}\boldsymbol{\omega}_i^\top \boldsymbol{\Sigma}_{iT}\boldsymbol{\omega}_i \leq c_{T,N} \right\},$$

with $s_N = o(\sqrt{N}/c_{T,N})$. By observing that $\mathcal{S}(s_N, c_{T,N}) \subset \mathcal{A}(16+\epsilon)$, the Max test is consistent over $\mathcal{S}(s_N, c_{T,N})$, according to Theorem 3.2. In contrast, in Section 3.2 of [18], the authors showed that, under $\mathcal{S}(s_N, c_{T,N})$, $\tilde{\Delta}_{\text{adj}}$ would suffer from trivial power. Hence the Max test is more efficient than the Sum test in such situations.

3.3 Combo test

To investigate the limiting null distribution and power property of the proposed Combo test, we first demonstrate the asymptotic independence between $\tilde{\Delta}_{\text{adj}}$ and T_{Max} under the null hypothesis.

Theorem 3.3 *Suppose conditions (C1)–(C4) hold. Under H_0, if $\log(N) = o(T^{1/3})$, then $\tilde{\Delta}_{\text{adj}}$ and T_{Max} are asymptotically independent in the sense that*

$$P\left(\tilde{\Delta}_{\text{adj}} \leq x, T_{\text{Max}} - 2\log(N) - (p-2)\log(\log(N)) + 2\log(\Gamma(p/2)) \leq y\right) \to \Phi(x)F(y),$$

as $N, T \to \infty$.

As a corollary, we derive the limiting null distribution of the Combo test.

Corollary 3.4 *Assume the conditions in Theorem 3.3 hold. Then T_{Combo} converges in distribution to $W = \min\{U, V\}$, as $N, T \to \infty$, where U, V are independent and identically distributed (iid) as a standard uniform random variable, and thus W has the density $G(w) = 2(1-w)I(0 \leq w \leq 1)$.*

By Corollary 3.4, given a significance level α, we can reject H_0 if $T_{\text{Combo}} < 1 - \sqrt{1-\alpha} \approx \alpha/2$ for a relatively small α.

The power function of the Combo test is $\beta_C(\boldsymbol{\delta}, \alpha) = P(T_{\text{Combo}} < 1 - \sqrt{1-\alpha})$. It can be verified that

$$\begin{aligned}\beta_C(\boldsymbol{\delta}, \alpha) =& P\left(p_M < 1 - \sqrt{1-\alpha}\right) + P\left(p_S < 1 - \sqrt{1-\alpha}\right) \\ &- P\left(p_M < 1 - \sqrt{1-\alpha},\ p_S < 1 - \sqrt{1-\alpha}\right) \\ \geq& \max\left\{P\left(p_S < 1 - \sqrt{1-\alpha}\right), P\left(p_M < 1 - \sqrt{1-\alpha}\right)\right\} \\ \approx& \max\left\{\beta_S(\boldsymbol{\delta}, \alpha/2), \beta_M(\boldsymbol{\delta}, \alpha/2)\right\},\end{aligned} \tag{3.3}$$

where $\beta_M(\boldsymbol{\delta}, \alpha)$ and $\beta_S(\boldsymbol{\delta}, \alpha)$ are the power functions of T_{Max} and $\tilde{\Delta}_{\text{adj}}$, respectively, at the significance level α. As shown in [18], $\beta_S(\boldsymbol{\delta}, \alpha) = \Phi(-z_\alpha + \psi)$, where $\psi = \lim_{N,T\to\infty} \frac{1}{\sqrt{2pN}} \sum_{i=1}^N T\sigma_i^{-2} \boldsymbol{\omega}_i^\top \boldsymbol{\Sigma}_{iT} \boldsymbol{\omega}_i$ and z_α is the upper $(1-\alpha)$-quantile of the standard normal distribution. According to (3.3), we have $\beta_C(\boldsymbol{\delta}, \alpha) \geq \Phi\left(-z_{\alpha/2} + \psi\right)$.

To compare the power of tests based on $\tilde{\Delta}_{\text{adj}}$, T_{Max} and T_{Combo}, consider a simplified scenario where $\boldsymbol{\Sigma}_{iT} = \mathbf{I}_p$ and $\sigma_i^2 = 1$. Moreover, m elements of $\boldsymbol{\delta}_i = (\delta_{i1}, \ldots, \delta_{ip})^\top$ are randomly sampled from $U(-\gamma, \gamma)$ for some $\gamma > 0$, and the rest are set to be 0, where $U(a, b)$ is the uniform distribution with the support $[a, b]$.

1. Assume that $m \to \infty$. By noticing that $m^{-1} \sum_{i=1}^N \sigma_i^{-2} \boldsymbol{\Sigma}_{iT} \boldsymbol{\delta}_i \xrightarrow{p} \mathbf{0}$ and $m^{-1} \sum_{i=1}^N \sigma_i^{-2} \boldsymbol{\omega}_i^\top \boldsymbol{\Sigma}_{iT} \boldsymbol{\omega}_i \xrightarrow{p} \frac{1}{3} p\gamma^2$, we have

$$\beta_S(\boldsymbol{\delta}, \alpha) = \Phi\left(-z_\alpha + \frac{Tmp^{1/2}\gamma^2}{3\sqrt{2N}}\right).$$

In addition, we have $\epsilon\gamma^2 < \max_{1\leq i\leq N} \sigma_i^{-2} \boldsymbol{\omega}_i^\top \boldsymbol{\Sigma}_{iT} \boldsymbol{\omega}_i \leq p\gamma^2$, for any positive constant $\epsilon < p$, with probability approaching one. We consider two special cases:

(1) *Dense case* $\gamma = O(T^{-\xi})$ and $m = O(N^{1/2}T^{2\xi-1})$ with some $\xi > 1/2$. In this case, $T\gamma^2 = o(1)$, and it can be verified that $\beta_M(\boldsymbol{\delta}, \alpha) \approx \alpha$. Thus, the Max test lacks power. Then, $\beta_C(\boldsymbol{\delta}, \alpha) \approx \beta_S(\boldsymbol{\delta}, \alpha/2) \approx \beta_S(\boldsymbol{\delta}, \alpha)$, if α is small. Hence, the Combo test performs similarly to $\tilde{\Delta}_{\text{adj}}$.

(2) *Sparse case* $\gamma = c\sqrt{\log N/T}$ for a sufficient large constant c and $m = o((\log N)^{-1}N^{1/2})$. In this case, $\frac{Tmp^{1/2}\gamma^2}{3\sqrt{2N}} \to 0$ and $\beta_S(\boldsymbol{\delta}, \alpha) \approx \alpha$; in other words, the Sum test based on $\tilde{\Delta}_{\text{adj}}$ lacks power. According to Theorem 3.2, $\beta_M(\boldsymbol{\delta}, \alpha) \to 1$. Consequently, the Combo test has the power $\beta_C(\boldsymbol{\delta}, \alpha) \to 1$.

2. Assume that m is fixed. By noticing that $\sum_{i=1}^N \sigma_i^{-2} \boldsymbol{\omega}_i^\top \boldsymbol{\Sigma}_{iT} \boldsymbol{\omega}_i = O_p(\gamma^2)$ and $\max_{1\leq i\leq N} \sigma_i^{-2} \boldsymbol{\omega}_i^\top \boldsymbol{\Sigma}_{iT} \boldsymbol{\omega}_i = O_p(\gamma^2)$, we can similarly show that, if $\gamma = c\sqrt{\log N/T}$ for a sufficient large constant c, then $\beta_S(\boldsymbol{\delta}, \alpha) \approx \alpha$, $\beta_M(\boldsymbol{\delta}, \alpha) \to 1$, and $\beta_C(\boldsymbol{\delta}, \alpha) \to 1$.

4 Numerical studies

4.1 Simulation

In this section, we investigate the finite-sample performance of the proposed Max and Combo tests based on T_{Max} and T_{Combo}, respectively. We choose some benchmark approaches, i.e., the tests based on $\tilde{\Delta}_{\text{adj}}$ [18], T_{CLM} [13], and T_{LM} [4]; see (2.3), (2.5), and (2.6), respectively. We consider the following three examples with independent, correlated, and structured noises, respectively. All simulation results are based on 1000 replications. We set the nominal significance level as $\alpha = 5\%$ in all examples.

Example 4.1 We revisit the model in [18]

$$
\begin{aligned}
y_{it} &= \alpha_i + \sum_{l=1}^{p} x_{ilt}\beta_{il} + u_{it},\ i = 1, \ldots, N,\ t = 1, \ldots, T, \\
x_{ilt} &= \alpha_i(1-\rho_{il}) + \rho_{il}x_{1i,t-1} + (1-\rho_{il}^2)^{1/2}v_{ilt},\ t = -48, \ldots, 0, \ldots, T,
\end{aligned}
$$

where $v_{ilt} \overset{iid}{\sim} N(0, \sigma_{ilx}^2)$. We fix some $\rho_{il} \overset{iid}{\sim} U(0.05, 0.95)$ and $\sigma_{ilx}^2 \overset{iid}{\sim} \chi^2(1)$ throughout the simulation study. We generate $\alpha_i \overset{iid}{\sim} N(1, 1)$, and discard the first 49 observations to reduce the effect of initial values. Three scenarios to generate $u_{it} = \sigma_i z_{it}$, with $\sigma_i^2 \overset{iid}{\sim} \frac{p}{2}\chi^2(2)$, are considered as follows:

(I) Normal distribution, $z_{it} \overset{iid}{\sim} N(0, 1)$;
(II) t-distribution, $z_{it} \overset{iid}{\sim} t(3)/\sqrt{3}$;
(III) Mixture of normals, $z_{it} \overset{iid}{\sim} \{0.9N(0, 1) + 0.1N(0, 100)\}/\sqrt{10.9}$.

Under H_0, $\beta_{il} \equiv 1$, for all i and l. While under H_1, we set $\beta_{il} = \beta_{i1}$, for $l \neq 1$, and for $\{\beta_{11}, \ldots, \beta_{N1}\}$, we first randomly choose $l_1 < \cdots < l_m$ from $\{1, \ldots, N\}$, and then generate $\beta_{l_i 1} \sim U(1 - 1.1m^{-0.65}, 1 + 1.1m^{-0.65})$, for $i = 1, \ldots, m$, and set the rest $\beta_{i1},\ i \notin \{l_1, \ldots, l_m\}$, to be 1.

Example 4.2 To study the impact of correlated errors on the testing procedures, we generated $u_{it} = \sigma_i z_{it}$ with $z_t = (z_{1t}, \ldots, z_{Nt})^\top \sim N(\mathbf{0}, \mathbf{\Sigma}_z)$, where $\mathbf{\Sigma}_z = (0.5^{|i-j|})_{1\le i,j\le N}$. The other settings are the same as in Example 4.1.

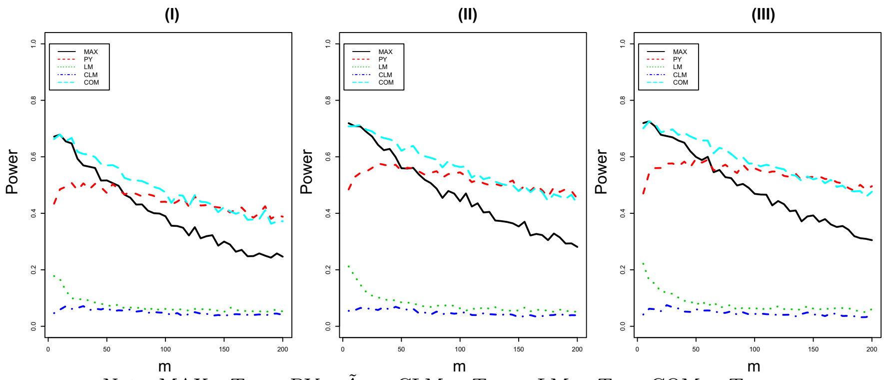
Example 4.3 We consider a high-dimensional panel data model with interactive fixed effects [2]. We generated $u_{it} = \boldsymbol{f}_t^\top \boldsymbol{\lambda}_i + \sigma_i z_{it}$ [1], where $\boldsymbol{f}_t$ are 2-dimensional factor vectors with iid $N(0, 1)$ entries, $\boldsymbol{\lambda}_i \overset{iid}{\sim} N(\mathbf{0}, 0.25\mathbf{I}_2)$ are factor-loading vectors, and $z_t = (z_{1t}, \ldots, z_{Nt})^\top \sim N(\mathbf{0}, \mathbf{\Sigma}_z)$, with $\mathbf{\Sigma}_z = (0.5^{|i-j|})_{1\le i,j\le N}$, are noises. The other settings are the same as in Example 4.1, except that, under H_1, $\beta_{l_i 1} \overset{iid}{\sim} U(1 - 2m^{-0.6}, 1 + 2m^{-0.6})$.

Table 1 presents empirical sizes of various slope homogeneity tests under Example 4.1, with the configuration that $p \in \{2, 3, 4\}$, $T \in \{50, 100\}$ and $N \in \{50, 100, 200\}$. We can see that the Max test is a little conservative when the sample size is small. This is not strange, because the convergence rate of the extreme value distribution is rather slow [16]. In most cases, the $\tilde{\Delta}_{\text{adj}}$, T_{LM} and T_{Combo} tests can maintain the sizes at the nominal level, while the T_{CLM} test tends to be fairly conservative.

Figures 1-3 report the power of various tests with $p = 2$, 3 or 4 under Example 4.1, respectively, when $T = 100$ and $N = 200$. We observe that the T_{LM} and T_{CLM} tests perform not very well, in terms of low power. As expected, the Max test outperforms the $\tilde{\Delta}_{\text{adj}}$-based Sum test when m is relatively small, and as m becomes large, the Sum test performs better than the Max test. This is consistent with our theoretical result, that is, the Sum test is favorable for detecting dense signals, while the Max test is preferred for sparse scenarios. The proposed Combo test performs similarly to the Max test

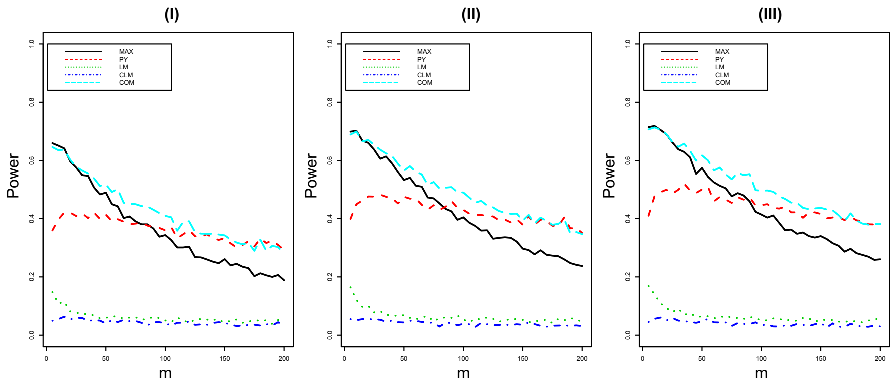
Table 1 Empirical Sizes of various slope homogeneity tests under Example 4.1

Scenarios		$p=2$			$p=3$			$p=4$		
		N			N			N		
T	Method	50	100	200	50	100	200	50	100	200
		(I)								
50	T_{Max}	1.9	1.7	0.8	3.2	1.5	1.6	3.8	3.0	2.3
	$\tilde{\Delta}_{\text{adj}}$	4.1	3.7	4.7	3.7	4.4	4.3	3.4	3.6	3.9
	T_{LM}	5.5	6.3	5.8	6.6	4.4	4.7	5.3	4.3	5.9
	T_{CLM}	1.4	1.0	1.3	1.3	0.7	1.2	1.1	1.0	1.5
	T_{Combo}	4.0	3.5	4.2	3.6	4.2	4.1	3.7	3.5	3.8
100	T_{Max}	3.1	3.7	2.6	3.8	4.2	3.3	5.3	6.8	4.5
	$\tilde{\Delta}_{\text{adj}}$	3.4	4.6	4.9	2.8	4.2	5.2	3.0	3.5	3.7
	T_{LM}	5.1	5.8	5.7	5.3	5.8	6.5	5.5	5.9	5.2
	T_{CLM}	0.6	0.8	1.3	0.3	1.8	1.9	1.2	1.4	1.9
	T_{Combo}	3.5	4.5	4.4	3.3	4.2	4.6	4.7	4.5	3.9
		(II)								
50	T_{Max}	2.3	1.1	1.3	1.3	1.8	1.2	3.9	3.0	1.7
	$\tilde{\Delta}_{\text{adj}}$	4.8	4.5	4.5	4.0	3.0	3.7	3.0	4.0	2.6
	T_{LM}	5.7	4.3	5.0	4.4	4.4	3.9	5.6	4.7	5
	T_{CLM}	0.6	1.6	1.6	1.0	0.7	1.0	1.1	1.1	1.9
	T_{Combo}	4.3	4.1	4.0	3.8	3.2	3.7	3.9	3.8	2.8
100	T_{Max}	2.9	3	2.5	4.5	3.2	3.1	7.3	5.5	4.3
	$\tilde{\Delta}_{\text{adj}}$	4.1	5.0	6.0	4.9	4.3	4.6	2.3	4.0	5.2
	T_{LM}	5.8	4.3	5.6	6.2	4.9	4.3	6.7	5.7	5.0
	T_{CLM}	1.1	1.2	1.3	1.1	1.0	1.6	0.9	1.1	2.0
	T_{Combo}	4.0	4.7	4.8	4.6	4.1	4.1	5.7	5.1	4.8
		(III)								
50	T_{Max}	1.3	1.8	1.0	2.2	1.4	1.1	4.1	1.8	1.2
	$\tilde{\Delta}_{\text{adj}}$	4.0	4.0	3.9	3.4	3.5	3.9	2.7	3.1	4.0
	T_{LM}	4.5	3.2	4.4	3.7	4.0	3.9	4.5	4.6	3.2
	T_{CLM}	1.1	1.2	1.7	1.0	1.6	1.4	0.8	1.3	1.5
	T_{Combo}	3.7	3.5	3.8	3.3	3.2	3.4	3.7	3.2	3.8
100	T_{Max}	3.1	1.7	2.6	3.9	4.2	4.8	6.9	6.4	5.3
	$\tilde{\Delta}_{\text{adj}}$	4.9	3.5	4.0	3.8	4.3	4.5	3.8	3.4	5.2
	T_{LM}	5.1	3.9	4.4	3.4	4.6	5.3	4.5	4.5	3.5
	T_{CLM}	1.1	1.2	1.9	0.6	1.1	1.4	1.3	1.2	1.6
	T_{Combo}	4.3	3.7	4.0	3.6	4.0	4.4	4.7	4.5	5.2



Note: MAX—T_{Max}; PY—$\tilde{\Delta}_{\text{adj}}$; CLM— T_{CLM}; LM— T_{LM}; COM— T_{Combo}.

Fig. 1 Power of various slope homogeneity tests with $p = 2$ under Example 4.1



Note: MAX—T_{Max}; PY—$\tilde{\Delta}_{\text{adj}}$; CLM— T_{CLM}; LM— T_{LM}; COM— T_{Combo}.

Fig. 2 Power of various slope homogeneity tests with $p = 3$ under Example 4.1

for small m, and has a similar performance to the Sum test for large m. Moreover, it outperforms both the Max and Sum tests for moderate m. Our simulation results reveal that the Combo test is very efficient in most cases, and it adapts to different levels of the sparsity. In addition, both the proposed Max and Combo tests, together with $\tilde{\Delta}_{\text{adj}}$, are robust to non-normal noises.

Tables 2 and 3 present empirical sizes of various tests under Examples 4.2 and 4.3, respectively, with a wide range of (p, N, T) configurations. Figures 4, 5 depict the power of each test against the sparsity level m. Similar conclusions can be made as under Example 4.1. In particular, the Combo test adapts to the sparsity and has very good power. We also conduct a simulation study regarding some larger dimensions $N = 400$ and $T = 200$ under Example 4.2; see Fig. 6. It can be seen from Fig. 6 that the proposed tests perform satisfactorily.

Having observed that the Max test can be sometimes conservative (see, for example, Tables 1-3), we provide a bootstrap calibration procedure to accommodate such issues. Based on the residuals $\hat{\boldsymbol{u}}_i = \mathbf{M}(\boldsymbol{Y}_i - \mathbf{X}_i\hat{\boldsymbol{\beta}}_i) = (\hat{u}_{i1}, \ldots, \hat{u}_{iT})^\top$, for $i = 1, \ldots, N$, we

Table 2 Empirical Sizes of various slope homogeneity tests under Example 4.2

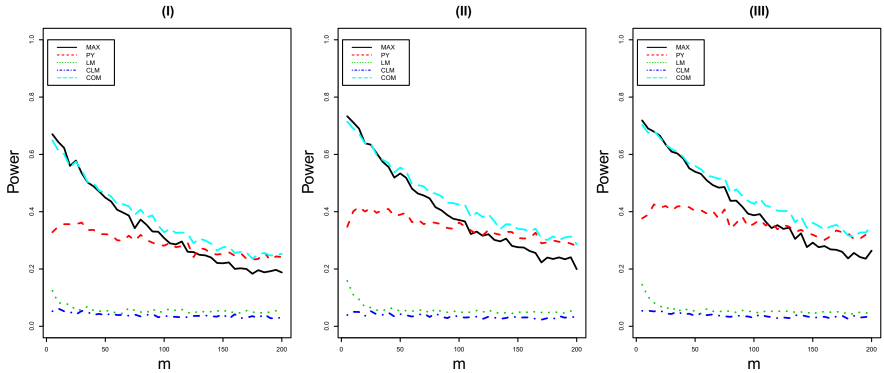
Scenarios		$p=2$			$p=3$			$p=4$		
		N			N			N		
T	Method	50	100	200	50	100	200	50	100	200
50	T_{Max}	1.4	1.1	1.3	3.8	1.7	2.0	3.7	2.8	1.8
	$\tilde{\Delta}_{\text{adj}}$	5.1	5.1	5.0	4.4	3.4	3.9	3.2	3.1	3.9
	T_{LM}	6.2	6.2	5.6	5.6	4.0	4.1	5.2	4.5	5.5
	T_{CLM}	0.7	1.5	1.2	1.2	1.6	1.2	0.7	1.3	1.5
	T_{Combo}	3.0	3.9	3.1	3.4	2.6	1.4	3.3	2.7	2.4
100	T_{Max}	2.5	2.8	1.9	4.7	4.5	2.4	6.6	6.2	5.2
	$\tilde{\Delta}_{\text{adj}}$	3.4	5.4	3.4	2.8	5.7	4.7	3.7	3.2	4.4
	T_{LM}	6.7	5.1	5.3	6.4	5.3	4.0	6.3	5.8	5.4
	T_{CLM}	0.7	1.0	1.2	0.7	1.4	1.6	1.4	1.2	1.1
	T_{Combo}	2.7	4.3	2.8	3.4	4	3.8	5.2	4.9	4.4

Table 3 Empirical Sizes of various slope homogeneity tests under Example 4.3

Scenarios		$p=2$			$p=3$			$p=4$		
		N			N			N		
T	Method	50	100	200	50	100	200	50	100	200
50	T_{Max}	1.9	1.7	1.5	2.2	2.6	1.1	2.7	1.7	1.8
	$\tilde{\Delta}_{\text{adj}}$	5.9	5.3	6.4	4.0	6.0	6.3	4.4	4.7	6.1
	T_{LM}	5.3	5.1	5.6	6.6	4.0	5.1	6.4	4.7	4.7
	T_{CLM}	0.7	1.2	2.0	0.9	1.5	1.6	1.6	0.8	1.9
	T_{Combo}	3	2.9	4.1	2.8	4.5	4.4	2.6	3.3	4.0
100	T_{Max}	3.3	2.9	3.4	5.3	3.3	3.1	7.2	7.9	4.7
	$\tilde{\Delta}_{\text{adj}}$	5.3	5.2	6.5	4.7	3.8	6.2	3.5	5.2	4.4
	T_{LM}	6.1	5.0	5.5	5.7	5.3	4.6	7.4	5.4	5.9
	T_{CLM}	0.9	0.9	1.7	1.1	1.0	1.2	0.9	1.3	1.7
	T_{Combo}	3.8	4.1	5.4	5	3.8	4.7	5	6.1	5.1

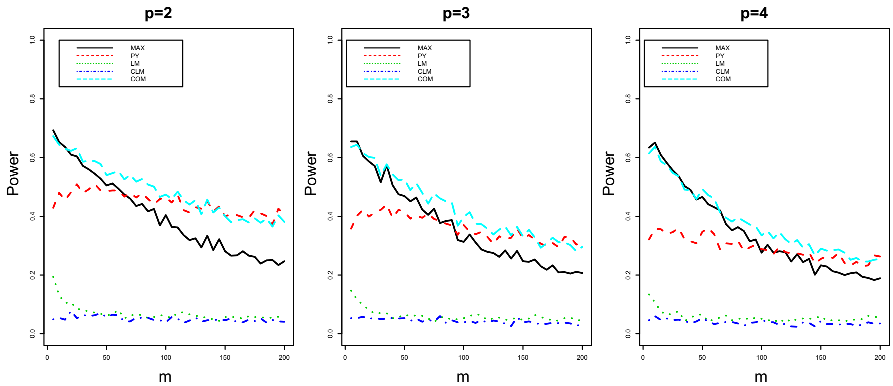
Table 4 Empirical sizes of the proposed tests with their bootstrap calibrations under Example 4.2

Scenarios		$p=2$			$p=3$			$p=4$		
		N			N			N		
T	Method	50	100	200	50	100	200	50	100	200
50	Max	1.4	1.1	1.3	3.8	1.7	2.0	3.7	2.8	1.8
	Sum	5.1	5.1	5.0	4.4	3.4	3.9	3.2	3.1	3.9
	Combo	3.0	3.9	3.1	3.4	2.6	1.4	3.3	2.7	2.4
	Max*	5.7	4.7	4.6	3.8	5.0	4.6	3.2	3.3	3.3
	Combo*	6.1	5.2	5.3	3.9	4.3	4.7	4.9	3.8	4.2



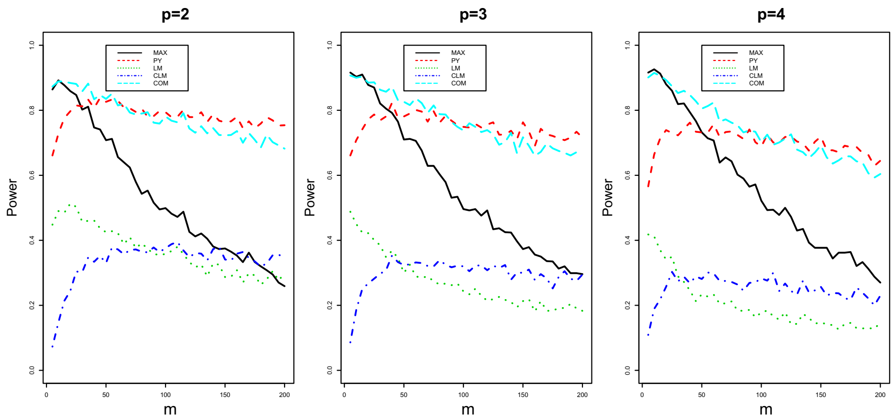
Note: MAX—T_{Max}; PY—$\tilde{\Delta}_{\text{adj}}$; CLM— T_{CLM}; LM— T_{LM}; COM— T_{Combo}.

Fig. 3 Power of various slope homogeneity tests with $p = 4$ under Example 4.1



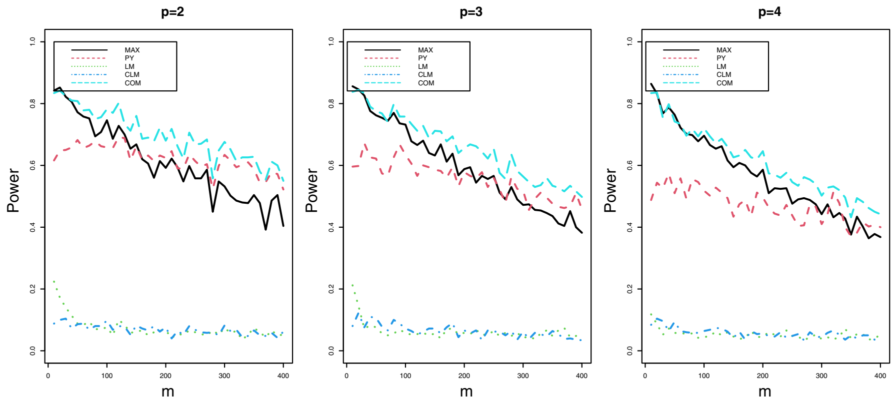
Note: MAX—T_{Max}; PY—$\tilde{\Delta}_{\text{adj}}$; CLM— T_{CLM}; LM— T_{LM}; COM— T_{Combo}.

Fig. 4 Power of various slope homogeneity tests under Example 4.2



Note: MAX—T_{Max}; PY—$\tilde{\Delta}_{\text{adj}}$; CLM— T_{CLM}; LM— T_{LM}; COM— T_{Combo}.

Fig. 5 Power of various slope homogeneity tests under Example 4.3



Note: MAX—T_{Max}; PY—$\tilde{\Delta}_{\text{adj}}$; CLM— T_{CLM}; LM— T_{LM}; COM— T_{Combo}.

Fig. 6 Power of various slope homogeneity tests under Example 4.2 with $N = 400$ and $T = 200$

generated N bootstrap samples

$$Y_i^* = \mathbf{X}_i \boldsymbol{\gamma}_0 + \boldsymbol{\eta}_i, i = 1, \ldots, N,$$

where $\boldsymbol{\gamma}_0 = (1, \ldots, 1)^\top \in \mathbb{R}^p$ and $\boldsymbol{\eta}_i = (\eta_{i1}, \cdots, \eta_{iT})^\top$ such that $\boldsymbol{\eta}_{\cdot t}$ are bootstrap samples from $\{\hat{\boldsymbol{u}}_{\cdot t}\}_{t=1}^T$ with $\hat{\boldsymbol{u}}_{\cdot t} = (\hat{u}_{1t}, \cdots, \hat{u}_{Nt})^\top$, and $\boldsymbol{\eta}_{\cdot t}$ is the tth column of $(\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_N)^\top$. Hence, a bootstrap calibrated Max statistic can be computed based on the bootstrap sample $(Y_i^*, \mathbf{X}_i)$, $i = 1, \cdots, N$. By repeating the sampling procedure $B = 500$ times, an empirical p-value can be obtained, say p_M^*. If $p_M^* < \alpha$, then we can reject the null hypothesis. We refer to this bootstrap calibrated procedure as Max*. In a similar way, we can define a bootstrap calibrated Combo test, by rejecting H_0 if $p_M^* < 1 - \sqrt{1-\alpha}$ or $p_S < 1 - \sqrt{1-\alpha}$, which is referred to as the Combo* test. Table 4 reports the empirical sizes of the proposed testing procedures, together with their bootstrap calibrations, under Example 4.2. We observe that both calibrated procedures perform very well. It is interesting to investigate the asymptotic validity of these tests for future researches.

4.2 Real data analysis

We study a real-data example of securities in stock markets to assess the performance of the proposed tests. To model the data, we use the Fama–French three-factor model [9], which adds size risk and value risk factors to the market risk factor in the capital asset pricing model. To be specific, assume

$$Y_{it} = r_{it} - r_{\text{f}t} = \alpha_i + \beta_{i1}(r_{\text{m}t} - r_{\text{f}t}) + \beta_{i2}\text{SMB}_t + \beta_{i3}\text{HML}_t + u_{it},$$

for $i \in \{1, \ldots, N\}$ and $t \in \{\tau, \ldots, \tau + T - 1\}$, where r_{it} is the return of portfolio i at time t, $r_{\text{f}t}$ is the risk-free rate at time t, $r_{\text{m}t}$ is the market portfolio return at time t, SMB_t is the size premium (small minus big), and HML_t is the value premium (high minus low). We are interested in testing $H_0 : \boldsymbol{\beta}_i = \boldsymbol{\beta}$ for all $i = 1, \ldots, N$ versus

Table 5 Rejection rates of different tests for beta testing, regarding the China and US stock markets

	$T=25$			$T=30$		
	N			N		
Method	30	50	80	30	50	800
China Stock Market						
T_{Max}	49.3	53.9	54.9	65.3	73.4	78.9
$\tilde{\Delta}_{\text{adj}}$	94.3	98.3	99.8	96.1	99.1	100
T_{LM}	48.5	69.3	83.9	53.5	72	87.8
T_{CLM}	63	78.5	87.1	67.4	82.7	92.4
T_{Combo}	92.6	97.5	99.5	95	99.1	100
US Stock Market						
T_{Max}	49.9	43.1	40.7	66.6	69.6	68.4
$\tilde{\Delta}_{\text{adj}}$	95.1	98.6	99.5	98.5	99.5	100
T_{LM}	66	83	93.4	74.5	89.9	95.1
T_{CLM}	86.6	94.9	97.2	92.4	97.1	98.9
T_{Combo}	93.3	98	99.3	98.3	99.2	99.9

$H_1 : \boldsymbol{\beta}_i \neq \boldsymbol{\beta}_j$ for some $1 \leq i \neq j \leq N$, where $\boldsymbol{\beta}_i = (\beta_{i1}, \beta_{i2}, \beta_{i3})^\top$, for all $i = 1, \ldots, N$.

Two data sets are investigated. One is the data set of securities in China's stock markets. We consider $N = 1,340$ securities during the period from June 2005 to May 2019, measured in percentages per month. Hence, we have a total amount of $T = 144$ temporal observations. The rate of China's 10-year government bond is chosen as the risk-free rate $r_{\text{f}t}$, the value-weighted return on the stocks of Shanghai Stock Exchange and Shenzhen Stock Exchange is used as a proxy for the market return $r_{\text{m}t}$, the average return on three small portfolios minus that on three big portfolios is calculated as SMB_t, and the average return on two value portfolios minus that on two growth portfolios is used as HML_t. The other data set is from the S&P 500 index. We specify the time range from January 2005 to November 2018 with $T = 165$, and collect $N = 374$ securities during this period.

We first apply five tests based on T_{Max}, $\tilde{\Delta}_{\text{adj}}$, T_{CLM}, T_{LM} and T_{Combo} to each data set with full samples. All tests reject the null hypothesis significantly, which shows that different stocks have different beta values. Next, we consider a restricted data size by randomly sampling $T \in \{25, 30\}$ and $N \in \{30, 50, 80\}$ observations from each data set, and then repeating the process 1,000 times for each (T, N) combination. Table 5 reports the rejection rates of different tests for each data set. We observe that the Max test performs not very well, which indicates the signal may be dense. This is further verified by the fact that the $\tilde{\Delta}_{\text{adj}}$-based Sum test performs the best among all (T, N) combinations. The proposed Combo test (i.e., T_{Combo}) performs very similarly to the Sum test, consistent with our theoretical and simulation findings.

5 Discussion

In this paper, we propose two approaches for slope homogeneity testing in high-dimensional panel data models, that is, the Max and Combo tests. The Max is more efficient compared to traditional Sum tests under sparse alternatives, while the Combo is robust to different levels of sparsity. We established the limiting null distributions of both test statistics. Two limitations of the present work are: (1) the errors are assumed to be normal; and (2) the cross-sectional units are assumed to be independent. Our simulation studies show that the proposed tests may perform satisfactorily under non-normal and/or correlated errors, but the theoretical properties deserve further studies. Recent developments of slope homogeneity testing with cross-sectional dependence and/or serially correlated errors, such as [1, 3] and [4], could be extended for our methods. We leave it as future researches.

6 Appendix

6.1 Some useful lemmas

Lemma 6.1 restates a result in Table 3.4.4 in [8].

Lemma 6.1 *Suppose* $z_i \overset{iid}{\sim} \Gamma(k, \theta)$, *for* $i = 1, \ldots, n$. *Then* $a_n(\max_{1\le i\le n} z_i - b_n) \xrightarrow{d} \Lambda$, *as* $n \to \infty$, *where* Λ *is the Gumbel distribution with* $P(\Lambda < x) = e^{-e^{-x}}$, $a_n = 1/\theta$, *and* $b_n = \theta(\log(n) + (k-1)\log(\log(n)) - \log(\Gamma(k)))$.

Lemma 6.2 *Suppose* $\boldsymbol{\varepsilon}_i \overset{iid}{\sim} N(0, \mathbf{I}_p)$, *for* $i = 1, \ldots, N$. *Let* $\mathbf{A}_i \in \mathbb{R}^{p\times p}$ *be positive definite matrices, for all* $i = 1, \ldots, N$, *and* $\max_{1\le i\le N} \lambda_{\max}(\mathbf{A}_i) \le C$, *for some positive constant* C. *Then*, $\max_{1\le i\le N} \boldsymbol{\varepsilon}_i^\top \mathbf{A}_i \boldsymbol{\varepsilon}_i = O_p(\log(N))$, *as* $N \to \infty$.

Proof Consider the eigenvalue decomposition $\mathbf{A}_i = \boldsymbol{\Omega}_i^\top \mathbf{D}_i \boldsymbol{\Omega}_i$, where $\mathbf{D}_i = \mathrm{diag}(\lambda_{i1}, \ldots, \lambda_{ip})$, λ_{ik} are the eigenvalues of $\mathbf{A}_i$, and $\boldsymbol{\Omega}_i$ is an orthogonal matrix. Then, $\boldsymbol{\varepsilon}_i^\top \mathbf{A}_i \boldsymbol{\varepsilon}_i = \boldsymbol{\varepsilon}_i^\top \boldsymbol{\Omega}_i^\top \mathbf{D}_i \boldsymbol{\Omega}_i \boldsymbol{\varepsilon}_i$. Since $\boldsymbol{\varepsilon}_i \sim N(0, \mathbf{I}_p)$, $\boldsymbol{\Omega}_i \boldsymbol{\varepsilon}_i \sim N(0, \mathbf{I}_p)$. Thus, $\boldsymbol{\varepsilon}_i^\top \mathbf{A}_i \boldsymbol{\varepsilon}_i$ equals $\boldsymbol{\varepsilon}_i^\top \mathbf{D}_i \boldsymbol{\varepsilon}_i = \sum_{k=1}^p \lambda_{ik} \boldsymbol{\varepsilon}_{ik}^2$ in distribution. Then,

$$\max_{1\le i\le N} \boldsymbol{\varepsilon}_i^\top \mathbf{A}_i \boldsymbol{\varepsilon}_i \overset{d}{=} \max_{1\le i\le N} \sum_{k=1}^p \lambda_{ik} \boldsymbol{\varepsilon}_{ik}^2 \le \max_{1\le i\le N} \lambda_{\max}(A_i) \sum_{k=1}^p \boldsymbol{\varepsilon}_{ik}^2 \le C \max_{1\le i\le N} \sum_{k=1}^p \boldsymbol{\varepsilon}_{ik}^2.$$

Obviously, $\sum_{k=1}^p \boldsymbol{\varepsilon}_{ik}^2 \sim \chi_p^2 = \Gamma(\frac{p}{2}, 2)$. Thus, by Lemma 6.1, we have

$$P\left(\max_{1\le i\le N} \sum_{k=1}^p \boldsymbol{\varepsilon}_{ik}^2 \le 3\log(N)\right) \sim \exp(-\exp(-\log(N)/2)) \to 1,$$

as $N \to \infty$. Then, $\max_{1\le i\le N} \sum_{k=1}^p \boldsymbol{\varepsilon}_{ik}^2 = O_p(\log(N))$, and thus $\max_{1\le i\le N} \boldsymbol{\varepsilon}_i^\top \mathbf{A}_i \boldsymbol{\varepsilon}_i = O_p(\log(N))$. □

Lemma 6.3 restates Lemma 6.1 in [14].

Lemma 6.3 *Suppose* $X \sim \chi_k^2$, *we have* $P(X \ge k + \sqrt{2kx} + 2x) \le \exp(-x)$ *and* $P(k - X \ge \sqrt{2kx}) \le \exp(-x)$.

Lemma 6.4 *Suppose conditions* $(C1)$–$(C4)$ *holds. Under* H_0, $\max_{1\le i\le N}|\tilde{\sigma}_i^2-\sigma_i^2| = O_p(\sqrt{\log(N)/T})$.

□

Proof According to (A.15) in [18], we have

$$\tilde{\sigma}_i^2 = \frac{\boldsymbol{\varepsilon}_i^\top \mathbf{M}\boldsymbol{\varepsilon}_i}{T-1} + \frac{1}{N(T-1)}\boldsymbol{\xi}_A^\top \boldsymbol{\Sigma}_A^{-1}\boldsymbol{\Sigma}_{iT}\boldsymbol{\Sigma}_A^{-1}\boldsymbol{\xi}_A + \frac{2}{\sqrt{N}(T-1)}\boldsymbol{\xi}_A^\top \boldsymbol{\Sigma}_A \boldsymbol{\xi}_i,$$

where $\boldsymbol{\xi}_i = T^{-1/2}\mathbf{X}_i^\top \mathbf{M}\boldsymbol{\varepsilon}_i$ and $\boldsymbol{\xi}_A = N^{-1/2}\sum_{i=1}^N \boldsymbol{\xi}_i$. Thus,

$$\max_{1\le i\le N}|\tilde{\sigma}_i^2-\sigma_i^2| \le \max_{1\le i\le N}\left|\frac{\boldsymbol{\varepsilon}_i^\top \mathbf{M}\boldsymbol{\varepsilon}_i}{T-1}-\sigma_i^2\right| + \max_{1\le i\le N}\frac{1}{N(T-1)}\boldsymbol{\xi}_A^\top \boldsymbol{\Sigma}_A^{-1}\boldsymbol{\Sigma}_{iT}\boldsymbol{\Sigma}_A^{-1}\boldsymbol{\xi}_A$$
$$+ \max_{1\le i\le N}\frac{2}{\sqrt{N}(T-1)}|\boldsymbol{\xi}_A^\top \boldsymbol{\Sigma}_A \boldsymbol{\xi}_i|.$$

By Condition (C1), we have $\sigma_i^{-2}\boldsymbol{\varepsilon}_i^\top \mathbf{M}\boldsymbol{\varepsilon}_i \sim \chi_{T-1}^2$. Let $\sigma_{max}^2 = \max_{1\le i\le N}\sigma_i^2$. By Lemma 6.3, we have

$$\begin{aligned}
&P\left(\max_{1\le i\le N}\frac{\boldsymbol{\varepsilon}_i^\top \mathbf{M}\boldsymbol{\varepsilon}_i}{T-1}-\sigma_i^2 > 3\sigma_{\max}^2\sqrt{\log(N)/T}\right)\\
&\le NP\left(\frac{\boldsymbol{\varepsilon}_i^\top \mathbf{M}\boldsymbol{\varepsilon}_i}{T-1}-\sigma_i^2 > 3\sigma_{\max}^2\sqrt{\log(N)/T}\right)\\
&\le NP\left(\frac{\sigma_i^{-2}\boldsymbol{\varepsilon}_i^\top \mathbf{M}\boldsymbol{\varepsilon}_i}{T-1}-1 > 3\sigma_i^{-2}\sigma_{\max}^2\sqrt{\log(N)/T}\right)\\
&\le NP\left(\chi_{T-1}^2-(T-1) \ge \sqrt{2.5(T-1)\log(N)}+2.5\sqrt{\log(N)}\right)\\
&\le N\exp(-1.25\log(N)) = N^{-1/4}\to 0.
\end{aligned}$$

Similarly, we have

$$P\left(\max_{1\le i\le N}\sigma_i^2-\frac{\boldsymbol{\varepsilon}_i^\top \mathbf{M}\boldsymbol{\varepsilon}_i}{T-1} > 3\sigma_{\max}^2\sqrt{\log(N)/T}\right)\to 0.$$

Thus,

$$P\left(\max_{1\le i\le N}\left|\sigma_i^2-\frac{\boldsymbol{\varepsilon}_i^\top \mathbf{M}\boldsymbol{\varepsilon}_i}{T-1}\right| > 3\sigma_{\max}^2\sqrt{\log(N)/T}\right)\to 0.$$

By Condition (C1), we have

$$\boldsymbol{\xi}_A^\top \boldsymbol{\Sigma}_A^{-1}\boldsymbol{\Sigma}_{iT}\boldsymbol{\Sigma}_A^{-1}\boldsymbol{\xi}_A \overset{d}{=} \boldsymbol{z}^\top \boldsymbol{\Sigma}_\sigma^{1/2}\boldsymbol{\Sigma}_A^{-1}\boldsymbol{\Sigma}_{iT}\boldsymbol{\Sigma}_A^{-1}\boldsymbol{\Sigma}_\sigma^{1/2}\boldsymbol{z},$$

where $z \sim N(0, \mathbf{I}_p)$ and $\mathbf{\Sigma}_\sigma = N^{-1}\sum_{i=1}^N \sigma_i^2 \mathbf{\Sigma}_{iT}$. By Condition (C3), the eigenvalues of $\mathbf{\Sigma}_\sigma^{1/2}\mathbf{\Sigma}_A^{-1}\mathbf{\Sigma}_{iT}\mathbf{\Sigma}_A^{-1}\mathbf{\Sigma}_\sigma^{1/2}$ are bounded. Thus,

$$\max_{1\le i\le N} \frac{1}{N(T-1)}\boldsymbol{\xi}_A^\top \mathbf{\Sigma}_A^{-1}\mathbf{\Sigma}_{iT}\mathbf{\Sigma}_A^{-1}\boldsymbol{\xi}_A \le \frac{C}{N(T-1)}\max_{1\le i\le N} z^\top z = \frac{C}{N(T-1)} z^\top z.$$

Because $z^\top z \sim \chi_p^2$, $\max_{1\le i\le N} \frac{1}{N(T-1)}\boldsymbol{\xi}_A^\top \mathbf{\Sigma}_A^{-1}\mathbf{\Sigma}_{iT}\mathbf{\Sigma}_A^{-1}\boldsymbol{\xi}_A = O_p(N^{-1}T^{-1})$.

Next, notice that

$$\max_{1\le i\le N} |\boldsymbol{\xi}_A^\top \mathbf{\Sigma}_A \boldsymbol{\xi}_i| \le \max_{1\le i\le N} N^{-1/2}\boldsymbol{\xi}_i^\top \mathbf{\Sigma}_A \boldsymbol{\xi}_i| + \max_{1\le i\le N} N^{-1/2}|\sum_{j\ne i}\boldsymbol{\xi}_j^\top \mathbf{\Sigma}_A \boldsymbol{\xi}_i|.$$

By condition (C1), we have

$$\begin{aligned}
\boldsymbol{\xi}_i^\top \mathbf{\Sigma}_A \boldsymbol{\xi}_i &\overset{d}{=} \sigma_i^2 z_i^\top \mathbf{\Sigma}_{iT}^{1/2}\mathbf{\Sigma}_A \mathbf{\Sigma}_{iT}^{1/2} z_i, \\
N^{-1/2}\sum_{j\ne i}\boldsymbol{\xi}_j^\top \mathbf{\Sigma}_A \boldsymbol{\xi}_i &\overset{d}{=} N^{-1/2}\sum_{j\ne i}\sigma_i\sigma_j z_j^\top \mathbf{\Sigma}_{iT}^{1/2}\mathbf{\Sigma}_A \mathbf{\Sigma}_{iT}^{1/2} z_i \\
&\overset{d}{=} (1-1/N)^{1/2}\sigma_i \boldsymbol{e}_i \mathbf{\Sigma}_{Ai}^{1/2}\mathbf{\Sigma}_A \mathbf{\Sigma}_{iT}^{1/2} z_i,
\end{aligned}$$

where $\boldsymbol{e}_i \sim N(0, \mathbf{I}_p)$ is independent of z_i and $\mathbf{\Sigma}_{Ai} = \frac{1}{(N-1)}\sum_{j\ne i}\sigma_j^2 \mathbf{\Sigma}_{jT}$. By condition (C3) and Lemma 6.2, $\max_{1\le i\le N} N^{-1/2}\boldsymbol{\xi}_i^\top \mathbf{\Sigma}_A \boldsymbol{\xi}_i = O_p(N^{-1/2}\log(N))$. Note that

$$\begin{aligned}
&\max_{1\le i\le N}\left(N^{-1/2}\sum_{j\ne i}\boldsymbol{\xi}_j^\top \mathbf{\Sigma}_A \boldsymbol{\xi}_i\right)^2 \\
&\le \max_{1\le i\le N}\left(\sigma_i \boldsymbol{e}_i \mathbf{\Sigma}_{Ai}^{1/2}\mathbf{\Sigma}_A \mathbf{\Sigma}_{iT}^{1/2} z_i\right)^2 \\
&\le \max_{1\le i\le N}\left(\boldsymbol{e}_i \mathbf{\Sigma}_{Ai}^{1/2}\mathbf{\Sigma}_A \mathbf{\Sigma}_{iT}^{1/2} \boldsymbol{e}_i\right)\max_{1\le i\le N}\sigma_i^2\left(z_i \mathbf{\Sigma}_{Ai}^{1/2}\mathbf{\Sigma}_A \mathbf{\Sigma}_{iT}^{1/2} z_i\right).
\end{aligned}$$

By Lemma 6.2, $\max_{1\le i\le N}\left(z_i \mathbf{\Sigma}_{Ai}^{1/2}\mathbf{\Sigma}_A \mathbf{\Sigma}_{iT}^{1/2} z_i\right) = O_p(\log(N))$. Note that

$$\max_{1\le i\le N}\left(\boldsymbol{e}_i \mathbf{\Sigma}_{Ai}^{1/2}\mathbf{\Sigma}_A \mathbf{\Sigma}_{iT}^{1/2} \boldsymbol{e}_i\right) \le C\max_{1\le i\le N}\boldsymbol{e}_i^\top \boldsymbol{e}_i \overset{d}{=} C\max_{1\le i\le N}\boldsymbol{\zeta}_i,$$

where $\boldsymbol{\zeta}_i \sim \chi_p^2$. By Lemma 6.3,

$$\begin{aligned}P\left(\max_{1\le i\le N} \boldsymbol{\zeta}_i > 3\log(N)\right) &\le \sum_{i=1}^{N} P(\boldsymbol{\zeta}_i > 3\log(N))\\ &= NP(\chi_p^2 > 3\log(N))\\ &\le NP(\chi_p^2 > p + 2\sqrt{2.5p\log(N)} + 2.5\log(N))\\ &\le N\exp(-1.25\log(N)) = N^{-1/4} \to 0.\end{aligned}$$

Thus, $\max_{1\le i\le N} \boldsymbol{\zeta}_i = O_p(\log(N))$. Then, we have

$$\max_{1\le i\le N}\left(\left|N^{-1/2}\sum_{j\ne i}\boldsymbol{\xi}_j^\top \boldsymbol{\Sigma}_A \boldsymbol{\xi}_i\right|\right) = O_p(\log(N)).$$

Consequently, $\max_{1\le i\le N} \frac{2}{\sqrt{N(T-1)}}|\boldsymbol{\xi}_A^\top \boldsymbol{\Sigma}_A \boldsymbol{\xi}_i| = O_p\left(\frac{\log(N)}{\sqrt{NT}}\right)$.

Combining these facts together, we have

$$\max_{1\le i\le N}|\tilde{\sigma}_i^2 - \sigma_i^2| = O_p\left(\sqrt{\frac{\log(N)}{T}}\right) + O_p\left(\frac{1}{NT}\right) + O_p\left(\frac{\log(N)}{\sqrt{NT}}\right),$$

which completes the proof. □

6.2 Proof of Theorem 3.1

Under H_0, we have

$$\begin{aligned}\hat{\boldsymbol{\beta}}_i - \tilde{\boldsymbol{\beta}}_{\text{WFE}} = T^{-1/2}\boldsymbol{\Sigma}_{iT}^{-1}\boldsymbol{\xi}_i - T^{-1/2}N^{-1/2}\\ \left(N^{-1}\sum_{j=1}^{N}\tilde{\sigma}_i^{-2}\boldsymbol{\Sigma}_{iT}\right)^{-1}\left(N^{-1/2}\sum_{j=1}^{N}\tilde{\sigma}_i^{-2}\boldsymbol{\xi}_i\right).\end{aligned}$$

Define $\tilde{\boldsymbol{\Sigma}}_A = N^{-1}\sum_{j=1}^{N}\tilde{\sigma}_i^{-2}\boldsymbol{\Sigma}_{iT}$ and $\tilde{\boldsymbol{\xi}}_A = N^{-1/2}\sum_{j=1}^{N}\tilde{\sigma}_i^{-2}\boldsymbol{\xi}_i$. Then,

$$\begin{aligned}\tilde{S}_i &= \left(\hat{\boldsymbol{\beta}}_i - \tilde{\boldsymbol{\beta}}_{\text{WFE}}\right)^\top \frac{\mathbf{X}_i^\top \mathbf{M}\mathbf{X}_i}{\tilde{\sigma_i}^2}\left(\hat{\boldsymbol{\beta}}_i - \tilde{\boldsymbol{\beta}}_{\text{WFE}}\right)\\ &= \tilde{\sigma}_i^{-2}\boldsymbol{\xi}_i^\top\boldsymbol{\Sigma}_{iT}^{-1}\boldsymbol{\xi}_i - 2N^{-1/2}\tilde{\sigma}_i^{-2}\boldsymbol{\xi}_i^\top\tilde{\boldsymbol{\Sigma}}_A^{-1}\tilde{\boldsymbol{\xi}}_A + N^{-1}\tilde{\boldsymbol{\xi}}_A^\top\tilde{\boldsymbol{\Sigma}}_A^{-1}\boldsymbol{\Sigma}_{iT}\tilde{\boldsymbol{\Sigma}}_A^{-1}\tilde{\boldsymbol{\xi}}_A.\end{aligned}$$

By Condition (C1), we have $\sigma_i^{-2}\boldsymbol{\xi}_i^\top \boldsymbol{\Sigma}_{iT}^{-1}\boldsymbol{\xi}_i \sim \chi_p^2$. By Lemma 6.1, we have

$$P_{H_0}\left\{ \max_{1\le i\le N} \sigma_i^{-2}\boldsymbol{\xi}_i^\top \boldsymbol{\Sigma}_{iT}^{-1}\boldsymbol{\xi}_i - 2\log(N) - (p-2)\log(\log(N)) + 2\log\left(\Gamma(\frac{p}{2})\right) \le x \right\}$$
$$\to \exp(-\exp(-x/2)).$$

By Lemmas 6.2 and Lemma 6.4,

$$\begin{aligned}
&\left| \max_{1\le i\le N} \tilde{\sigma}_i^{-2}\boldsymbol{\xi}_i^\top \boldsymbol{\Sigma}_{iT}^{-1}\boldsymbol{\xi}_i - \max_{1\le i\le N} \sigma_i^{-2}\boldsymbol{\xi}_i^\top \boldsymbol{\Sigma}_{iT}^{-1}\boldsymbol{\xi}_i \right| \\
&\qquad \le \max_{1\le i\le N} \left|\tilde{\sigma}_i^2 - \sigma_i^2\right| \tilde{\sigma}_i^{-2} \max_{1\le i\le N} \sigma_i^{-2}\boldsymbol{\xi}_i^\top \boldsymbol{\Sigma}_{iT}^{-1}\boldsymbol{\xi}_i \\
&\qquad = O_p\left(\sqrt{\frac{\log(N)}{T}}\right) O_p(\log(N)) \\
&\qquad = O_p\left(\frac{\log^{3/2}(N)}{\sqrt{T}}\right) = o_p(1).
\end{aligned}$$

Next, we show that

$$\max_{1\le i\le N} N^{-1}\tilde{\boldsymbol{\xi}}_A^\top \tilde{\boldsymbol{\Sigma}}_A^{-1}\boldsymbol{\Sigma}_{iT}\tilde{\boldsymbol{\Sigma}}_A^{-1}\tilde{\boldsymbol{\xi}}_A = O_p(N^{-1}) = o_p(1).$$

By Condition (C1), we have

$$\tilde{\boldsymbol{\xi}}_A^\top \tilde{\boldsymbol{\Sigma}}_A^{-1}\boldsymbol{\Sigma}_{iT}\tilde{\boldsymbol{\Sigma}}_A^{-1}\tilde{\boldsymbol{\xi}}_A \overset{d}{=} \boldsymbol{z}^\top \tilde{\boldsymbol{\Sigma}}_A^{-1/2}\boldsymbol{\Sigma}_{iT}\tilde{\boldsymbol{\Sigma}}_A^{-1/2}\boldsymbol{z},$$

and by Condition (C3), the eigenvalues of $\tilde{\boldsymbol{\Sigma}}_A^{-1/2}\boldsymbol{\Sigma}_{iT}\tilde{\boldsymbol{\Sigma}}_A^{-1/2}$ are bounded. Thus,

$$\begin{aligned}
&\max_{1\le i\le N} N^{-1}\tilde{\boldsymbol{\xi}}_A^\top \tilde{\boldsymbol{\Sigma}}_A^{-1}\boldsymbol{\Sigma}_{iT}\tilde{\boldsymbol{\Sigma}}_A^{-1}\tilde{\boldsymbol{\xi}}_A \\
&\quad \le N^{-1}C\boldsymbol{z}^\top\boldsymbol{z} = O_p(N^{-1}).
\end{aligned}$$

By Cauchy's inequality, we have

$$\begin{aligned}
&\max_{1\le i\le N} N^{-1}\left(\tilde{\sigma}_i^{-2}\boldsymbol{\xi}_i^\top \tilde{\boldsymbol{\Sigma}}_A^{-1}\tilde{\boldsymbol{\xi}}_A\right)^2 \\
&\quad \le \max_{1\le i\le N} \tilde{\sigma}_i^{-2}\boldsymbol{\xi}_i^\top \boldsymbol{\Sigma}_{iT}^{-1}\boldsymbol{\xi}_i \times \max_{1\le i\le N} N^{-1}\tilde{\boldsymbol{\xi}}_A^\top \tilde{\boldsymbol{\Sigma}}_A^{-1}\boldsymbol{\Sigma}_{iT}\tilde{\boldsymbol{\Sigma}}_A^{-1}\tilde{\boldsymbol{\xi}}_A \\
&\quad = O_p(\log(N))O_p(N^{-1}) = o_p(1).
\end{aligned}$$

Thus,

$$P_{H_0}\left\{\max_{1\le i\le N}\tilde{S}_i - 2\log(N) - (p-2)\log(\log(N)) + 2\log\left(\Gamma\left(\frac{p}{2}\right)\right) \le x\right\}$$
$$\to \exp(-\exp(-x/2)).$$

6.3 Proof of Theorem 3.2

Under H_1, we have

$$\hat{\boldsymbol{\beta}}_i - \tilde{\boldsymbol{\beta}}_{\mathrm{WFE}} = T^{-1/2}\boldsymbol{\Sigma}_{iT}^{-1}\boldsymbol{\xi}_i - T^{-1/2}N^{-1/2}\left(N^{-1}\sum_{j=1}^{N}\tilde{\sigma}_i^{-2}\boldsymbol{\Sigma}_{iT}\right)^{-1}$$
$$\left(N^{-1/2}\sum_{j=1}^{N}\tilde{\sigma}_i^{-2}\boldsymbol{\xi}_i\right) + \hat{\boldsymbol{\omega}}_i,$$

where

$$\hat{\boldsymbol{\omega}}_i = \boldsymbol{\delta}_i - \left(N^{-1}\sum_{j=1}^{N}\tilde{\sigma}_i^{-2}\boldsymbol{\Sigma}_{iT}\right)^{-1}\left(N^{-1}\sum_{j=1}^{N}\tilde{\sigma}_i^{-2}\boldsymbol{\Sigma}_{iT}\boldsymbol{\delta}_i\right).$$

By Lemma 6.4 and Condition (C3), we have $\max_{1\le i\le N}|\hat{\boldsymbol{\omega}}_i - \boldsymbol{\omega}_i| = O_p(\log(N)/T)$. By the triangle inequality,

$$\begin{aligned}
&\max_{1\le i\le N} T\tilde{\sigma}_i^{-2}\hat{\boldsymbol{\omega}}_i^\top\boldsymbol{\Sigma}_{iT}\hat{\boldsymbol{\omega}}_i \ge \frac{1}{2}\max_{1\le i\le N} T\tilde{\sigma}_i^{-2}\boldsymbol{\omega}_i^\top\boldsymbol{\Sigma}_{iT}\boldsymbol{\omega}_i\\
&\quad - \max_{1\le i\le N} T\tilde{\sigma}_i^{-2}(\hat{\boldsymbol{\omega}}_i - \boldsymbol{\omega}_i)^\top\boldsymbol{\Sigma}_{iT}(\hat{\boldsymbol{\omega}}_i - \boldsymbol{\omega}_i)\\
&\ge \frac{1}{2}\max_{1\le i\le N} T\sigma_i^{-2}\boldsymbol{\omega}_i^\top\boldsymbol{\Sigma}_{iT}\boldsymbol{\omega}_i - \frac{1}{2}\max_{1\le i\le N} T\sigma_i^{-2}\boldsymbol{\omega}_i^\top\boldsymbol{\Sigma}_{iT}\boldsymbol{\omega}_i \times \max_{1\le i\le N}\left|\tilde{\sigma}_i^{-2} - \sigma_i^{-2}\right|\\
&\quad - \max_{1\le i\le N} T\tilde{\sigma}_i^{-2}(\hat{\boldsymbol{\omega}}_i - \boldsymbol{\omega}_i)^\top\boldsymbol{\Sigma}_{iT}(\hat{\boldsymbol{\omega}}_i - \boldsymbol{\omega}_i)\\
&\ge (8 + \frac{1}{2}\epsilon)\log(N) - O_p(\log^{3/2}(N)/T^{-1/2}) - O_p(\log^2(N)/T)\\
&\ge (8 + \frac{1}{4}\epsilon)\log(N),
\end{aligned}$$

as $N\to\infty$. According to the proof of Theorem 3.1, we have

$$P\left\{\max_{1\le i\le N} T\tilde{\sigma}_i^{-2}(\hat{\boldsymbol{\beta}}_i - \tilde{\boldsymbol{\beta}}_{\mathrm{WFE}} - \hat{\boldsymbol{\omega}}_i)^\top\boldsymbol{\Sigma}_{iT}(\hat{\boldsymbol{\beta}}_i - \tilde{\boldsymbol{\beta}}_{\mathrm{WFE}} - \hat{\boldsymbol{\omega}}_i)\right.$$
$$\left. - 2\log(N) - (p-2)\log(\log(N)) + 2\log(\Gamma(\frac{p}{2})) \le x\right\} \to \exp(-\exp(-x/2)).$$

Hence,

$$P\left\{\max_{1\le i\le N} T\tilde{\sigma}_i^{-2}(\hat{\boldsymbol{\beta}}_i-\tilde{\boldsymbol{\beta}}_{\mathrm{WFE}}-\hat{\boldsymbol{\omega}}_i)^\top\boldsymbol{\Sigma}_{iT}(\hat{\boldsymbol{\beta}}_i-\tilde{\boldsymbol{\beta}}_{\mathrm{WFE}}-\hat{\boldsymbol{\omega}}_i)\right.$$
$$\left.\le 2\log(N)+(p-1)\log(\log(N))\right\}\to 1,$$

by setting $x=\log(\log(N))+2\log(\Gamma(\frac{p}{2}))$. By the triangle inequality, we have

$$\begin{aligned}
&\max_{1\le i\le N} T\tilde{\sigma}_i^{-2}(\hat{\boldsymbol{\beta}}_i-\tilde{\boldsymbol{\beta}}_{\mathrm{WFE}})^\top\boldsymbol{\Sigma}_{iT}(\hat{\boldsymbol{\beta}}_i-\tilde{\boldsymbol{\beta}}_{\mathrm{WFE}})\\
&\qquad\ge \frac{1}{2}\max_{1\le i\le N} T\tilde{\sigma}_i^{-2}\hat{\boldsymbol{\omega}}_i^\top\boldsymbol{\Sigma}_{iT}\hat{\boldsymbol{\omega}}_i-\max_{1\le i\le N} T\tilde{\sigma}_i^{-2}(\hat{\boldsymbol{\beta}}_i\\
&\qquad\quad-\tilde{\boldsymbol{\beta}}_{\mathrm{WFE}}-\hat{\boldsymbol{\omega}}_i)^\top\boldsymbol{\Sigma}_{iT}(\hat{\boldsymbol{\beta}}_i-\tilde{\boldsymbol{\beta}}_{\mathrm{WFE}}-\hat{\boldsymbol{\omega}}_i)\\
&\qquad\ge (4+\frac{1}{8}\epsilon)\log(N)-2\log(N)-(p-1)\log(\log(N))\\
&\qquad\ge \frac{1}{16}\epsilon\log(N)+2\log(N)+(p-2)\log(\log(N))-2\log\left(\Gamma\left(\frac{p}{2}\right)\right)+q_\alpha,
\end{aligned}$$

with probability approaching one, as $N\to\infty$. Hence, $P(\Phi_\alpha=1)\to 1$.

6.4 Proof of Theorem 3.3

Lemma 6.5 *Suppose $Z_1,\ldots,Z_N$ are independent and identically distributed random sample from χ^2_p. Set $S_N=Z_1+\cdots+Z_N$, $\upsilon_N=(2pN)^{1/2}$ and $A_N=\{\frac{S_N-pN}{\upsilon_N}\le x\}$. For $y\in\mathbb{R}$, denote $l_N=2\log(N)+(p-2)\log(\log(N))-2\log(\Gamma(\frac{p}{2}))+y$ and $B_i=\{Z_i>l_N\}$. Then, for each $n\ge 1$,*

$$\sum_{1\le i_1<\cdots<i_n\le N}\left|P(A_NB_{i_1}\cdots B_{i_n})-P(A_N)\cdot P(B_{i_1}\cdots B_{i_n})\right|\to 0,$$

as $N\to\infty$.

Proof Write

$$S_N=\sum_{i=1}^N Z_i=\sum_{i=n+1}^N Z_i+\sum_{i=1}^n Z_i\doteq U_N+\Theta_N.$$

We will show the last term on the right hand side is negligible. By the definition, we have $\Theta_N\sim\chi^2_{pn}$. By Lemma 6.3, for any $n\ge 1$ and $\epsilon>0$, there exists $t=t_N>0$ with $\lim_{N\to\infty}t_N=\infty$ and N_0, depending on n,ϵ, such that

$$P(\Theta_N\ge\epsilon\upsilon_N)\le\frac{1}{N^t},$$

for $N \geq N_0$. Define

$$A_N(x) = \left\{ \frac{1}{\upsilon_N}(S_N - pN) \leq x \right\}, \quad x \in \mathbb{R},$$

for $N \geq 1$. From the fact $S_N = U_N + \Theta_N$ we see that

$$\begin{aligned} P(A_N(x)B_1 \cdots B_n) &\leq P\Big(A_N(x)B_1 \cdots B_n,\ \frac{|\Theta_N|}{\upsilon_N} < \epsilon\Big) + \frac{1}{N^t} \\ &\leq P\Big(\frac{1}{\upsilon_N}(U_N - pN) \leq x + \epsilon,\ B_1 \cdots B_n\Big) + \frac{1}{N^t} \\ &= P\Big(\frac{1}{\upsilon_N}(U_N - pN) \leq x + \epsilon\Big) \cdot P\big(B_1 \cdots B_n\big) + \frac{1}{N^t}, \end{aligned}$$

by the independence between U_N and Θ_N. Now,

$$\begin{aligned} P\Big(\frac{1}{\upsilon_N}(U_N - pN) \leq x + \epsilon\Big) &\leq P\Big(\frac{1}{\upsilon_N}(U_N - pN) \leq x + \epsilon,\ \frac{|\Theta_N|}{\upsilon_N} < \epsilon\Big) + \frac{1}{N^t} \\ &\leq P\Big(\frac{1}{\upsilon_N}(U_N + \Theta_N - pN) \leq x + 2\epsilon\Big) + \frac{1}{N^t} \\ &\leq P\big(A_N(x + 2\epsilon)\big) + \frac{1}{N^t}. \end{aligned}$$

Combining the two inequalities,

$$P(A_N(x)B_1 \cdots B_n) \leq P\big(A_N(x + 2\epsilon)\big) \cdot P\big(B_1 \cdots B_n\big) + \frac{2}{N^t}. \tag{6.1}$$

Similarly,

$$\begin{aligned} &P\Big(\frac{1}{\upsilon_N}(U_N - pN) \leq x - \epsilon,\ B_1 \cdots B_n\Big) \\ &\leq P\Big(\frac{1}{\upsilon_N}(U_N - pN) \leq x - \epsilon, B_1 \cdots B_n, \frac{|\Theta_N|}{\upsilon_N} < \epsilon\Big) + \frac{1}{N^t} \\ &\leq P\Big(\frac{1}{\upsilon_N}(S_N - pN) \leq x,\ B_1 \cdots B_n\Big) + \frac{1}{N^t}. \end{aligned}$$

By the independence between U_N and Θ_N,

$$P(A_N(x)B_1 \cdots B_n) \geq P\Big(\frac{1}{\upsilon_N}(U_N - pN) \leq x - \epsilon\Big) \cdot P(B_1 \cdots B_n) - \frac{1}{N^t}.$$

Furthermore,

$$P\Big(\frac{1}{\upsilon_N}(S_N - pN) \le x - 2\epsilon\Big) \le P\Big(\frac{1}{\upsilon_N}(S_N - pN) \le x - 2\epsilon,\ \frac{|\Theta_N|}{\upsilon_N} < \epsilon\Big) + \frac{1}{N^t}$$
$$\le P\Big(\frac{1}{\upsilon_N}(U_N - pN) \le x - \epsilon\Big) + \frac{1}{N^t},$$

due to the fact $S_N = U_N + \Theta_N$. Combining the two inequalities, we get

$$P(A_N(x)B_1 \cdots B_n) \ge P(A_N(x - 2\epsilon)) \cdot P(B_1 \cdots B_n) - \frac{2}{N^t}.$$

This, together with (6.1), concludes

$$\big|P(A_N(x)B_1 \cdots B_n) - P(A_N(x)) \cdot P(B_1 \cdots B_n)\big|$$
$$\le \Delta_{N,\epsilon} \cdot P(B_1 \cdots B_n) + \frac{2}{N^t},$$

for $N \ge N_0$, where

$$\Delta_{N,\epsilon} := |P(A_N(x)) - P(A_N(x + 2\epsilon))| + |P(A_N(x)) - P(A_N(x - 2\epsilon))|.$$

Similarly, for any $1 \le i_1 < i_2 < \cdots < i_n \le N$, we have

$$\big|P(A_N(x)B_{i_1} \cdots B_{i_n}) - P(A_N(x)) \cdot P(B_{i_1} \cdots B_{i_n})\big|$$
$$\le \Delta_{N,\epsilon} \cdot P(B_{i_1} \cdots B_{i_n}) + \frac{2}{N^t},$$

for $N \ge N_0$. As a result,

$$\begin{aligned}\zeta(N, n) :&= \sum_{1\le i_1<\cdots<i_n\le N} \big[P(A_N(x)B_{i_1} \cdots B_{i_n}) - P(A_N(x)) \cdot P(B_{i_1} \cdots B_{i_n})\big] \\ &\le \sum_{1\le i_1<\cdots<i_n\le N} \Big[\Delta_{N,\epsilon} \cdot P(B_{i_1} \cdots B_{i_n}) + \frac{2}{N^t}\Big] \\ &\le \Delta_{N,\epsilon} \cdot H(n, N) + \binom{N}{n} \cdot \frac{2}{N^t}, \end{aligned} \tag{6.2}$$

where

$$H(n, N) := \sum_{1\le i_1<\cdots<i_n\le N} P(B_{i_1} \cdots B_{i_n}).$$

First, by the central limit theorem,

$$\frac{S_N - pN}{\upsilon_N} \to N(0, 1) \text{ weakly},$$

as $N \to \infty$, and hence

$$\Delta_{N,\epsilon} \to |\Phi(x+2\epsilon) - \Phi(x)| + |\Phi(x-2\epsilon) - \Phi(x)|,$$

as $N \to \infty$, where $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt$. This implies that $\lim_{\epsilon \downarrow 0} \lim_{N\to\infty} \Delta_{N,\epsilon} = 0$. Second, by the independence of Z_i, we have

$$\begin{aligned} H(n, N) &= \sum_{1 \le i_1 < \cdots < i_n \le N} P(B_{i_1} \cdots B_{i_n}) \\ &= \binom{N}{n} P(B_1 \cdots B_n) = \binom{N}{n} P(B_1)^n \\ &= \binom{N}{n} \{P(\chi_p^2 > l_N)\}^n. \end{aligned}$$

As $N \to \infty$,

$$\log P(\max_{1\le i\le N} \chi_p^2 \le l_N) = N \log P(\chi_p^2 \le l_N) = N \log(1 - P(\chi_p^2 > l_N)) \sim N P(\chi_p^2 > l_N).$$

By Lemma 6.1, we have $P(\chi_p^2 > l_N) \sim \frac{1}{N} e^{-y/2}$. Thus,

$$\lim_{N\to\infty} H(n, N) = \frac{1}{n!} e^{-ny/2}, \tag{6.3}$$

for each $n \ge 1$. By using $\binom{N}{n} \le N^n$ and (6.2), for fixed $n \ge 1$, sending $N \to \infty$ first, then sending $\epsilon \downarrow 0$, we get $\lim_{N\to\infty} \zeta(N, n) = 0$, for each $n \ge 1$. The proof is completed. $\square$

Lemma 6.6 *Suppose $Z_1, \ldots, Z_N$ are independent and identically distributed random sample from χ_p^2, we have $\frac{\sum_{i=1}^N Z_i - pN}{\sqrt{2pN}}$ and $\max_{1\le i\le N} Z_i - 2\log(N) - (p-2)\log(\log(N)) + 2\log(\Gamma(\frac{p}{2}))$ are asymptotically independent, as $N \to \infty$.*

Proof Define $S_N = \sum_{i=1}^N Z_i$ and $\upsilon_N = \sqrt{2pN}$. By the central limit theorem, we have

$$\frac{S_N - pN}{\upsilon_N} \to N(0, 1) \text{ weakly}, \tag{6.4}$$

as $N \to \infty$. By Lemma 6.1, we have

$$\begin{aligned} &\max_{1\le i\le N} Z_i - 2\log(N) - (p-2)\log(\log(N)) + 2\log(\Gamma(\frac{p}{2})) \\ &\quad \to F(y) = \exp\Big\{ - e^{-y/2} \Big\} \end{aligned} \tag{6.5}$$

in distribution, as $N \to \infty$. To show the asymptotic independence, it suffices to prove

$$P\Big(\frac{S_N - pN}{\upsilon_N} \le x,\ \max_{1\le i\le N} Z_i - 2\log(N) - (p-2)\log(\log(N)) + 2\log(\Gamma(\frac{p}{2})) \le y\Big)$$
$$\to \Phi(x)\cdot F(y),$$

as $N \to \infty$, for any $x \in \mathbb{R}$ and $y \in \mathbb{R}$, where $\Phi(x) = (2\pi)^{-1/2}\int_{-\infty}^{x} e^{-t^2/2}\,dt$. Set

$$L_N = \max_{1\le i\le N} Z_i \ \text{ and } \ l_N = 2\log(N) + (p-2)\log(\log(N)) - 2\log(\Gamma(\frac{p}{2})) + y.$$

Because of (6.4) and (6.5), it is equivalent to show

$$\lim_{N\to\infty} P\Big(\frac{S_N - pN}{\upsilon_N} \le x,\ L_N > l_N\Big) = \Phi(x)\cdot[1 - F(y)], \tag{6.6}$$

for any $x \in \mathbb{R}$ and $y \in \mathbb{R}$. Define

$$A_N = \Big\{\frac{S_N - pN}{\upsilon_N} \le x\Big\} \quad \text{and} \quad B_i = \big\{Z_i > l_N\big\},$$

for $1 \le i \le N$. Therefore,

$$P\Big(\frac{1}{\upsilon_N}(S_N - pN) \le x,\ L_N > l_N\Big) = P\Big(\bigcup_{i=1}^{N} A_N B_i\Big). \tag{6.7}$$

Here the notation $A_N B_i$ stands for $A_N \cap B_i$. By the inclusion–exclusion principle,

$$P\Big(\bigcup_{i=1}^{N} A_N B_i\Big) \le \sum_{1\le i_1\le N} P(A_N B_{i_1}) - \sum_{1\le i_1<i_2\le N} P(A_N B_{i_1} B_{i_2}) + \cdots + \sum_{1\le i_1<\cdots<i_{2k+1}\le N} P(A_N B_{i_1}\cdots B_{i_{2k+1}}) \tag{6.8}$$

and

$$P\Big(\bigcup_{i=1}^{N} A_N B_i\Big) \ge \sum_{1\le i_1\le N} P(A_N B_{i_1}) - \sum_{1\le i_1<i_2\le N} P(A_N B_{i_1} B_{i_2}) + \cdots - \sum_{1\le i_1<\cdots<i_{2k}\le N} P(A_N B_{i_1}\cdots B_{i_{2k}}), \tag{6.9}$$

for any integer $k \geq 1$. Define

$$H(N,n) = \sum_{1 \leq i_1 < \cdots < i_n \leq N} P(B_{i_1} \cdots B_{i_n}),$$

for $n \geq 1$. From (6.3) we know

$$\lim_{n\to\infty} \limsup_{N\to\infty} H(N,n) = 0. \tag{6.10}$$

Set

$$\zeta(N,n) = \sum_{1 \leq i_1 < \cdots < i_n \leq N} \left[P(A_N B_{i_1} \cdots B_{i_n}) - P(A_N) \cdot P(B_{i_1} \cdots B_{i_n})\right],$$

for $n \geq 1$. By Lemma 6.5,

$$\lim_{N\to\infty} \zeta(N,n) = 0, \tag{6.11}$$

for each $n \geq 1$. The assertion (6.8) implies that

$$\begin{aligned} P\Big(\bigcup_{i=1}^{N} A_N B_i\Big) \leq P(A_N)\Big[&\sum_{1\leq i_1 \leq N} P(B_{i_1}) - \sum_{1 \leq i_1 < i_2 \leq N} P(B_{i_1} B_{i_2}) + \cdots - \\ &\sum_{1 \leq i_1 < \cdots < i_{2k} \leq N} P(B_{i_1} \cdots B_{i_{2k}})\Big] + \Big[\sum_{n=1}^{2k} \zeta(N,n)\Big] + H(N, 2k+1) \\ &\leq P(A_N) \cdot P\Big(\bigcup_{i=1}^{N} B_i\Big) + \Big[\sum_{n=1}^{2k} \zeta(N,n)\Big] + H(N,2k+1), \end{aligned} \tag{6.12}$$

where the inclusion–exclusion formula is used again in the last inequality, that is,

$$\begin{aligned} P\Big(\bigcup_{i=1}^{N} B_i\Big) \geq \sum_{1 \leq i_1 \leq N} P(B_{i_1}) - \sum_{1 \leq i_1 < i_2 \leq N} P(B_{i_1} B_{i_2}) + \cdots - \\ \sum_{1 \leq i_1 < \cdots < i_{2k} \leq N} P(B_{i_1} \cdots B_{i_{2k}}), \end{aligned}$$

for all $k \geq 1$. By the definition of l_N and (6.5),

$$\begin{aligned} P\Big(\bigcup_{i=1}^{N} B_i\Big) &= P\big(L_N > l_N\big) \\ &= P\Big(\max_{1\leq i \leq N} Z_i - 2\log(N) - (p-2)\log(\log(N)) + 2\log(\Gamma(\frac{p}{2})) > y\Big) \\ &\to 1 - F(y), \end{aligned}$$

as $N \to \infty$. By (6.4), $P(A_N) \to \Phi(x)$, as $N \to \infty$. From (6.7), (6.11) and (6.12), by fixing k first and sending $N \to \infty$, we obtain that

$$\limsup_{N\to\infty} P\Big(\frac{1}{\upsilon_N}(S_N - pN) \le x,\ L_N > l_N\Big) \le \Phi(x)\cdot[1-F(y)] + \lim_{N\to\infty} H(N, 2k+1).$$

Now, by letting $k \to \infty$ and using (6.10), we have

$$\limsup_{N\to\infty} P\Big(\frac{1}{\upsilon_N}(S_N - pN) \le x,\ L_N > l_N\Big) \le \Phi(x)\cdot[1-F(y)]. \tag{6.13}$$

By applying the same argument to (6.9), we see that the counterpart of (6.12) becomes

$$\begin{aligned} P\Big(\bigcup_{i=1}^{N} A_N B_i\Big) &\ge P(A_N)\Big[\sum_{1\le i_1\le N} P(B_{i_1}) - \sum_{1\le i_1<i_2\le N} P(B_{i_1}B_{i_2}) + \cdots + \\ &\qquad \sum_{1\le i_1<\cdots<i_{2k-1}\le N} P(B_{i_1}\cdots B_{i_{2k-1}})\Big] + \Big[\sum_{n=1}^{2k-1}\zeta(N,n)\Big] - H(N,2k) \\ &\ge P(A_N)\cdot P\Big(\bigcup_{i=1}^{N} B_i\Big) + \Big[\sum_{n=1}^{2k-1}\zeta(N,n)\Big] - H(N,2k), \end{aligned}$$

where in the last step we use the inclusion–exclusion principle, i.e.,

$$\begin{aligned} P\Big(\bigcup_{i=1}^{N} B_i\Big) \le \sum_{1\le i_1\le N} P(B_{i_1}) - \sum_{1\le i_1<i_2\le N} P(B_{i_1}B_{i_2}) + \cdots + \\ \sum_{1\le i_1<\cdots<i_{2k-1}\le N} P(B_{i_1}\cdots B_{i_{2k-1}}), \end{aligned}$$

for all $k \ge 1$. Review (6.7) and repeat the earlier procedure to see

$$\liminf_{N\to\infty} P\Big(\frac{1}{\upsilon_N}(S_N - pN) \le x,\ L_N > l_N\Big) \ge \Phi(x)\cdot[1-F(y)],$$

by sending $N \to \infty$ and then sending $k \to \infty$. This and (6.13) yield (6.6). The proof is completed.

□

Proof of Theorem 3.3 According to the proof of Theorem 3.2 in [18], we have $\tilde{\Delta}_{\mathrm{adj}} = S_a + O_p(T^{-1/2}) + O_p(N^{-1/2})$, where $S_a = (2pN)^{-1/2}\sum_{i=1}^{N}(\sigma_i^{-2}\boldsymbol{\xi}_i^\top \boldsymbol{\Sigma}_{iT}^{-1}\boldsymbol{\xi}_i - p)$. According to the proof of Theorem 3.2, we have $T_{\mathrm{Max}} = M_a + O_p(\sqrt{\frac{\log N}{N}}) +$

$O_p(\frac{\log^{3/2} N}{T^{1/2}})$, where $M_a = \max_{1\le i\le N} \sigma_i^{-2} \boldsymbol{\xi}_i^\top \boldsymbol{\Sigma}_{iT}^{-1} \boldsymbol{\xi}_i$. Given $\epsilon \in (0, 1)$, set $\Omega_N = \{|\tilde{\Delta}_{\text{adj}} - S_a| < \epsilon, |M - M_a| < \epsilon\}$. We have $\lim_{N,T\to\infty} P(\Omega_N) = 1$. By Lemma 6.6,

$$\begin{aligned} P(\tilde{\Delta}_{\text{adj}} \le x, T_{\text{Max}} > l_N) &\le P(\tilde{\Delta}_{\text{adj}} \le x, T_{\text{Max}} > l_N, \Omega_N) + P(\Omega_N^c) \\ &\le P(S_a \le x + \epsilon, M_a > l_N - \epsilon) + P(\Omega_N^c) \\ &\to \Phi(x + \epsilon)(1 - F(y - \epsilon)), \end{aligned}$$

as $N, T \to \infty$. Similarly, by Lemma 6.6,

$$\begin{aligned} P(\tilde{\Delta}_{\text{adj}} \le x, T_{\text{Max}} > l_N) &\ge P(\tilde{\Delta}_{\text{adj}} \le x, T_{\text{Max}} > l_N, \Omega_N) \\ &\ge P(S_a \le x - \epsilon, M_a > l_N + \epsilon) \\ &\to \Phi(x - \epsilon)(1 - F(y + \epsilon)), \end{aligned}$$

as $N, T \to \infty$. So

$$\Phi(x - \epsilon)(1 - F(y + \epsilon)) \le \lim_{N,T\to\infty} P(\tilde{\Delta}_{\text{adj}} \le x, T_{\text{Max}} > l_N) \le \Phi(x + \epsilon)(1 - F(y - \epsilon)).$$

Sending $\epsilon \to 0$, the conclusion follows. □

Acknowledgements Guanghui Wang was supported by the Natural Science Foundation of Shanghai (No. 23ZR1419400) and the National Key R&D Program of China (No. 2021YFA1000100, 2021YFA1000101, 2022YFA1003801). Ping Zhao and Long Feng was partially supported by Shenzhen Wukong Investment Company, the Fundamental Research Funds for the Central Universities under Grant No. ZB22000105, the China National Key R&D Program (Grant Nos. 2019YFC1908502, 2022YFA1003703, 2022YFA1003802, 2022YFA1003803) and the National Natural S0cience Foundation of China Grants (Nos. 12271271, 11925106, 12231011, 11931001 and 11971247).

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Ando, T., Bai, J.: A simple new test for slope homogeneity in panel data models with interactive effects. Econom. Lett. **136**, 112–117 (2015)
2. Bai, J.: Panel data models with interactive fixed effects. Econometrica **77**(4), 1229–1279 (2009)
3. Blomquist, J., Westerlund, J.: Testing slope homogeneity in large panels with serial correlation. Econom. Lett. **121**(3), 374–378 (2013)
4. Breitung, J., Roling, C., Salish, N.: Lagrange multiplier type tests for slope homogeneity in panel data models. Econom. J. **19**(2), 166–202 (2016)
5. Breusch, T.S., Pagan, A.R.: A simple test for heteroscedasticity and random coefficient variation. Econometrica **47**(5), 1287–1294 (1979)
6. Cai, T.T., Liu, W., Xia, Y.: Two-sample test of high dimensional means under dependence. J. R Stat. Soc. Ser. B Stat. Methodol. **76**(2), 349–372 (2014)
7. Chow, T.L., Teugels, J.L.: The sum and the maximum of iid random variables. In: Proceedings of the 2nd Prague Symposium on Asymptotic Statistics, vol 45, pp 394–403 (1978)
8. Embrechts, P., Klüppelberg, C., Mikosch, T.: Modelling extremal events: for insurance and finance, vol 33. Springer Science & Business Media (2013)

9. Fama, E.F., French, K.R.: Common risk factors in the returns on stocks and bonds. J. Financ. Econ. **33**(1), 3–56 (1993)
10. Feng, L., Jiang, T., Liu, B., Xiong, W.: Max-sum tests for cross-sectional independence of high-dimensional panel data. Ann. Statist. **50**(2), 1124–1143 (2022)
11. Hausman, J.A.: Specification tests in econometrics. Econometrica **46**(6), 1251–1271 (1978)
12. Hsing, T.: A note on the asymptotic independence of the sum and maximum of strongly mixing stationary random variables. Ann. Probab. **23**(2), 938–947 (1995)
13. Juhl, T., Lugovskyy, O.: A test for slope heterogeneity in fixed effects models. Econom. Rev. **33**(8), 906–935 (2014)
14. Laurent, B., Massart, P.: Adaptive estimation of a quadratic functional by model selection. Ann. Statist. **28**(5), 1302–1338 (2000)
15. Li, D., Xue, L.: Joint limiting laws for high-dimensional independence tests. arXiv preprint arXiv:1512.08819 (2015)
16. Liu, W.D., Lin, Z., Shao, Q.M.: The asymptotic distribution and Berry-Esseen bound of a new test for independence in high dimension with an application to stochastic optimization. Ann. Appl. Probab. **18**(6), 2337–2366 (2008)
17. Pesaran, H., Smith, R., Im, K.S.: Dynamic linear models for heterogenous panels. In: The Econometrics of Panel Data, Springer, pp 145–195 (1996)
18. Pesaran, M.H., Yamagata, T.: Testing slope homogeneity in large panels. J Econom. **142**(1), 50–93 (2008)
19. Phillips, P.C.B., Sul, D.: Dynamic panel estimation and homogeneity testing under cross section dependence. Econom. J. **6**(1), 217–259 (2003)
20. Swamay, P.A.V.B.: Efficient inference in a random coefficient regression model. Econometrica **38**, 311–323 (1970)
21. Wang, H.J., McKeague, I.W., Qian, M.: Testing for marginal linear effects in quantile regression. J. R Stat. Soc. Ser. B Stat. Methodol. **80**(2), 433–452 (2018)
22. Zellner, A.: An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. J. Amer. Statist. Assoc. **57**, 348–368 (1962)