A High Dimensional Two-Sample Test under a Low Dimensional Factor Structure

Yingying Ma[1], Wei Lan [2] and Hansheng Wang[1]

[1]*Peking University and* [2]*Southwestern University of Finance and Economics*

Abstract

Existing high dimensional two-sample tests usually assume that different elements of a high dimensional predictor are weakly dependent. Such a condition can be violated when data follow a low dimensional latent factor structure. As a result, the recently developed two-sample testing methods are not directly applicable. To fulfill such a theoretical gap, we propose here a Factor Adjusted two-Sample Testing (FAST) procedure to accommodate the low dimensional latent factor structure. Under the null hypothesis, together with fairly weak technical conditions, we show that the proposed test statistic is asymptotically distributed as a weighted chi-square distribution with a finite number of degrees of freedom. This leads to a totally different test statistic and inference procedure, as compared with those of Bai and Saranadasa (1996) and Chen and Qin (2010). Simulation studies are carried out to examine its finite sample performance. A real example on China stock market is analyzed for illustration purpose.

KEY WORDS: China Stock Market; High-dimensional data; Hypothesis testing; Latent factor structure; Two-Sample Test

*Both Yingying Ma and Hansheng Wang are from Guanghua School of Management, Peking University, Beijing, 100871, P. R. China. Wei Lan is from Statistics School and Center of Statistical Research, Southwestern University of Finance and Economics, Chengdu, Sichuan, 610074, P. R. China. The corresponding author: Wei Lan (lanwei@swufe.edu.cn).

1. INTRODUCTION

Let $X_{ki} = (X_{ki1}, \cdots, X_{kip})^\top \in \mathbb{R}^p$ be the information collected from the ith subject in the kth($k = 1, 2$) group, with $E(X_{ki}) = \mu_k \in \mathbb{R}^p$ and $\mathrm{cov}(X_{ki}) = \Sigma$. Then, the classical two-sample test concerns about testing the null hypothesis $H_0 : \mu_1 = \mu_2$ against the alternative $H_1 : \mu_1 \neq \mu_2$. This is a problem of fundamental importance, and has been commonly encountered in many scientific applications, including economics, finance, genetics, and many others. The corresponding testing procedures (e.g., Hotelling's T^2 test) are well studied when p is fixed (Anderson, 2003). However, when p is much larger than n, the traditional methods, such as Hotelling's T^2 test, cannot be computed, as the sample covariance matrix is not invertible. This makes two-sample test for ultra high dimensional data a problem of importance; see, for example, Bai and Saranadasa (1996), Chen and Qin (2010), Srivastava et al. (2013), and Cai et al. (2014).

Recently, a number of useful methods have been developed for high dimensional two-sample test. Nevertheless, their applicability relies on one critical assumption. That is different elements of X_{ki} for $k = 1, 2$ should be weakly dependent. Mathematically, this requires that $tr(\Sigma^4) = o\{tr^2(\Sigma^2)\}$; see, for example, condition (2.8) in Zhong and Chen (2011) and condition (3.7) in Chen and Qin (2010). It is worthy mentioning that such an assumption can be violated if the eigenvalues of Σ are dominated by a few top ones. This can happen if X_{ki} follows a low dimensional latent factor structure (Fan et al., 2011; Wang, 2012). For the purpose of illustration, we assume that $X_{ki} = BZ_{ki} + \varepsilon_{ki}$, where each element of the factor loading $B \in \mathbb{R}^{p\times d}$, common factor $Z_{ki} \in \mathbb{R}^d$, and random error ε_{ki} are all independently generated from a standard normal distribution, with $d > 0$ is the finite number of common factors. In this setting, one can verify that $tr(\Sigma^4) = dp^4\{1 + o(1)\}$ and $tr(\Sigma^2) = dp^2\{1 + o(1)\}$. As a result, $tr(\Sigma^4)/tr^2(\Sigma^2) \to 1/d \neq 0$, which makes condition (2.8) in Zhong and Chen (2011) and condition (3.7) in Chen and Qin (2010) invalid. As a result, how to construct a testing procedure for data of this type becomes a problem of interest.

The aforementioned problem is also empirically motivated. Consider, for example, if our interest is to test Monday effect in China stock market, for assessing stock market efficiency (Keim and Ziemba, 2000). To this end, we treat trading days (not stocks) as samples, and classify different trading days into two groups, according to whether they are Monday ($k = 1$) or not ($k = 2$). For a given trading day i in the kth group, we use X_{kij} to stand for the jth stock return on this particular trading day in percentage (%). Thus, $X_{ki} = (X_{ki1}, \cdots, X_{kip})^\top \in \mathbb{R}^p$ records all the stock returns on the (k, i)th day. Then, one way to evaluate Monday effect is to test whether $\mu_1 = \mu_2$. It is remarkable that for this problem the existing methods of Bai and Saranadasa (1996) and Chen and Qin (2010) cannot be directly applied. The reason is that different stock returns are all heavily correlated with at least one common factor, that is the market index (Sharpe, 1964; Fama and French, 1993). This makes the leading eigenvalue of the covariance matrix Σ extremely large, which violates condition (3.8) in Bai and Saranadasa (1996) and condition (3.7) in Chen and Qin (2010) seriously. Thus, this application calls for a new method, which can accommodate such a highly singular eigenvalue structure.

Motivated by the theoretical and practical demand, we aim to develop a two-sample testing procedure for data admitting a low dimensional latent factor structure. Specifically, we investigate the asymptotic distribution of $\|\bar{X}_1 - \bar{X}_2\|$ under a low dimensional latent factor model setup, where $\bar{X}_k$ stands for the sample mean of the kth group. We demonstrate theoretically that such a simple discrepancy measure is asymptotically distributed as a weighted chi-square distribution with a finite degrees of freedom. That leads to a totally different test statistic and inference procedure, as compared with those of Bai and Saranadasa (1996) and Chen and Qin (2010). Extensive simulation studies are conducted to demonstrate its finite sample performance. A real data example is also presented for illustration purpose.

The rest of the paper is organized as follows. Section 2 introduces the methodology with both model assumptions and asymptotic theories. Section 3 presents numerical studies based on both simulation and real dataset. The article is concluded with a short discussion in Section 4. All technical details are relegated to the Appendix.

2. METHODOLOGY

2.1. Model and Notation

Let $X_{ki} = (X_{ki1}, \cdots, X_{kip})^\top \in \mathbb{R}^p$ be a p-dimensional vector collected from the ith subject ($1 \le i \le n_k$) in the kth ($1 \le k \le 2$) group. Write $E(X_{ki}) = \mu_k = (\mu_{k1}, \cdots, \mu_{kp})^\top \in \mathbb{R}^p$ and assume $\mathrm{cov}(X_{ki}) = \Sigma = (\sigma_{j_1 j_2}) \in \mathbb{R}^{p\times p}$. Then, the hypotheses of interest are given by

$$H_0 : \mu_1 = \mu_2 \quad \text{vs.} \quad H_1 : \mu_1 \ne \mu_2. \tag{2.1}$$

To test the null hypotheses in (2.1), it is natural to consider the following Hotelling's test statistic. That is,

$$T_{\text{Hotelling}} = \left(\frac{n_1 n_2}{n_1 + n_2}\right) \left(\bar{X}_1 - \bar{X}_2\right)^\top \hat{\Sigma}^{-1} \left(\bar{X}_1 - \bar{X}_2\right),$$

where $\hat{\Sigma} = \sum_{k=1}^2 \sum_{i=1}^{n_k} (X_{ki} - \bar{X}_k)(X_{ki} - \bar{X}_k)^\top/(n_1 + n_2 - 2)$ is the sample covariance matrix. Assuming p is fixed and X_{ki} is normally distributed. Under the null hypothesis of (2.1), $T_{\text{Hotelling}}$ follows a Hotelling's T^2 distribution with (p, $n_1 + n_2 - 2$) degrees of freedom. Nevertheless, if the data dimension p is considerably larger than the sample size n, the story changes. In that situation, $\hat{\Sigma}$ is not invertible. As a result, the test statistic $T_{\text{Hotelling}}$ is no longer computable.

2.2. Existing Methods

To fix the aforementioned problem with $p \gg n$, Bai and Saranadasa (1996) proposed the following test statistic

$$T_{BS} = \|\bar{X}_1 - \bar{X}_2\|^2 - \left(\frac{n_1 + n_2}{n_1 n_2}\right) tr(\widehat{\Sigma}).$$

Assuming appropriate regularity conditions and also $p/n \to c$ for some constant $c > 0$, Bai and Saranadasa (1996) demonstrates that $T_{BS}/\text{var}^{1/2}(T_{BS})$ follows a standard normal distribution asymptotically. Such a test was further improved by Chen and Qin (2010) to the following test statistic

$$T_{CQ} = \frac{\sum_{i \neq j}^{n_1} X_{1i}^\top X_{1j}}{n_1(n_1 - 1)} + \frac{\sum_{i \neq j}^{n_2} X_{2i}^\top X_{2j}}{n_2(n_2 - 1)} - 2\frac{\sum_{i=1}^{n_1} \sum_{j=1}^{n_2} X_{1i}^\top X_{2j}}{n_1 n_2}.$$

Under the null hypothesis of (2.1) , Chen and Qin (2010) proved that $T_{CQ}/\text{var}^{1/2}(T_{CQ})$ follows a standard normal distribution, even if $p/n \to \infty$.

For the statistical validity of the aforementioned tests, appropriate technical conditions have to be assumed. Among all the conditions, Bai and Saranadasa (1996) requires that

$$\lambda_{\max}(\Sigma) = o(p^{1/2}), \tag{2.2}$$

where $\lambda_{\max}(A)$ stands for the largest eigenvalue of an arbitrary positive definite matrix A. In contrast, the method of Chen and Qin (2010) replaced the condition (2.2) by another regularity assumption, as

$$tr(\Sigma^4) = o\Big\{tr^2(\Sigma^2)\Big\}. \tag{2.3}$$

We find that both (2.2) and (2.3) are reasonable assumptions if the covariance matrix Σ is not highly singular. This is particularly true if the data are generated according to a diverging latent factor model; see Bai and Saranadasa (1996) and Chen and Qin (2010). Nevertheless, depending on the real problem, we do encounter situations where a few top eigenvalues of Σ completely dominate the rest. This can happen if the data admit a low dimensional factor structure. We then investigate its theoretical properties in the next section.

2.3. A Factor Model

To model the covariance structure of X_{ki}, we consider the following standard latent factor model (Fan et al., 2008)

$$X_{ki} = \mu_k + BZ_{ki} + \varepsilon_{ki}, \quad k = 1, 2, \tag{2.4}$$

where the loading matrix $B = (b_1, \cdots, b_p)^\top \in \mathbb{R}^{p\times d}$ and $b_j = (b_{j1}, \cdots, b_{jd})^\top \in \mathbb{R}^{d\times 1}$ for each $1 \leq j \leq p$. $Z_{ki} = (Z_{ki1}, \cdots, Z_{kid})^\top \in \mathbb{R}^d$ is the unobserved latent vector, and $\varepsilon_{ki} = (\varepsilon_{ki1}, \cdots, \varepsilon_{kip})^\top \in \mathbb{R}^p$ is the random noise. Here, d is a fixed number, standing for the latent factor dimension. Because Z_{ki} is unobserved latent vector, we can assume that $E(Z_{ki}) = 0$ and $\mathrm{cov}(Z_{ki}) = I_d$, where I_d stands for a d-dimensional identity matrix. Otherwise, we can always re-define $B := B\mathrm{cov}^{1/2}(Z_{ki})$ and $Z_{ki} := \mathrm{cov}^{-1/2}(Z_{ki})Z_{ki}$, so that these assumptions are well satisfied. We assume further that Z_{ki} and ε_{ki} are mutually independent. Lastly, let $E(\varepsilon_{ki}) = 0$ and $\mathrm{cov}(\varepsilon_{ki}) = D$. We require different components of ε_{ki} to be mutually independent, and their fourth moment are finite. Thus, D is a diagonal covariance matrix with positive diagonal entries. We further assume that D's diagonal components are all bounded as $p \to \infty$. Denote $\Sigma_{B,p} = p^{-1}B^\top B$, we then assume that

$$\Sigma_{B,p} = \Sigma_B + o(1) \tag{2.5}$$

for some positive definite matrix $\Sigma_B \in \mathbb{R}^{d\times d}$ as $p \to \infty$. To study Σ's eigenvalue structure, we define $\lambda_l(A)$ to be the l-th largest eigenvalue of an arbitrary positive definite matrix A. We know immediately $\lambda_1(\Sigma) = \lambda_{\max}(\Sigma)$. We then have the following theorem.

Theorem 1. *Assuming the factor model (2.4) and the technical condition (2.5), we should have: (1) $\lambda_l(\Sigma)/p = \lambda_l(\Sigma_B) + o(1)$ for any $1 \leq l \leq d$; and (2) $\lambda_l(\Sigma) \leq C$ for any $l > d$ and some constant C.*

By Theorem 1, we know that the condition (2.2) is obviously violated. Specifically, from model (2.4), one can verify that $\Sigma = BB^\top + D$. Then, $tr(\Sigma^4) = \sum_{l=1}^p \lambda_l^4(\Sigma) \geq tr\{(BB^\top)^4\} = tr\{(B^\top B)^4\} = p^4 \cdot tr(\Sigma_{B,p}^4)$ and $tr(\Sigma^2) \leq 2p^2(tr\Sigma_{B,p}^2 + trD^2/p^2)$ by Cauchy-Schwarz inequality. Thus, via condition (2.5), we have

$$\frac{tr(\Sigma^4)}{tr^2(\Sigma^2)} \geq \frac{p^4 \cdot tr(\Sigma_{B,p}^4)}{4p^4(tr\Sigma_{B,p}^2 + trD^2/p^2)^2} \not\to 0.$$

This also violates the condition (2.3).

2.4. A Preliminary Investigation

To gain some intuition, we temporarily assume that X_{ki}s are normally distributed. Then, under the null hypothesis of (2.1), we should have $\bar{X}_1 - \bar{X}_2$ normally distributed with mean 0 and covariance matrix $(1/n_1 + 1/n_2)\Sigma$. Consequently, we should have

$\|\bar{X}_1 - \bar{X}_2\|^2(n_1n_2)/(n_1+n_2)$ following a weighted chi-square distribution, where the weights are given by the eigenvalues of Σ. If the eigenvalues are approximately of the same order, such a a weighted chi-square distribution can be further approximated by a normal distribution, after appropriate linear normalization; see, for example, Bai and Saranadasa (1996) and Chen and Qin (2010). Nevertheless, under the factor model (2.4) and assumption (2.5), we know that Σ's first d eigenvalues (or weights) are considerably larger than the rest. As a result, the weighted chi-square distribution for $\|\bar{X}_1 - \bar{X}_2\|^2(n_1n_2)/(n_1+n_2)$ cannot be approximated by any normal distribution. In fact, by the following Theorem 2, we know that $\|\bar{X}_1 - \bar{X}_2\|^2(n_1n_2)/(n_1+n_2)$ admits a weighted chi-square distribution with a finite degrees of freedom, after appropriate linear adjustment.

Theorem 2. *Assuming the factor model (2.4), $n_1/n_2 \to \hbar$ for some constant $\hbar > 0$, the technical condition (2.5), and the null hypothesis of (2.1), then $p^{-1}\|\bar{X}_1 - \bar{X}_2\|^2(n_1n_2)/(n_1+n_2) - p^{-1}tr(D)$ is asymptotically distributed as $\sum_{l=1}^{d} \lambda_l(\Sigma_B) \cdot \chi_l^2(1)$, where $\chi_l^2(1)$ stands for the l-th independent chi-square random variable with 1 degree of freedom.*

2.5. Factor Adjusted Two-Sample Test

To make use of the previous theoretical result, we need to estimate $tr(D)$ and the latent factor number d consistently. Recall that $\Sigma = BB^\top + D$. Thus, we have $tr(\Sigma)/p = tr(BB^\top)/p + tr(D)/p = tr(B^\top B)/p + tr(D)/p = tr(\Sigma_{B,p}) + tr(D)/p$. Next, by condition (2.5), we know that $tr(\Sigma_{B,p}) = tr(\Sigma_B) + o(1)$. This leads to $tr(\Sigma)/p = tr(\Sigma_B) + tr(D)/p + o(1)$. By Theorem 1(1), we further have $tr(\Sigma)/p = p^{-1}\sum_{l=1}^{d} \lambda_l(\Sigma) + tr(D)/p + o(1)$. As a result, we are motivated to estimate $tr(D)/p$ by $\{tr(\widehat{\Sigma}) - \sum_{l=1}^{d} \lambda_l(\widehat{\Sigma})\}/p$, when p is very large. Moreover, the number of common factors d can be selected by maximizing the following eigenvalue ratio method as

$\hat{d} = \text{argmax}_{l \le n_0 - 1} \lambda_l(\widehat{\Sigma})/\lambda_{l+1}(\widehat{\Sigma})$, where $n_0 = \min\{n_1, n_2\}$; see Lam and Yao (2012) and Wang (2012). Then, we can demonstrate the following results,

Theorem 3. *Assuming the factor model (2.4) and the technical condition (2.5), We then have (1)* $P(\hat{d} = d) \to 1$, *and (2)* $\{tr(\widehat{\Sigma}) - \sum_{l=1}^{\hat{d}} \lambda_l(\widehat{\Sigma})\}/p - tr(D)/p \to_p 0$.

By Theorem 3, we know that the latent factor dimension d can be consistently estimated by $\hat{d}$. We then propose the following test statistic, as

$$T_{\text{FAST}} = p^{-1}\Big(\frac{n_1 n_2}{n_1 + n_2}\Big)\Big\|\bar{X}_1 - \bar{X}_2\Big\|^2 - p^{-1}\Big\{tr\Big(\widehat{\Sigma}\Big) - \sum_{l=1}^{\hat{d}} \lambda_l(\widehat{\Sigma})\Big\}.$$

T_{FAST} should follow a weighted chi-square distribution as $\sum_{l=1}^{d} \lambda_l(\Sigma_B) \cdot \chi_l^2(1)$, whose upper α quantile is denoted by q_α.

By Theorem 1(1), we know that $p^{-1}\lambda_l(\Sigma) = \lambda_l(\Sigma_B) + o(1)$ for any $1 \le l \le d$. This suggests that $\sum_{l=1}^{d} \lambda_l(\Sigma_B) \cdot \chi_l^2(1)$ and $p^{-1}\sum_{l=1}^{d} \lambda_l(\Sigma) \cdot \chi_l^2(1)$ should have identical asymptotic distribution as $p \to \infty$. Next, by Lemma A.11 of Ahn and Horenstein (2013), we know that $p^{-1}\lambda_l(\hat{\Sigma}) - p^{-1}\lambda_l(\Sigma) \to_p 0$ for every $1 \le l \le d$. Thus, $p^{-1}\sum_{l=1}^{d} \lambda_l(\Sigma) \cdot \chi_l^2(1)$ can be empirically approximated by $p^{-1}\sum_{l=1}^{d} \lambda_l(\hat{\Sigma}) \cdot \chi_l^2(1)$. Denote $\hat{q}_\alpha$ be the upper α quantile of $p^{-1}\sum_{l=1}^{d} \lambda_l(\hat{\Sigma}) \cdot \chi_l^2(1)$, where $\hat{q}_\alpha$ is estimated through a simulation-based numerical method. More specifically, we simulate a large number of independent and identically distributed random variables, according to the weighted chi-square distribution $p^{-1}\sum_{l=1}^{d} \lambda_l(\hat{\Sigma}) \cdot \chi_l^2(1)$. Then, the empirical αth quantile of these simulated random variables can be treated as $\hat{q}_\alpha$. We then reject the null hypothesis of (2.1), if $T_{\text{FAST}} > \hat{q}_\alpha$. We refer to the above testing procedure as a Factor Adjusted two Sample Testing (FAST) method.

Let $p^{-1}n_1 n_2\|\mu_1 - \mu_2\|^2/(n_1 + n_2) = \delta$ measure the discrepancy between μ_1 and μ_2 under the alternative hypothesis. The asymptotic power of the test is given by the following theorem, where $F(\cdot)$ stands for the cumulative distribution function of $\sum_{l=1}^{d} \lambda_l(\Sigma_B) \cdot \chi_l^2(1)$.

Theorem 4. *Assuming the factor model (2.4), the technical condition (2.5), and the alternative hypothesis of (2.1) hold with $\delta \to \infty$ as $n_0 \to \infty$. Then we have the power $P(T_{\scriptscriptstyle FAST} > \hat{q}_\alpha) \to 1$, as n goes to infinity.*

According to Theorem 4, the power of the test is largely affected by δ. The asymptotic power of the test tends to 1, as $\delta \to \infty$. Specifically, if n_1 and n_2 are of comparable size such that $n_1 n_2/(n_1 + n_2) = O(n)$, where $n = n_1 + n_2$ is the total sample size. Then the proposed test is consistent as n goes to infinity), as long as $p^{-1}\|\mu_1 - \mu_2\|^2$ is of larger order of $O(1/n)$.

3. NUMERICAL STUDIES

3.1. Simulation Results

To gauge the finite sample performance of the FAST test statistic, we conduct extensive simulation studies in this section. We consider five different predictor dimensions ($p = 100, 200, 300, 400, 500$), two different sample sizes ($n = 100, 200$), and two different numbers of latent factors ($d = 1, 3$). The data are simulated according to (2.4), where the elements of the random error ε_{ki}, the common factor Z_{ki}, and the factor loading B are all independently generated from a standard normal distribution.

For each fixed parameter setup (i.e., n, p, and d), a total of 1,000 random replications have been conducted with the nominal level $\alpha = 5\%$. For each replication, we use the aforementioned eigenvalue ratio method to estimate the number of common factors d. Recall that $\hat{d}$ represents the estimated number of common factors. We find that the percentage of $\hat{d} = d$ is always 100%. Thus, the eigenvalue ratio method is indeed a reliable method for factor number estimation. This finding perfectly corroborates the first part of Theorem 3. For comparison purpose, we also present the results for testing the null hypothesis $H_0 : \mu_1 = \mu_2$ by using the methods proposed by Bai and Saranadasa (1996) and Chen and Qin (2010) respectively. The sizes of the three tests are reported in Figure 1. Here, we use "FAST" to stand for our proposed test, "CQ" for that of Chen and Qin (2010), and "BS" for that of Bai and Saranadasa (1996).

A well-behaved test should have its empirical size around 5%. As one can see from Figure 1, FAST can generally control the size well across different simulation scenarios. The reported empirical sizes of FAST are always very close to its nominal level 5%, with no more than one percentage point difference in most cases. By contrast, the sizes of the CQ and BS tests are alarmingly larger than the nominal significance level. Such a finding is not surprising, as both CQ and BS tests are not designed for data with a low dimensional latent factor structure.

To give a more clear picture about the asymptotic properties of the three tests under the null hypothesis, we next compare the empirical density of the three tests against their asymptotic counterparts in Figure 2. To save space, we only consider the case with $n = 200, p = 500$, and $d = 3$ for the purpose of illustration. As we can see from the first graph, the density curve of FAST is reasonably close to that of the weighted chi-square distribution. In contrast, the empirical densities of the other two test statistics are apparently deviated from the standard normal distribution, which is consistent with our theoretical findings.

We next study the power of the proposed test. For simplicity, we fix $\mu_1 = 0$ and the percentages of $\mu_{2j} \neq 0$ are chosen to be 5%, 10%, 25%, 50% respectively. For each $\mu_{2j} \neq 0$, we generate it independently from $U(0, 1)$. Table 1 reports the empirical power of FAST with different (n, p, d) combinations. As one can see, the power of FAST rises up steadily with the percentage of nonzero μ_{2j}s increases. When the percentage of nonzero μ_{2j}s approaches 50%, the empirical power of FAST becomes very close to 1 for most simulation cases.

3.2. Real Data Analysis

To further demonstrate the practical usefulness of the proposed method, we consider here an empirical example to test Monday effect in China stock market (Keim and Ziemba, 2000). One way to do this is to test whether Monday stock returns are equal to those of other trading days on average. Let μ_1 and μ_2 represent the population stock returns on Monday and the other four trading days (Tuesday, Wednesday, Thursday and Friday), respectively. Then the hypothesis of interest is $H_0 : \mu_1 = \mu_2$ vs. $H_1 : \mu_1 \neq \mu_2$. To solve the problem, we download a dataset from the Resset Database (*http://www.resset.cn/*). After deleting the stocks with missing values, we obtain a total of 111 Monday records, and 449 non-Monday records on a total of $p = 895$ number of stocks, ranging from 01/01/2012 to 01/05/2013.

Due to the existence of market factor, the leading eigenvalue of $\widehat{\Sigma}$ is much larger than the rest. For this case, we have $\lambda_1(\widehat{\Sigma}) = 1810.4$, which is much larger than $\lambda_2(\widehat{\Sigma}) = 160.1$. Furthermore, $\lambda_1(\widehat{\Sigma})$ itself accounts for $\lambda_1(\widehat{\Sigma})/\sum_{l=1}^{p}\lambda_l(\widehat{\Sigma}) = 32.7\%$ of the total volatility. We further verify $tr(\widehat{\Sigma}^4)/tr^2(\widehat{\Sigma}^2) = 0.9 \not\to 0$, which clearly violates condition (2.8) in Zhong and Chen (2011) and condition (3.7) in Chen and Qin (2010). This makes our method useful for this case. To implement FAST, we firstly need to estimate the dimension of latent factors. By the maximum eigenvalue ratio method, we obtain $\hat{d} = 1$. Based on $\hat{d} = 1$, we compute the FAST test statistic and compare this value with the upper 5% quantile for the weighted chi-square distribution $p^{-1}\sum_{l=1}^{\hat{d}}\lambda_l(\widehat{\Sigma})\chi_l^2(1)$. Here, our test statistic value $T_{\text{FAST}} = 2.4$ and the corresponding $\hat{q}_\alpha = 7.8$. As a result, no evidence supporting the existence of Monday effect was detected.

4. CONCLUDING REMARKS

In this article, we propose a FAST procedure for high dimensional two-sample test when the data of the two groups follow a low dimensional latent factor structure (Fan et al., 2011; Wang, 2012). Under the null hypothesis, together with fairly weak conditions, we show that the proposed test statistic is asymptotically distributed as a weighted chi-square distribution with a finite degrees of freedom. The new method is different from those of Bai and Saranadasa (1996) and Chen and Qin (2010). The finite sample performance and the usefulness of the new method is further demonstrated by extensive simulation studies and a real example on China stock market.

It is worth mentioning that when the covariance matrix is highly singular, another common approach to do high dimensional two-sample test is to conduct principal component analysis first. Thereafter, Hotelling's T^2 test can be carried out based on the first few principal components. However, this approach may suffer from some problems. To fix the idea, consider for example there is only one latent factor (i.e., $d = 1$). In this case, only one principal component should be used for Hotelling's T^2 test. Assuming further the direction of the first principal component is given by an eigenvector $\gamma \in \mathbb{R}^p$. Then, the corresponding principal component scores for each subject can be computed as $s_{ki} = X_{ki}^\top \gamma$ for $k = 1, 2$, whose expectation is $E(s_{ki}) = \mu_k^\top \gamma$. The resulting mean difference in terms of the principal component score is $\Delta = E(s_{1i}) - E(s_{2i}) = (\mu_1 - \mu_2)^\top \gamma$, which equals to 0 if $\mu_1 - \mu_2$ and γ are orthogonal with each other. This suggests that, even if the alternative hypothesis is correct (i.e., $\mu_1 \neq \mu_2$), the test conducted based on principal component scores might not have much power to detect it, if the principal component direction γ happens to be orthogonal or nearly orthogonal to the mean difference vector $\mu_1 - \mu_2$. This explains why T^2 test cannot be the only best choice for high dimensional mean test.

To conclude the article, we identify the following possible research avenues. First, in this article, we assume that the two populations have the same covariance matrix. Thus, in order to use FAST test, one needs to firstly test $H_0 : \Sigma_1 = \Sigma_2$, where $\Sigma_i = \mathrm{cov}(X_i)$ represents the covariance matrix of the two populations respectively. To this end, some existing testing procedures can be used; see, for example, Li and Chen (2012) and Cai et al. (2013). If the null hypothesis of $H_0 : \Sigma_1 = \Sigma_2$ has been rejected, FAST cannot be applied immediately and some modifications are needed. Second, it is also interesting to have a more general test like $H_0 : R\mu = r$, where $\mu = (\mu_1, \mu_2)^\top$, $R \in \mathbb{R}^{q\times 2}$ is a known matrix, and $r \in \mathbb{R}^q$ is a constant vector. Notice that, the null hypotheses of (2.1) is a special case with $R = (1, -1)$ and $r = 0$. More residual efforts along this direction are needed.

APPENDIX

Appendix A. Two Useful Lemmas

To prove the main theorems, the following two lemmas are useful. Lemma 1 can be derived directly through Weyl's inequality for Hermitian matrices, while Lemma 2 is slightly modified from Lemma 2 of Lan et al. (2014), so that the proofs of the two lemmas are thus omitted to save space.

Lemma 1. *For any symmetric matrix $A \in \mathbb{R}^{n\times n}$ and positive semi-definite matrix $B \in \mathbb{R}^{n\times n}$, we have (1). $\lambda_k(A+B) \geq \lambda_k(A)$ for $k = 1, \cdots, n$, and (2). $\lambda_{j+k-1}(A+B) \leq \lambda_j(A) + \lambda_k(B)$ for any $j + k \leq n + 1$.*

Lemma 2. *Let $V = (V_1, \cdots, V_n)^\top \in \mathbb{R}^n$ be a random vector with $E(V) = 0$ and $cov(V) = I_n$, and V_ks to be mutually independent. Then, for any $n \times n$ positive definite* *matrix A, we have that $var(V^\top AV) = 2tr(A^2) + \phi tr(A^{\otimes 2}) \le (2+\phi)tr(A^2)$, where $A = (a_{j_1j_2})$, $A^{\otimes 2} = (a^2_{j_1j_2})$, and $\phi = E(V_i^4) - 3 < \infty$.*

Appendix B. Proof of Theorem 1

By (2.4), we know that $\Sigma = \mathrm{cov}(X_{ki}) = E\big\{(X_{ki}-\mu_k)(X_{ki}-\mu_k)^\top\big\} = E\big\{(BZ_{ki}+\varepsilon_{ki})(BZ_{ki}+\varepsilon_{ki})^\top\big\} = BB^\top + D$. Since D is a diagonal matrix with positive entries, by Lemma 1 (2), for any $1 \le l \le d$, we firstly have $\lambda_l(\Sigma) \le \lambda_l(BB^\top)+\lambda_1(D)$. On the other hand, by Lemma 1 (1), we have $\lambda_l(\Sigma) \ge \lambda_l(BB^\top)$. The above results, together with the fact that $\lambda_1(D) < \infty$ and assumption (2.5), we know that $\lambda_l(\Sigma)/p = \lambda_l(\Sigma_B)+o(1)$, which completes the first part of Theorem 1. Since $B \in \mathbb{R}^{p\times d}$, then for any $l > d$, we have $\lambda_l(BB^\top) = 0$, the above inequality implies that $\lambda_l(\Sigma) \le \lambda_1(D) < \infty$ for $l > d$. Thus, Theorem 1 (2) also holds.

Appendix C. Proof of Theorem 2

Under the null hypothesis of (2.1) and the latent factor model (2.4), we have

$$
\begin{aligned}
&\|\bar{X}_1 - \bar{X}_2\|^2 = \|\mu_1 - \mu_2 + B(\bar{Z}_1 - \bar{Z}_2) + \bar{\varepsilon}_1 - \bar{\varepsilon}_2\|^2 \\
=\ & (\bar{Z}_1 - \bar{Z}_2)^\top B^\top B(\bar{Z}_1 - \bar{Z}_2) + 2(\bar{\varepsilon}_1 - \bar{\varepsilon}_2)^\top B(\bar{Z}_1 - \bar{Z}_2) + (\bar{\varepsilon}_1 - \bar{\varepsilon}_2)^\top(\bar{\varepsilon}_1 - \bar{\varepsilon}_2) \\
\doteq\ & I_1 + 2I_2 + I_3,
\end{aligned}
$$

where $I_1 = (\bar{Z}_1 - \bar{Z}_2)^\top B^\top B(\bar{Z}_1 - \bar{Z}_2)$, $I_2 = (\bar{\varepsilon}_1 - \bar{\varepsilon}_2)^\top B(\bar{Z}_1 - \bar{Z}_2)$ and $I_3 = (\bar{\varepsilon}_1 - \bar{\varepsilon}_2)^\top(\bar{\varepsilon}_1 - \bar{\varepsilon}_2)$. Here, we define $\bar{Z}_k = n_k^{-1}\sum_i Z_{ki}$ and $\bar{\varepsilon}_k = n_k^{-1}\sum_i \varepsilon_{ki}$, for $k = 1, 2$. We next consider the three terms I_1, I_2 and I_3 separately by the following three steps accordingly.

STEP I. We firstly consider I_1. By the central limit theorem, $\sqrt{n_1}\bar{Z}_1 \to_d N(0, I_d)$ and $\sqrt{n_2}\bar{Z}_2 \to_d N(0, I_d)$ hold as long as $\min(n_1, n_2) \to \infty$. Denote $Y = \sqrt{n_1 n_2}(\bar{Z}_1 -$

$\bar{Z}_2)/\sqrt{n_1+n_2}$, then $Y \rightarrow_d N(0, I_d)$. Consequently, by assumption (2.5), $p^{-1}n_1n_2/(n_1+n_2)I_1 = \text{tr}(p^{-1}Y^\top B^\top BY) = \text{tr}\Big[\big\{\Sigma_B + o(1)\big\}YY^\top\Big] = \text{tr}(W^\top \Lambda_B W)\big\{1+o(1)\big\}$, where $W = P^\top Y \rightarrow_d N(0, I_d)$ and $\Sigma_B = P\Lambda_B P^\top$ is the resulting spectral decomposition. Therefore, $p^{-1}n_1n_2/(n_1+n_2)I_1 = \sum_{l=1}^d \lambda_l(\Sigma_B)\chi_l^2(1) + o_p(1)$.

Step II. We next prove that I_2 is negligible. By the model assumptions, $(\bar{Z}_1, \bar{Z}_2)$ is independent of $(\bar{\varepsilon}_1, \bar{\varepsilon}_2)$. Then $E(I_2) = 0$ and

$$\begin{aligned}
&\text{var}(I_2) = \text{E}(I_2^2) = \text{E}\Big\{(\bar{\varepsilon}_1 - \bar{\varepsilon}_2)^\top B(\bar{Z}_1 - \bar{Z}_2)(\bar{Z}_1 - \bar{Z}_2)^\top B^\top(\bar{\varepsilon}_1 - \bar{\varepsilon}_2)\Big\} \\
&= \text{tr}\Big[\text{E}\big\{B(\bar{Z}_1 - \bar{Z}_2)(\bar{Z}_1 - \bar{Z}_2)^\top B^\top(\bar{\varepsilon}_1 - \bar{\varepsilon}_2)(\bar{\varepsilon}_1 - \bar{\varepsilon}_2)^\top\big\}\Big] \\
&= \text{tr}\Big[\text{E}\big\{B(\bar{Z}_1 - \bar{Z}_2)(\bar{Z}_1 - \bar{Z}_2)^\top B^\top\big\}\text{E}\big\{(\bar{\varepsilon}_1 - \bar{\varepsilon}_2)(\bar{\varepsilon}_1 - \bar{\varepsilon}_2)^\top\big\}\Big] \\
&= \text{tr}\big[(n_1+n_2)/(n_1n_2)BB^\top(n_1^{-1}+n_2^{-1})D\big] \\
&= \text{tr}\big[(n_1+n_2)^2/(n_1n_2)^2BB^\top D\big].
\end{aligned}$$

As a result, by assumption (2.5) and the fact that $\lambda_1(D) < \infty$, there exists some finite positive constant C_1 such that

$$\begin{aligned}
p^{-1}\text{var}(I_2) = (n_1+n_2)^2/(n_1n_2)^2\text{tr}(p^{-1}B^\top DB) \leq (n_1+n_2)^2/(n_1n_2)^2\lambda_1(D)tr(p^{-1}B^\top B) \\
\leq C_1\text{tr}(\Sigma_B)(n_1+n_2)^2/(n_1n_2)^2.
\end{aligned}$$

This suggests that $I_2 = O_p\{p^{-1/2}(n_1+n_2)/(n_1n_2)\}$. Consequently, $n_1n_2p^{-1}(n_1+n_2)^{-1}I_2 = O_p(p^{-1/2})$, so that I_2 is negligible.

Step III. We lastly prove that I_3 is negligible. By definition, we have $\text{E}\big\{n_1n_2I_3/(n_1+n_2) - tr(D)\big\} = 0$. We next evaluate $\text{var}(I_3)$. By Lemma 2, there exists some finite positive constant C_2 such that

$$\begin{aligned}&\mathrm{var}(I_3) = \mathrm{E}\Big\{(\bar{\varepsilon}_1 - \bar{\varepsilon}_2)^\top(\bar{\varepsilon}_1 - \bar{\varepsilon}_2)(\bar{\varepsilon}_1 - \bar{\varepsilon}_2)^\top(\bar{\varepsilon}_1 - \bar{\varepsilon}_2)\Big\} - \big\{\mathrm{E}(I_3)\big\}^2 \\ &\leq C_2 \mathrm{tr}\big\{(n_1^{-1} + n_2^{-1})^2 D^2\big\} = O(n^{-2}p),\end{aligned}$$

where the last equality is due to the fact that $\lambda_1(D) < \infty$. As a result, we have $\mathrm{var}(p^{-1}n_1n_2I_3/(n_1+n_2)) = O(p^{-1})$, which further implies $p^{-1}n_1n_2I_3/(n_1+n_2) - p^{-1}tr(D) = o_p(1)$. Thus, I_3 is also negligible. Combines the results above, we have completed the entire proof.

Appendix D. Proof of Theorem 3

The proof of (1) can be directly obtained by Wang (2012). As a result, we only need to prove (2). Note that $\Sigma = BB^\top + D$. Then, $tr(D)/p = tr(\Sigma)/p - tr(BB^\top)/p = tr(\Sigma)/p - tr(\Sigma_B) + o(1)$, where the last equality is implied by condition (2.5). Thus, the desired conclusion can be proved in two steps. In the first step, we show that $tr(\hat{\Sigma})/p - tr(\Sigma)/p \to_p 0$. In the second step, we prove $tr(\Sigma_B) - \sum_{l=1}^d \lambda_l(\hat{\Sigma})/p \to_p 0$.

Step I. Note that $X_{ki} - \bar{X}_k = B(Z_{ki} - \bar{Z}_k) + \varepsilon_{ki} - \bar{\varepsilon}_k$. Consequently, $\hat{\Sigma}$ can be decomposed into three parts, i.e. $\hat{\Sigma} \doteq \Delta_1 + \Delta_2 + \Delta_3$.

$$\begin{aligned}\Delta_1 &= \sum_{k=1}^2\sum_{i=1}^{n_k} B(Z_{ki} - \bar{Z}_k)(Z_{ki} - \bar{Z}_k)^\top B^\top/(n_1+n_2-2), \\ \Delta_2 &= \sum_{k=1}^2\sum_{i=1}^{n_k}(\varepsilon_{ki} - \bar{\varepsilon}_k)(\varepsilon_{ki} - \bar{\varepsilon}_k)^\top/(n_1+n_2-2), \\ \Delta_3 &= 2\sum_{k=1}^2\sum_{i=1}^{n_k} B(Z_{ki} - \bar{Z}_k)(\varepsilon_{ki} - \bar{\varepsilon}_k)^\top/(n_1+n_2-2).\end{aligned}$$

We next consider the above three terms separately. We firstly consider Δ_1. By the identification condition $\mathrm{cov}(Z_{ki}) = I_d$, we have $n_k^{-1}\sum_{i=1}^{n_k}(Z_{ki} - \bar{Z}_k)(Z_{ki} - \bar{Z}_k)^\top = I_d\{1+$

$o_p(1)\}$ for $k = 1, 2$. As a result,

$$\Delta_1 = B\Big\{\sum_{k=1}^{2}\sum_{i=1}^{n_k}(Z_{ki} - \bar{Z}_k)(Z_{ki} - \bar{Z}_k)^\top/(n_1 + n_2 - 2)\Big\}B^\top = BB^\top\{1 + o_p(1)\},$$

which leads to $tr(p^{-1}\Delta_1) = tr(p^{-1}B^\top B)\{1 + o_p(1)\} \to_p tr(\Sigma_B)$. We next consider Δ_2. One can easily verify that $E(\Delta_2) = D$, we next evaluate $\text{var}\big\{p^{-1}tr(\Delta_2)\big\}$. Define $\widetilde{\varepsilon}_{ki} = D^{-1/2}(\varepsilon_{ki} - \bar{\varepsilon}_k)$. Consequently, by Lemma 2, there exists some finite positive constant C_3 such that

$$p^{-2}\text{var}\big\{tr(\Delta_2)\big\} = p^{-2}(n_1 + n_2 - 2)^{-2}\sum_{k=1}^{2}\sum_{i=1}^{n_k}\text{var}\Big\{\widetilde{\varepsilon}_{ki}^\top D\widetilde{\varepsilon}_{ki}\Big\}$$

$$\leq C_3 p^{-2}(n_1 + n_2 - 2)^{-2}(n_1 + n_2)tr(D^2) = o(1),$$

which leads to $p^{-1}tr(\Delta_2) - p^{-1}tr(D) \to_p 0$. Lastly, using the similar techniques, one can verify that $\text{var}\{tr(\Delta_3)\} = O(p) = o(p^2)$, which leads to $p^{-1}tr(\Delta_3) = o_p(1)$. Combing the results above, we have completed the first part of the proof.

Step II. We next and lastly show that $tr(\Sigma_B) - \sum_{l=1}^{d}\lambda_l(\hat{\Sigma})/p \to_p 0$. It suffices to prove that $\lambda_l(\hat{\Sigma})/p \to_p \lambda_l(\Sigma_B)$ for every $1 \leq l \leq d$. By Lemma 1 (1), we have $\lambda_l(p^{-1}\Sigma) \geq \lambda_l(p^{-1}BB^\top) = \lambda_l(p^{-1}B^\top B)$ and $\lambda_l(p^{-1}B^\top B) \to_p \lambda_l(\Sigma_B)$, which leads to $\lambda_l(p^{-1}\Sigma) \geq \lambda_l(\Sigma_B)$ with probability approaching to 1. Moreover, by Lemma A.5 in Ahn and Horenstein (2013), we have $\lambda_l(p^{-1}\Sigma) \leq \lambda_l(p^{-1}B^\top B) + \lambda_1(D)/p$ and $\lambda_l(p^{-1}B^\top B) + \lambda_1(D)/p \to_p \lambda_l(\Sigma_B)$. Besides, by Lemma A.11 of Ahn and Horenstein (2013), we know that $p^{-1}\lambda_l(\hat{\Sigma}) - p^{-1}\lambda_l(\Sigma) \to_p 0$ for every $1 \leq l \leq d$. Combining the results above, we have $\lambda_l(\hat{\Sigma})/p \to_p \lambda_l(\Sigma_B)$ for every $1 \leq l \leq d$, which completes the entire proof.

Appendix E.Proof of Theorem 4

First, one can verify that $\hat{q}_\alpha$ is a consistent estimator for q_α. As a result, theorem conclusion follows if we can show that $T_{\text{FAST}} \to_p \infty$. By the definition, we have

$$T_{\text{FAST}} = p^{-1}\|\bar{X}_1 - \bar{X}_2\|^2(n_1n_2)/(n_1+n_2) - p^{-1}\Big\{tr(\widehat{\Sigma}) - \sum_{j=1}^{\hat{d}}\lambda_j(\widehat{\Sigma})\Big\}, \tag{A.1}$$

where $\|\bar{X}_1 - \bar{X}_2\|$ can also be expressed as $\|(\bar{X}_1 - \mu_1) - (\bar{X}_2 - \mu_2) + (\mu_1 - \mu_2)\|^2 = \|(\bar{X}_1 - \mu_1) - (\bar{X}_2 - \mu_2)\|^2 + \|\mu_1 - \mu_2\|^2 + 2(\mu_1 - \mu_2)^\top\{(\bar{X}_1 - \mu_1) - (\bar{X}_2 - \mu_2)\}$. Define $n_{12} = (n_1n_2)/(n_1+n_2)$, $\omega_1 = p^{-1}\|(\bar{X}_1-\mu_1)-(\bar{X}_2-\mu_2)\|^2 n_{12} - p^{-1}\{tr(\widehat{\Sigma}) - \sum_{j=1}^{\hat{d}}\lambda_j(\widehat{\Sigma})\}$, and $Q = (\mu_1-\mu_2)^\top\big\{(\bar{X}_1-\mu_1)-(\bar{X}_2-\mu_2)\big\}$. Recall $\delta = p^{-1}\|\mu_1-\mu_2\|^2 n_{12}$. We can then re-write T_{FAST} in (A.1) as $T_{\text{FAST}} = \omega_1 + \delta + 2Q$. Then, by Theorem 2, one can verify that ω_1 is asymptotically distributed as a weighted chi-square distribution of finite degrees of freedom, that is $\sum_{l=1}^{d}\lambda_l(\Sigma_B)\cdot\chi_l^2(1) = O_p(1)$. Thus, it suffices to show $\delta + 2Q \to_p \infty$. To this end, we verify that $E(Q) = 0$ and $\text{var}(Q) = n_{12}^{-1}(\mu_1 - \mu_2)^\top\Sigma(\mu_1 - \mu_2) \leq n_{12}^{-1}\lambda_1(\Sigma)\|\mu_1 - \mu_2\|^2$. By Theorem 1(1), we know that $\lambda_1(\Sigma) = O(p)$. We then have $Q = \|\mu_1 - \mu_2\|O_p(\sqrt{p/n_{12}})$. This immediately suggests that $Q = O_p(\delta^{1/2})$. By theorem assumption, we know that $\delta \to \infty$. We thus have $\delta + 2Q = \delta\{1 + O_p(\delta^{-1/2})\} \to_p \infty$. This completes the entire proof.

REFERENCES

Ahn, S. C. and Horenstein, A. (2013). "Eigenvalue Ratio Test for the Number of Factors," *Econometrica*, 81, 1203–1227.

Anderson, T. W. (2003), *An Introduction to Multivariate Statistical Analysis*, Wiley, Hoboken, NJ.

Bai, Z. D. and Saranadasa, H. (1996), "Effect of high dimension: by an example of a two sample problem," *Statistica Sinica*, 6, 311–329.

Bendat, J. S. and Piersol, A. G. (1966), *Measurement and Analysis of Random Data*, John Wiley: New York.

Cai, T. T., Liu, W. and Xia, Y. (2014), "Two-Sample Covariance Matrix Testing and Support Recovery in High-Dimensional and Sparse Settings," *Journal of the American Statistical Association*, 108, 265–277.

Cai, T. T., Liu, W. and Xia, Y. (2014), "Two-sample test of high dimensional means under dependence," *Journal of the Royal Statistical Society, Series B* , 76, 349-372.

Chen, S. X. and Qin, Y. L.(2010), "A two sample test for high dimensional data with applications to gene-set testing," *The Annals of Statistics*, 38, 808–835.

Fama, E. F. and French, K. R. (1993), "Common risk factors in the return on stocks and bonds," *Journal of Financial Economics*, 33, 3–56.

Fan, J., Fan, Y. and Lv, J. (2008), "High dimensional covariance matrix estimation using a factor model," *Journal of Econometrics*, 147, 186–197.

Fan, J., Liao, Y. and Mincheva, M. (2011), "High Dimensional Covariance Matrix Estimation in Approximate Factor Models," *Annals of Statistics*, 39, 3320-3356.

Keim, D.B. and Ziemba, W. T. (2000), *Security Market Imperfections in World Wide Equity Markets*, Cambridge Universtiy Press.

Lam, C. and Yao, Q. W. (2012), "Factor modeling for high-dimensional time series: inference for the number of factors," *The Annals of Statistics*, 40, 694-726.

Lan, W., Wang, H. and Chih-Ling Tsai. (2014), "Testing covariates in high dimensional regression," *Annals of the Institute of Statistical Mathematics*, 66, 279-301.

Li, J. and Chen, S. X. (2012), “Two sample tests for high dimensional covariance matrices,” *The Annals of Statistics*, 40, 908-940.

G. A. Milliken. and D. E. Johnson (1993), *Analysis of Messy Data: Designed Experiments, Vol 1*, CRC Press.

Maiwald, D. and Kraus, D. (2000), “Calculation of moments of complex Wishart and complex inverse Wishart distributed matrices,” *IEE Proc-Radav Sonar Nuvig*, 147, 162–168.

Ravishanker, N. and Dey, D. K. (2001), *A First Course in Linear Model Theory,* Chapman & Hall/CRC.

Sharpe, W. F. (1964), “Capital asset prices: A theory of market equilibrium under conditions of risk,” *Journal of Finance*, 19, 425–442.

Spirates, P., Glymour, C. and Scheines, R. (2000), *Causation, Prediction, and Search*, Cambridge: MIT Press.

Srivastava, M. S., Katayama, S. and Kano, Y. (2013), “ A two sample test in high dimensional data,” *Journal of Multivariate Analysis*, 114, 349–358.

Wang, H. (2012), “Factor profiled sure independence screening,” *Biometrika*, 99, 15–28.

Zhong, P. S. and Chen, S. X. (2011), “Tests for high dimensional regression coefficients with factorial designs,” *Journal of the American Statistical Association*, 106, 260–274.

Table 1: Empirical power of FAST based on 1,000 simulation replications.

		$n = 100$		$n = 200$	
p	% of true alternative	$d = 1$	$d = 3$	$d = 1$	$d = 3$
200	5%	0.098	0.065	0.171	0.112
	10%	0.167	0.109	0.545	0.244
	25%	0.838	0.315	1.000	0.932
	50%	1.000	0.937	1.000	1.000
500	5%	0.079	0.061	0.145	0.116
	10%	0.146	0.108	0.568	0.236
	25%	0.887	0.299	1.000	0.970
	50%	1.000	0.973	1.000	1.000
1000	5%	0.094	0.072	0.164	0.111
	10%	0.136	0.122	0.558	0.229
	25%	0.918	0.329	1.000	0.985
	50%	1.000	0.988	1.000	1.000