Determining the Rank of the Beta Matrix in a Linear Asset Pricing Model*

Seung C. Ahn[a]
Arizona State University and Sogang University

Alex R. Horenstein[b]
ITAM School of Business

Na Wang[c]
Hofstra University

This Version: September 2011

Abstract

This paper proposes a new estimation method for the rank of the beta matrix in a linear factor model. We consider the case in which possible factor variables, which we call factor-candidate variables, are observed. For a factor model, estimating the rank of the beta matrix is equivalent to estimating the number of the relevant but unobservable factors that are correlated with factor-candidate variables. The estimator we propose is easy to use because it is computed with the eigenvalues of the inner product of an estimated beta matrix. Simulation results show that the proposed method works well even in small samples. Our analysis of US individual stock returns is consistent with the notion that the three factors of Fama and French (1993) capture three different risk sources. We also find that thirteen candidate variables proposed by previous studies including Fama and French are correlated with six common risk factors in US individual stock returns. In addition, our analysis of portfolio returns reveals that the estimated number of common factors changes depending on how the portfolios are constructed. The number of risk sources found from the analysis of portfolio returns is generally smaller than the number found in individual stock returns.

Key Words: factor models, beta matrix, rank, eigenvalues.

JEL Classification: C01, C23, C31, G12

---

*Formerly called “Determining the Rank of the Beta Matrix in a Factor Model with Factor-Candidate Regressors”. We are grateful to Richard Roll, Todd Pronno, seminar participants at the 2010 MFA annual meeting, the 2010 EFA annual meeting, the 2010 FMA annual meeting, the 2010 Korean Econometric Society meeting, the Econometric Society 2009 Far Eastern and South Asia meeting, the 2009 Japan Association of Applied Economics meeting, Columbia University, Instituto Tecnológico Autónomo de México, Universidad Autónoma de Nuevo León, Sogang University and Arizona State University.

[a] Department of Economics, W.P. Carey School of Business, Arizona State University, Tempe, AZ, 85287. Email: miniahn@asu.edu.

[b] ITAM School of Business, Instituto Tecnológico Autónomo de México, México, D.F. 01080, México. Email: alex.horenstein@itam.mx. Alex Horenstein thanks the Asociación Mexicana de Cultura A.C. for its support.

[c] Department of Finance, Frank G. Zarb School of Business, Hofstra University, Hempstead, NY, 11549. Email: na.wang@hofstra.edu.

1. Introduction

Jack Treynor (1962), William Sharpe (1964), John Lintner (1965) and Jan Mossin (1966) developed the Capital Asset Pricing Model (CAPM). The model laid out the foundations of modern asset pricing theory. Since the advent of the CAPM, it has become an important question whether a small number of economic or financial variables can capture the sources of non-diversifiable risk. If the answer is affirmative, then the variables should be priced and the information contained in them is crucial for the agents' portfolio strategies.

Determining whether a factor is priced or not became more important with the development of multifactor asset pricing models, like Merton's Intertemporal CAPM (1972) and the Arbitrage Price Theory (APT) of Ross (1976). These multifactor models tell us that if there exist multiple (r) factors determining non-diversifiable sources of risks, then the factors should properly price the risky assets. However, these models do not tell us what the factors are.

In the empirical asset pricing literature many time-series variables have been proposed as possible risk factors (see Chapter 6 of Campbell, Lo and MacKinlay (1997), Chen, Roll, and Ross (1986), and Fama and French (1992)), which we call factor-candidate variables. Several important questions arise with respect to these factor candidates. Which ones should be included in the pricing equation? Are they capturing different risk sources? By estimating the rank of the beta matrix, we can answer these questions. Adding an additional factor-candidate variable into a model which does not explain asset returns at all is equivalent to adding a column of zeros to the corresponding beta matrix. Thus, the rank of the beta matrix remains the same. Adding an additional variable correlated only with the same risks that the other candidate variables are correlated with amounts to adding a column of betas that is spanned by other columns of betas. Again, the rank remains the same. The rank increases only if the additional variable has explanatory power for individual asset results *and* is correlated with the true factors uncorrelated with other factors. Thus, estimating the rank of the beta matrix is equivalent to estimating the number of true latent factors (true risk sources) correlated with the factor candidate variables used.

Estimating the rank of beta matrix is also a necessary condition for the two-pass (TP) risk premium estimation. The two-pass estimation developed by Fama and MacBeth (1973) has been widely used to estimate the risk premium of each factor-candidate variable. Using this method, the betas of candidate variables are first estimated using asset-by-asset time-series regressions, and then the risk premiums related to the variables are estimated by the cross sectional regression of the mean asset returns on the estimated betas. Whether a factor-candidate variable is priced or not is determined by the significance of the estimated risk premium.

An important condition for the consistency of the TP estimator is that the matrix of the true beta values has full columns. However, there are two cases in which the beta matrix may fail to have full columns. The first case is the true betas related to a factor are all zeros. Kan and Zhang (1999) name such a factor “useless” factor. For a one-factor model in which the factor is useless, Kan and Zhang (1999) have investigated the asymptotic properties of the TP estimator. The useless factor cannot be priced; that is, the premium of the useless factor should be undefined. However, Kan and Zhang show that the TP estimator of an undefined risk premium is asymptotically significant. This happens because the estimated betas are not zeros although the true betas are. The second case is when relevant factors are not the factor-candidate variables themselves, but rather a few linear combinations of them. For such cases, the true beta matrix is not full column, but the estimated matrix may appear to be of full column. Accordingly, some TP premium estimates could falsely appear to be statistically significant, although the corresponding premiums are in fact undefined. Thus, when using the two-pass estimation method researchers need to check the rank of the beta matrix before continuing the second pass cross sectional regression.

This paper proposes a new estimation method for the rank of the beta matrix in an approximate factor model, which we call "threshold" method. We allowed the idiosyncratic components of individual asset returns to be both serially and cross-sectionally correlated. Specifically, we estimate the rank using the eigenvalues of the inner product of the estimated beta matrix. Our estimator equals the number of the eigenvalues greater than a specified threshold value. This threshold method produces consistent estimation as the time series dimension T goes to infinity, as long as the number of cross sectional units (N) is greater than or equal to the number of factor candidates used.

A few papers in the literature have developed estimation methods for the rank of a matrix. Zhou (1995) proposes a Wald test in samples with small N to test the hypothesis of a given rank. Cragg and Donald (1997) provide the tests for the rank of a matrix based on a minimum chi-squared criterion. Robin and Smith (2000) consider the tests based on certain estimated characteristic roots, and show that the limiting distributions of the test statistics are a weighted sum of independent chi-square variables. Kleibergen and Paap (2006) propose a rank statistic using a consistent estimator of the unrestricted matrix, and the proposed rank statistic has a standard χ^2 limiting distribution. However, all these methods are applicable only to data with small *N*. When *N* is large, too many parameters need to be estimated. This is very restrictive for asset prcing applications in which the number of cross-sectional observations, *N*, is usually large.

A method closely related to our method is proposed by Connor and Korajzcyk (1993). Their method is designed to be appropriate for the analysis of the data with large *N* and relatively small *T* observations. Autocorrelation is not allowed for the idiosyncratic components of stock returns. For such data, the number of relevant factors is estimated by evaluating whether adding one more factor results in a significant decrease in the sum of the squares of estimated error terms. To use this sequential method, one needs to determine the order of the factor variables to be tested in an arbitrary matter. In contrast, the threshold method we propose requires weaker restrictions in data. In addition, no ordering of the factors is necessary.

Estimating the rank of the beta matrix is also related to estimating the number of factors. They are related in the sense that the number of the common factors in return data equals to the rank of the beta matrix corresponding to the factors. Bai and Ng (2002), Onatski (2010), and Ahn and Horenstein (2011) have developed formal statistical procedures to estimate the number of the true factors in approximate factor models. Our approach is different from their approaches in one important aspect. Our threshold method is for the case in which the factor-candidate variables are available, while their methods are designed for the cases in which factor-candidate variables are not observed. Our interest is not to estimate the number of all common factors in asset return data, but to estimate the number of relevant factors contained in observed factor-candidate variables. For this purpose, we estimate the number of relevant factors using the estimated betas corresponding to the candidate variables.

The threshold estimator we propose possesses several good properties. First, its consistency does not require any particular restriction on the relation between N and T. Its consistency only requires data with large *T*. Second, the threshold estimator allows idiosyncratic error terms to have serial and cross-sectional dependence. Third, it has power to detect weak factors which have only limited explanatory power. Fourth, it can be applied to the zero factor case. Finally, our simulation exercises indicate that the threshold estimator has good finite sample properties.

Application of the threshold estimation is conducted first on the US individual stock returns. We confirm that the Fama-French (1993) three factors capture three different risk sources. We also find that the five factors of Chen, Roll, and Ross (1986) captures only one risk factor not captured by the Fama-French three factors. Usual momentum and reversal factors and two factors proposed by Chen, Novy-Marx, and Zhang (2010, CNZ) capture two additional risk factors. Overall, our results are consistent with the notion that thirteen factor candidates proposed by previous studies are correlated with six latent factors in US individual stock returns. When we use industrial portfolio returns, results remain the same. However, when we use portfolios that are better diversified such as the ones sorted on characteristics like Size and Book to Market, the Fama-French factors seem to be enough to capture all the common sources of risks, except for the 100 Size and Book to Market portfolios in which an extra factor appears when adding the CNZ factors. Overall, our analysis of portfolio returns reveals that the estimated number of common factors changes depending on how the portfolios are constructed. The rank of the beta matrix found from the analysis of portfolio returns is generally smaller than the one found in individual stock returns, except for the industry portfolios. This result suggest that some industry specific factors disappear when well diversified portfolios are used.

The rank estimation proposed in the paper has two implications for the asset pricing literature. First, it emphasizes the over-identification problem that may arises when all the available factors are used to explain asset returns. The rank estimation produces the number of independent sources of commovement that we should include from all the factor candidates when searching for priced risk premiums. The estimator works very well even when some important factors are not included in the set of factor candidates since we allow for a factor structure in the residuals. Another implication is that the rank estimation method is free of the debate whether or not firm characteristics are priced risk factors. Since we use the double demeaned data set, we exclude the effect of firm characteristics. If priced, the risk sources captured by estimating the rank of the beta matrix can only be systematic risk.

The rest of this paper is presented as follows. Section 2 introduces the factor model we investigate and the assumptions imposed on it. Section 3 derives the asymptotic properties of the threshold estimator. Simulation results are reported in section 4. Section 5 reports our estimation results applied to US individual stock returns and portfolio returns. Concluding remarks follow in section 6. All of the proofs are given in the appendix.

2. Model and Assumptions

We begin by defining an approximate factor model as the one considered by Chamberlain and Rothschild (1983) and Bai and Ng (2002). Let x_{it} be the response variable for the i^{th} cross-section unit at time t, where $i = 1, 2, \cdots, N$, and $t = 1, 2, \cdots, T$. Explicitly, x_{it} can be the (excess) return on asset i at time t. The response variables x_{it} depend on the individual effect α_i, the time effect δ_t and the k factor-candidate variables in $f_t = (f_{1t}, f_{2t}, ..., f_{kt})'$. That is,

$$x_{it} = \alpha_i + \delta_t + \beta_i' f_t + \varepsilon_{it}, \tag{1}$$

where $\beta_i = (\beta_{i1}, \beta_{i2}, \cdots, \beta_{ik})'$ is the beta vector for cross section unit *i*. The product $\beta_i' f_t$ is the common component of x_{it}, and the ε_{it} are idiosyncratic components or idiosyncratic risks.[1]

Our interest for model (1) is to estimate rank of the beta matrix B, where $\mathrm{B} = (\beta_1, \beta_2, ..., \beta_N)'$. However, because of the presence of the time effects δ_t, we are unable to estimate β_i. Instead we can estimate the demeaned betas, $\dot{\beta}_i = \beta_i - \overline{\beta}$, where $\overline{\beta} = N^{-1}\Sigma_{i=1}^{N}\beta_i$. Use of the demeaned beta estimated instead of the raw beta estimates does not cause any technical problem. As long as any β_{ij} is varying over different cross-section units, $rank(\mathrm{B}) = rank(\mathrm{B}^d)$, where $\mathrm{B}^d = (\dot{\beta}_1, \dot{\beta}_2, ..., \dot{\beta}_N)'$. In addition, the rank of B^d matters more than the rank of B for the two-pass regression, because the risk premiums corresponding to the factors in f_t are estimated by the cross-section regression of the individual mean of x_{it} ($\overline{x}_i = T^{-1}\Sigma_{t=1}^{T} x_{it}$) on one and β_i. If any beta in β_i is constant over *i*, the risk premiums are undefined. The premiums are identified only if the demeaned beta matrix B^d has full column.

The demeaned betas can be estimated by estimating the following double demeaned model,

$$\ddot{x}_{it} = \dot{\beta}_i' \dot{f}_t + \ddot{\varepsilon}_{it}, \tag{2}$$

where $\ddot{x}_{it} = x_{it} - \overline{x}_t - \overline{x}_i + \overline{x}$, $\dot{f}_t = f_t - \overline{f}$, $\ddot{\varepsilon}_{it} = \varepsilon_{it} - \overline{\varepsilon}_t - \overline{\varepsilon}_i + \overline{\varepsilon}$, $\overline{x}_t = N^{-1}\Sigma_{i=1}^{N} x_{it}$, $\overline{x}_i = T^{-1}\Sigma_{t=1}^{T} x_{it}$, $\overline{x} = (NT)^{-1}\Sigma_{i=1}^{N}\Sigma_{t=1}^{T} x_{it}$, $\overline{f} = (\Sigma_{t=1}^{T} f_t)/T$, and $\overline{\varepsilon}_t$, $\overline{\varepsilon}_i$, and $\overline{\varepsilon}$ are similarly defined. For each time period t, model (2) can be written as

$$\underset{(T\times 1)}{\ddot{x}_{i\bullet}} = \underset{(T\times k)}{\dot{F}} \; \underset{(k\times 1)}{\dot{\beta}_i} + \underset{(T\times 1)}{\ddot{\varepsilon}_{i\bullet}},$$

where $\ddot{x}_{i\bullet} = (\ddot{x}_{i1}, \ddot{x}_{i2}, \cdots, \ddot{x}_{iT})'$, $\ddot{\varepsilon}_{i\bullet}$ is similarly defined, and $\dot{F} = (\dot{f}_1, \dot{f}_2, \cdots, \dot{f}_T)'$. For all data, we have

---

[1] In this model, we consider only the case of time invariant betas. Our method can be easily extended to the case of time-varying betas since the rank estimation is based on the estimated beta matrix.

$$\begin{array}{ccccccc} \ddot{X} & = & \dot{F} & \mathrm{B}^{d\prime} & + & \ddot{\mathrm{E}} \\ (T\times N) & & (T\times k) & (k\times N) & & (T\times N) \end{array},$$

where $\ddot{X} = (\ddot{x}_{1\cdot}, \ddot{x}_{2\cdot}, \cdots, \ddot{x}_{N\cdot})$, , and $\ddot{\mathrm{E}} = (\ddot{\varepsilon}_{1\cdot}, \ddot{\varepsilon}_{2\cdot}, \cdots, \ddot{\varepsilon}_{N\cdot})$. Then, the demeaned beta matrix B^d can be estimated by the OLS estimator $\hat{\mathrm{B}}^d = \ddot{X}'\dot{F}(\dot{F}'\dot{F})^{-1}$.

In what follows, we use $\lambda_j(A)$ to denote the j^{th} largest eigenvalue of a matrix A, and the norm of A is denoted by $\|A\| = [tr(A'A)]^{1/2}$. We define c as a generic positive constant. With this notation, we make the following assumptions:

Assumption A (factors): $\dot{F}'\dot{F}/T = \Sigma_{t=1}^{T}(f_t - \bar{f})(f_t - \bar{f})'/T \to_p \Sigma_f$, and $\bar{f} \to_p \mu_f$, where Σ_f is finite and positive definite matrix and μ_f is a finite vector.

Assumption B (betas): (i) $\|\beta_i\| \le c$ for all $i = 1, 2, \cdots, N$. (ii) $\mathrm{B}^{d\prime}\mathrm{B}^d/N$ is positive semi-definite and $rank(\mathrm{B}^d) = rank(\mathrm{B}^{d\prime}\mathrm{B}^d) = r \le k$ for all $N \ge r$. (iii) If $N \to \infty$, $\mathrm{B}^{d\prime}\mathrm{B}^d/N \to \Sigma_\beta$, where Σ_β is finite.

Assumption C (idiosyncratic errors): $E(\varepsilon_{it}) = 0$ and $E|\varepsilon_{it}|^4 \le c$ for all i and t, and

$$E\left(\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\varepsilon_{it}\right\|^2\right) = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}E(\varepsilon_{it}\varepsilon_{is}) \le c.$$

Assumption D (weak dependence between factors and idiosyncratic errors):

$$E(\frac{1}{NT}\|\mathrm{E}'F\|^2) = E(\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{1}{\sqrt{T}}\sum_{t=1}^{T}f_t\varepsilon_{it}\right\|^2) = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}E(\varepsilon_{it}\varepsilon_{is}f_t'f_s) \le c.$$

The four assumptions are a subset of the assumptions used in Bai and Ng (2002) and Ahn and Horenstein (2011). Assumption A implies that the factors should be stationary. Assumption B(i) ensures that each factor loading does not explode. Assumption B(ii) allows that the rank of B^d to be smaller than the number of the variables in f_t. Assumption B (iii) implies that for the cases where N is large, $\mathrm{B}^{d\prime}\mathrm{B}^d/N$ is asymptotically finite. That is, the explanatory power of each factor increases at the rate of *N*. The estimators we propose below do not require large *N*. Under Assumption B (iii), the estimators are consistent regardless of the size of *N*. Under Assumption B, we treat the betas as fixed constants. We can easily relax this assumption, but at the cost of more notation.

Assumption C allows weak time-series correlations and does not impose any restrictions on the cross-sectional correlation among the error terms ε_{it}. Our asymptotic results obtained below depend not on the covariance among the errors, but on the dependence between the errors and factors. Assumption C implies that $\Sigma_{t=1}^{T}\varepsilon_{it} / \sqrt{T}$ is a bounded random variable for all *i*. This assumption is weaker than Assumption C of Bai and Ng (2002):

$$\frac{1}{NT}\Sigma_{i=1}^{N}\Sigma_{j=1}^{N}\Sigma_{t=1}^{T}\Sigma_{s=1}^{T}\left|E(\varepsilon_{it}\varepsilon_{js})\right| < c\,.$$

Assumption D implies that the random vectors $\Sigma_{t=1}^{T}\varepsilon_{it} f_t / \sqrt{T}$ are bounded. This assumption is required for the consistency of the ordinary least squares (OLS) estimator of B^{d}. Assumption D is essentially the same assumption as Assumption D of Bai and Ng (2002).

Furthermore, Assumption D allows the errors ε_{it} to have a factor structure. To see why, consider a simple case in which the ε_{it} have an one-factor structure: $\varepsilon_{it} = \xi_i g_t$ where $E(g_t) = 0$, $E(g_t f_t) = 0$, $E(\left|g_t\right|^4) < c$, and $T^{-1}\Sigma_{t=1}^{T}\Sigma_{s=1}^{T}E(g_s g_t f_t' f_s) < c$ for all *t*, and $\left|\xi_i\right| < c$ for all *i*. For this case, the random variable $\Sigma_{t=1}^{T} g_t f_t / \sqrt{T}$ is bounded. Thus, we can easily show that Assumption C holds. In addition,

$$\frac{1}{T}\Sigma_{t=1}^{T}\Sigma_{s=1}^{T}E(\varepsilon_{it}\varepsilon_{is} f_t' f_s) = \xi_i^2 \frac{1}{T}\Sigma_{t=1}^{T}\Sigma_{s=1}^{T}E(g_t g_s f_t' f_s) < c^3\,.$$

Thus, Assumption D holds. Given that the ε_{it} can have a factor structure, estimating the rank of B^{d} is not equivalent to estimating the number of all of the common factors in response variables. The rank of B^{d} is the maximum number of the common components in response variables among the factor candidate variables f_t. Hence, the rank estimation method works well even when the factor candidates do not include all the common underlying factors. The missing information is captured in the error terms with a factor structure.

3. Rank Estimation using Eigenvalues

The Threshold estimator we propose below uses the eigenvalues of $\hat{\hat{\mathrm{B}}}^{d\prime}\mathrm{B}^{d} / N$. So, we begin this section by studying the asymptotic properties of the eigenvalues. Below, we use the notation $\hat{\mu}_{NT,j} = \lambda_j(\hat{\hat{\mathrm{B}}}^{d\prime}\mathrm{B}^d / N)$ where *j* indicates that $\hat{\mu}_{NT,j}$ is the *j*th largest eigenvalue of the matrix $\hat{\hat{\mathrm{B}}}^{d\prime}\mathrm{B}^d / N$. The following theorem presents the asymptotic properties of the eigenvalues.

Theorem 1: Under assumption A – D, (i) $p\lim_{T\to\infty} \hat{\mu}_{NT,j} > 0$ for $0 < j \le r$; and (ii) $\tilde{\mu}_{NT,j} = O_p(T^{-1})$, for $0 \le r < j \le k$.

Theorem 1 shows that the first $r > 0$ largest eigenvalues of $\hat{\hat{\mathrm{B}}}^{d\prime}\mathrm{B}^d / N$ have the same convergence rates, which are different from those of the other eigenvalues. This difference in convergence rate is used to identify the rank of the matrix B^d, r. Notice that the asymptotic properties of the eigenvalues do not require $N \to \infty$. Theorem 1 holds for any fixed number *N*. Therefore, the estimator we propose below does not require large *N*.

The following theorem defines the consistent estimator that we call “Threshold” estimator.

Theorem 2 (Threshold Estimator): For a given threshold function $g(T) > 0$ such that $g(T) \to 0$ and $Tg(T) \to \infty$ as $T \to \infty$, define $\hat{\hat{r}}_{TH} = \#\{1 \le j \le k : \mu_{NT,j} > g(T)\}$, where $\#\{\bullet\}$ is the cardinality of a set . Then, under Assumptions A – D, $\lim_{T\to\infty} \Pr(\hat{r}_{TH} = r) = 1$.

The result of Theorem 2 is quite intuitive. Observe that $g(T)$ converges to zero at a lower rate than the last $(k-r)$ eigenvalues of $\hat{\hat{\mathrm{B}}}^{d\prime}\mathrm{B}^d / N$ do. The first r eigenvalues converge to positive numbers. Accordingly, for sufficiently large T, the value of $g(T)$ is most likely to be smaller than the first r eigenvalues and larger than the rest of the eigenvalues. The threshold estimation procedure is similar to the methods suggested by Bai and Ng (2002) to estimate the number of unobservable common factors in an approximate factor model with a large number of response variables.

Note that we can also use the Threshold estimator proposed in Theorem 2 for the cases in which (i) the data is not generated by a factor model and/or (ii) all the factor candidates are useless. We will call this situation “no-factor” case. For such a case, $r = 0$.

A possible pitfall of the threshold estimator is that there are many possible choices for $g(T)$. Whenever a function is an appropriate choice for $g(T)$, so is a finite multiple of the function. If T is large, the estimation results would be insensitive to the choice of $g(T)$. However, for the data with relatively small T, the estimation result could change depending on the choice of $g(T)$. The optimal choice of the threshold function $g(T)$ may depend on the data generating processes. In the following paragraph we propose a specific function for $g(T)$ which provides reliable estimates for many different data generating processes we have considered in our Monte Carlo experiments.

Let $\hat{\sigma}^2 = [(N-1)(T-1)]^{-1}\Sigma_{i=1}^{N}\Sigma_{t=1}^{T}\ddot{e}_{it}^2$, where the $\ddot{e}_{it}$ are the OLS residuals from the regression of the double demeaned model (2). The estimator $\hat{\sigma}^2$ is a consistent estimator of $\operatorname{var}(\varepsilon_{it})$. Also, let $R^2 = 1-[\Sigma_{i=1}^{N}\Sigma_{t=1}^{T}\ddot{e}_{it}^2]/[\Sigma_{i=1}^{N}\Sigma_{t=1}^{T}\ddot{x}_{it}^2]$ be the R-square from the OLS regression of model (2). Then, the threshold function we suggest to use for the Threshold estimator is given by:

$$g(d,T) = d\hat{\sigma}^2 / T^d, \tag{3}$$

where $d = 1-R^2$ for $0.3 \le 1-R^2 \le 0.8$, $d = 0.3$ for $1-R^2 < 0.3$, and $d = 0.8$ for $1-R^2 > 0.8$.

The function $g(d,T)$ is designed to be a non-decreasing function of R^2 for sufficiently large T. Specifically, for $T > 28$, $g(d,T)$ is a monotonically decreasing function of d. Because d is a non-increasing function of R^2, $g(d,T)$ is an increasing (specifically, non-decreasing) function of R^2. The use of $g(d,T)$ is motivated by our findings from Monte Carlo simulations: when the data are generated by weak factors (that have low explanatory power), smaller threshold values are needed to better estimate the rank of B^d. Since $g(d,T)$ should satisfy the two conditions given in Theorem 2, we limit the range of d to be [0.3, 0.8]. The choice of the range is somewhat arbitrary. However, this range is the best choice we have found from simulations. The property of $g(d,T)$ could be stated from Graph 1. When R^2 is low, factors have low explanatory power, we need a small value of $g(d,T)$ to detect the weak factors. When R^2 is high, factors are stronger and a relative larger threshold function is needed. When T increases, the last $(k-r)$ eigenvalues of $\hat{\hat{\mathrm{B}}}^{d\prime}\mathrm{B}^d / N$ converges to zero faster, and a smaller $g(d,T)$ is needed.

4. Simulations

Our simulation data are drawn by the same model used in Bai and Ng (2002) and Ahn and Horenstein (2009):

$$x_{it} = \alpha_i + \delta_t + \Sigma_{j=1}^{k} \beta_{ij} f_{jt} + u_{it}\ ;\ u_{it} = v_{it}\sqrt{(1-\rho^2)/(1+2J\delta)}\ ,$$

where $v_{it} = \rho v_{i,t-1} + \xi_{it} + \Sigma_{h=\max(i-J,1)}^{i-1} \delta v_{ht} + \Sigma_{h=i+1}^{\min(i+J,N)} \delta v_{ht}$, and the ξ_{it} $(1 \le i \le N)$ and the factor candidate variables f_{jt} are randomly drawn from $N(0,1)$. In this setup, the variance of u_{it} is roughly equal to one.

For simplicity, we set $\alpha_i = 0$ for all i, and $\delta_t = 0$ for all t. The beta matrix B is drawn by the following way. We draw a $N \times r$ random matrix A, each entry of which is $N(0,1)$. We also draw a random $k \times k$ positive definite matrix, compute the first r orthonormalized eigenvectors of the matrix, and set a $k \times r$ matrix C using the eigenvectors.[2] Then, we set $\mathrm{B} = A\Lambda^{1/2}C'$, where $\Lambda = diag(\lambda_1, ..., \lambda_r)$. This setup is equivalent to the case in which the true factors are $f_t^* = \Lambda^{1/2}C'f_t$ with $Var(f_t^*) = \Lambda$ and the beta matrix corresponding to $f_t^* = (f_{1t}^*, ..., f_{rt}^*)'$ is A.

The parameter Λ controls the signal to noise ratio of each of the true factors (SNR, ratio of the variances of a factor and the idiosyncratic error, u_{it}). When the j^{th} true factor, f_{tj}^*, has the variance of $\lambda_j = 1/r$, its SNR equals $1/r$, where $r \ge 1$. In case of $r = 0$, we present the table separately. For benchmark simulations, we use $\lambda_j = 1/r$, for $1 \le j \le r$. In other simulations we try different λ_j's.

For the error terms, we consider four cases: (i) the cases with *i.i.d.* errors ($\rho = J = \delta = 0$), (ii) with both cross-sectional and auto-correlated errors ($\delta = 0.2$, $\rho = 0.5$, $J = 8$), (iii) with only cross-sectional correlated errors ($\rho = 0$), and (iv) with only auto-correlated errors ($J = \delta = 0$). For each case, we try 25 different combinations of *N* and *T*, where *N*, *T* $\in$ {50, 100, 200, 500, 1000}. 1,000 samples are drawn for each combination of *N* and *T*.

Tables 1 – 3 report the results from our benchmark simulations ($\lambda_1 = ... = \lambda_r = 1/r$). Table 1 shows the estimation results from the cases with *i.i.d.* idiosyncratic errors and both cross-sectional and auto-correlated errors. Specifically, for the correlated error cases, we set $\rho = 0.5$, $\delta = 0.2$, $J = 8$. All factors have the SNRs of $1/r$, where $r \ge 1$. The Threshold estimator performs very well, even in the case of small sample size (e.g., $T = 50$). For every case, the mean of the rank estimates is almost equal to the true rank. Also, only for a few

---

[2] We first generate a $N \times k$ matrix *M* whose entries are drawn from *N*(0,1), and then compute the r eigenvectors of $M'M$.

cases, the standard deviation of the estimates is larger than zero. The results with correlated errors are not noticeably different from those with *i.i.d.* errors.

Tables 2 shows the results from the cases of cross-sectional correlation only ($\rho = 0$, $\delta = 0.2$, $J = 8$) and auto-correlation only ($\rho = 0.5$, $\delta = 0.2$). The factors are generated with $\lambda_j = 1/r$, for $1 \le j \le r$. For all cases, the Threshold estimator performs very well even if *T* is small.

Table 3 shows the results for the cases in which all factors are weak with the same SNRs. The left part of the table reports the results from the cases with *i.i.d.* errors, while the right part presents the results from the cases with both cross-sectionally and auto-correlated errors. For small *T* (*T* = 60), the Threshold estimator does not perform well when the SNRs of the factors are as low as 0.025. But it works well in the cases with the SNRs larger than 0.05. For the case in which $T = 100$, the Threshold estimator performs very well even in the cases with the SNRs of 0.025. The estimation results from the data simulated with *i.i.d.* errors are more reliable than those from the data with correlated errors, especially when *T* is small and factors are weak. In fact, we can add one more dimension of the SNR to the threshold function. If the weak factors defined as important factors need SNRs at least larger than 1/5, we can adjust the threshold function with the simulated data to make our estimation capturing al the factors with SNRs larger than 1/5.

Table 4 is designed to investigate the performances of the Threshold estimator when both weak and strong factors coexist. As in table 3, the left part of the table reports the results from the cases with *i.i.d.* errors, while the other part presents the results from the cases with cross-sectionally and auto-correlated errors. We conduct the test with two different factor-candidates models, both of them with $k = 3$. In each of these models we study three different possible SNRs for the weak factor. In one model we construct a factor structure with $r = 2$, where the first true factor is strong with λ_1 fixed at one and the second true factor is weak with three different λ_2 values: λ_2 = 0.1, 0.2, and 0.3. In the other model we study the case with $r = 3$ where the first two true factors are strong with $\lambda_1 = \lambda_2 = 1$, the last one is weak with three different λ_3 values: λ_3 = 0.1, 0.2, and 0.3. From the table, we can see that the Threshold estimator performs very well in small samples even if the weak factor's SNR is ten times smaller than the SNRs of the strong ones ($\lambda_2 = 0.1$, in the first model and $\lambda_3 = 0.1$, in the second model). The structure of the error terms does not show significant difference in the results.

Table 5 is designed to investigate the performances of the Threshold estimator for the data generated without true factors. That is, all of the factor candidate factors used for Table 5 are “useless.” We consider the cases with different numbers of useless factors. Table 5 shows that the Threshold estimator correctly detects the cases in which all factors are useless, if the number of factor candidate variables is small (e.g., $k = 1$), or T is large, or errors are only weakly correlated. When the errors are highly correlated, the estimator has relatively low power to detect useless factors unless T is sufficiently large.

Our simulation results can be summarized as follows. First the Threshold estimator provides quite reliable inferences on the rank of the beta matrix even if the sample size is small. The SNR of each factor, the degrees of correlations among the errors, and the number of cross section units do not substantially influence the performances of the estimators. Second, the Threshold estimator can be used to check the possibility of all factor candidates’ being “useless.” The Threshold estimator is relatively less precise, if the number of the factor candidates analyzed is too large, or if the errors are highly correlated. However, it performs reasonably well even under such cases if the number of the time series observations is sufficiently large.

5. Application

In this section we estimate the rank of the beta matrix using different factor-candidates as regressors. More specifically, we use the three factors proposed in the model of Fama and French (1992, FF), the five factors of Chen, Roll, and Ross (1986, CRR),[3] the momentum and reversal factors (MOM) available on Kenneth French webpage: momentum, short-term reversal and long-term reversal, and the two new factors developed in Chen, Novy-Marx, and Zhang (CNZ, 2010): Investment to Asset (IA) and Return on Asset (ROA). [4]

As response variables we use the US monthly individual stock returns and portfolio returns.[5] Returns are calculated in excess over the risk free rate. The individual stock returns

---

[3] While the FF model may be more related to the APT, the CRR model is more related to Merton’s (1972) Intertemporal CAPM, in the sense that they try to find the macroeconomic (state) variables that may influence future investment opportunities. The factors proposed by CRR are industrial production (MP), unexpected inflation (UI), change in expected inflation (DEI), the term premium (UTS), and the default premium (UPR). Each of these factors is available from Laura Xiaolei Liu’s webpage from January 1960 to December 2004 (http://www.bm.ust.hk/~fnliu/research.html). For detailed information on how these factors have been constructed, see Liu and Zhang (2008). The FF factors are the proxy for the market risk premium, SMB and HML.

[4] We thank Long Chen for providing us the latest version of their factors.

[5] We do not use the daily returns since the data of some factor candidates are only available at monthly frequency.

are downloaded from CRSP. The returns include dividends. The risk free rate is the one-month Treasury bill rate, which is available from Kenneth French's webpage. For the individual stock returns, we exclude REITs (Real Estate Investment Trusts), ADRs (American Depositary Receipts) and the stocks that do not have information for every month over a sample period. We also exclude stocks that show more than 300% excess returns in a given month since we are trying to capture common variation. Excessively high or low returns are most likely to be idiosyncratic risks. US Stock portfolio returns are downloaded from Kenneth French's webpage. The portfolios used are 100 portfolios based on Size and Book to Market, 25 portfolios based on Size and Book to Market, 25 portfolios based on Size and Momentum, 49 Industrial portfolios and 30 Industrial portfolios. We use monthly returns in every data set.

Response variables are always double-demeaned as suggested in the Equation (2). We also use standardized factors for the following reason. The beta values corresponding to each factor change depending on the scale of the factor. For example, if we rescale a factor by multiplying 10, the (absolute) beta values corresponding to the factor are scaled down by the order of 0.1. In this case, even if the factor has a high explanatory power, the estimated betas obtained with the rescaled factor would not reflect the factor's true explanatory power.

5.1 Rank of beta matrices using individual US stock returns as response variables

The time span included in the analysis is from 1972 to 2004. We divide the individual stock returns into three samples: the entire time span (1972-2004), two subsamples (1972 – 1987 and 1988 – 2004) and three subsamples (1972 – 1978, 1979 – 1992, and 1993 – 2004). Under both subdivisions, we could fit a polynomial trend to the value weighted market portfolio to estimate the up and down cycles. We do so to examine how the estimation results may change depending on time intervals. We keep the time span T at around 100 or more since the simulation exercises show that the estimators are very accurate in this case. The number of cross-sectional observations N changes as T changes in order to maintain a balanced panel. The value of N depends on the available observations with complete data on CRSP for each sample period after the data has been cleaned.

The results from the estimation of the rank of the beta matrix for individual stock returns are shown in Table 6. Each line of the table represents a different estimated model. For each model we report the number of factor candidates used (*k*), the estimated number of factors among the factor candidates ($\hat{r}$) and the average R^2 of the regressing the response variables on the factor-candidates.

The first line of table 6 shows that the Threshold estimator predicts that the rank of the beta matrix equals three when using the three FF factors in different sample periods.

The second line of table 6 shows the results from the estimation of the five CRR factors. For any period, the estimated rank does not exceed two. This means that only one or two common sources of comovement in individual stock returns are explained by the CRR factors. This result provides strong evidence that the risk premiums of some factors in the CRR model are undefined.

Given that the CRR factors can identify one or two common factors in individual stock returns, a question we wish to answer is whether the CRR factors capture some sources of comovement that the FF factors fail to do. If the CRR factors capture different information from what the FF factors do, we could expect that the rank of the beta matrix from the joint model of CRR and FF would be equal to the sum of the ranks from the CRR and FF models separately. Indeed, the Threshold estimation results are consistent with this expectation in the entire sample and every subsample. In the third line of result in table 6 the Threshold estimation suggests that the risks captured by the CRR and FF factors are different.

Since the five CRR factors capture a common source of comovement that is not captured by the FF factors, an interesting question is which of the CRR factors contain the information missed by the FF factors. For this purpose we add to the FF factors each CRR factor individually in order to estimate the rank of the beta matrix of at most four. In unreported results we find that no individual CRR factor increases the rank of the beta matrix when combined with the FF factors. Then we use every possible combination of two CRR factors together added to the three FF factors. In this case we found that adding UI (unexpected inflation) and DEI (changes in expected inflation) increases the rank of the beta matrix to four. Results are shown in the 4th line of table 6. This shows that a factor related to inflation is missed by the FF factors.

Furthermore, we analyze if momentum factors (as constructed by Kenneth French) capture a different source of risk than the Fama-French factors. Results of estimating the rank of the beta matrix of the three momentum factors and the FF factors are presented in the 5th row of the table. The Threshold estimator finds strong evidence for an extra factor contained in the three momentum factors in most samples. However, if we add any one or any two possible combinations of the momentum factors to FF three factors, unreported results show that in most cases we find the rank equals three. We conclude that there is evidence for a momentum factor among the three momentum factors during the period under analysis that is not captured by the FF factors when using individual stock returns.

In the 6th row of table 6 we test the rank of the beta matrix when using the three FF factors and the two new factors of CNZ and find four factors in almost every subsample.[6] This is evidence that the CNZ factors capture one dimension missed by the FF factors.

Finally, the last row of the table show the results of using the ten factor-candidates that seem to contain different information together: the three FF factors, UI and DEI from CRR, the three momentum factors and the two CNZ factors. The table shows that there is evidence for at least six factors among the 10 factor candidates.

However, an open and important question is whether we need to use individual stock returns or portfolio returns to estimate the beta matrix in order to perform asset pricing tests[7]. For example, imagine a hypothetical situation in which half of the sample of the individual stock returns have betas of 0.5 with respect to a factor and the other half have betas of -0.5. In this case the factor will add a dimension to the rank of the beta matrix when using individual stock returns, but this factor will disappear in properly diversified portfolios (because the beta of the diversified portfolio with respect to the factor will be zero). In the next section we estimate the rank of the beta matrix using the same factor candidates as before but using portfolio returns as response variables.

5.2 Rank of beta matrices using US stock portfolio returns as response variables

In this section we use five sets of portfolios downloaded from Kenneth French website as response variables. Since the number of portfolios is fixed in each different set, we use for every estimation the full time span from January 1972 to December 2004 (T=396). The cross-sectional dimension N equals to the number of portfolios in each set. In table 7 we report the same statistics for portfolio returns as those in the previous table for individual stock returns.

When using the FF factors we find all the time an estimated rank of three except for the 25 Size and Book to Market portfolio set where we find a rank of two. When we use the five CRR factors we find the rank equals to one or two as in the case with individual stocks. When we test together the FF factors and the CRR factors (k=8), we do not find evidence of an extra factor except for the cases of the 49 and 30 Industrial Portfolios.

---

[6] Most of the time adding ROA to the FF factors is sufficient to get a rank equal to four while adding only IA never increases the estimated rank of the beta matrix. For this reason, we can conclude that ROA posses most of the information not captured by the FF factors.

[7] See Ang, Liu, and Schwarz (2008).

A common pattern observed in table 7 is that when testing the number of factors in Industrial Portfolios the results are similar to those obtained using individual stock returns. However, once we use portfolios based on Book to Market and Size or Size and Momentum, the rank of the beta matrix is at most four. The maximum rank we find for 100 Size and Book to Market portfolios is four, and for 25 Size and Book to Market portfolios and 25 Size and Momentum portfolios is three. This is evidence that the portfolios sorted based on these characteristics are better diversified (these portfolios also show less residual variance since their R^2 is higher than the one of the Industrial portfolios). A possible explanation is the existence of industry specific factors that are diversified away when constructing portfolios based on characteristics like Size and Book to Market. This is a useful result that can clarify the discussion of whether to use portfolios or individual stock returns when testing factors and also the discussion about which type of portfolios should be used. It is known that industry portfolios tend to have positive abnormal excess returns (intercepts are significantly larger than zero). According to our result this is because the existence of industry specific factors that disappear when well diversified portfolios are used. In other words, the positive α that appears in many of the Industry Portfolios should not be considered a models' mispricing since it is exposure to a source of diversifiable risk.

Our empirical results can be summarized as follows. When using individual stock returns we find evidence for the existence of six common factors among the thirteen factor candidates used. These factors are the three FF factors, a factor related to inflation from the CRR factors, a Momentum factor and a factor captured by the new CNZ factors. When we use Industrial Portfolio returns, results remain the same. However, when we use portfolios that are better diversified such as the ones sorted on characteristics like Size and Book to Market, the FF factors seem to be enough to capture all the common sources of risk among the thirteen factor candidates, except for the 100 Size and Book to Market portfolios in which an extra factor appears when adding the CNZ factors.

6. Conclusions

In this paper, we have proposed a new rank estimator, called Threshold estimator, for the beta matrix from a factor model with observed factor-candidate variables. Testing whether the beta matrix has full rank is important for the two-pass estimation of the risk premiums in empirical asset pricing models. The (demeaned) beta matrix needs to have full rank. Otherwise, risk premiums are undefined. The Threshold estimator is computed easily with the eigenvalues of the inner product of an estimated beta matrix. Our simulation exercises provide promising evidence that the Threshold estimator has good finite-sample properties. Different from the existing methods, this proposed method can be used to analyze the data with a large number of cross-section units.

In our empirical investigation we find that all of the Fama-French (1993) three factors have explanatory power when using US individual stock returns as response variables, In contrast, only one or two among the five factors of Chen, Roll, and Ross (1986) have explanatory power. When we combine the three factors of Fama-French (FF) together with the five factors of Chen, Roll, and Ross (CRR) we find that a factor not captured by FF is captured by CRR. Furthermore, we find that momentum and reversal factors capture a source of risk not captured by either FF or CRR. Similarly, the two factors proposed by Chen, Novy-Marx, and Zhang (2010, CNZ) capture a source of risk missed by all the other factors. We find evidence for six factors in US individual stock returns among the thirteen factor candidates used. When we use Industrial Portfolio returns, results remain the same. However, when we use portfolios that are better diversified such as the ones sorted on characteristics like Size and Book to Market, the FF factors seem to be enough to capture all the common sources of risk among the thirteen factor candidates, except for the 100 Size and Book to Market portfolios in which an extra factor appears when adding the CNZ factors.

Bai and Ng (2002), Onatski (2006), and Ahn and Horenstein (2009) have developed the estimators for the number of factors without using factor candidate variables. Their studies have found one or two factors from US individual stock return data. In contrast, our results provide evidence of at least six factors in individual stock returns. All of the estimation methods proposed by the above three studies are based on the analysis of principal components of response variables. Ahn and Horenstein (2009) found that principal components provide poor estimates of the true factors when the true factors are weak and the idiosyncratic errors are cross sectional correlated. From their results, we can conjecture that the analysis of principal components might have limited power to detect weak factors. In contrast, the Threshold estimator proposed in this paper utilizes observed factor candidate variables. Factors need not be estimated. Thus, we can expect that the new estimator would have a higher power to detect the weak factors hidden among the factor-candidate variables. Our estimation results are consistent with this expectation.

Appendix

The following two Lemmas are used to prove Theorem 1.

Lemma 1: Under Assumption B and D, for any $k\times p$ ($p\le k$) matrices A and G such that $\|A\| = O_p(1)$, and $\|G\| = O_p(1)$, we have two conclusions:

(i) $$\frac{1}{NT}\left|tr(A'\mathrm{B}^{d\prime}\ddot{\mathrm{E}}'\dot{F}^{d}G)\right| = O_p\left(T^{-1/2}\right);$$

(ii) $$\frac{1}{NT^2}\left|tr(A'\dot{F}'\ddot{\mathrm{E}}\ddot{\mathrm{E}}'\dot{F}A)\right| = O_p(T^{-1}).$$

Proof: Assumption B implies

$$\begin{aligned}\left(\frac{1}{\sqrt{N}}\left\|\mathrm{B}^{d}\right\|\right)^2 &= \frac{1}{N}\left\|\mathrm{B}-1_N1_N'\mathrm{B}/N\right\|^2 \le \frac{1}{N}\left\|\mathrm{B}\right\|^2+\frac{1}{N}\left\|1_N1_N'/N\right\|^2\left\|\mathrm{B}\right\|^2 \\ &= \frac{2}{N}\left\|\mathrm{B}\right\|^2 = \frac{2}{N}\sum_{i=1}^{N}\left\|\beta_i\right\|^2 \le 2c^2\end{aligned};$$

From Assumptions B – D, we obtain

$$\begin{aligned}&\left(\frac{1}{\sqrt{NT}}\left\|\ddot{\mathrm{E}}'\dot{F}\right\|\right)^2 \\ &= \frac{1}{NT}\left\|\ddot{\mathrm{E}}'\dot{F}\right\|^2 = \frac{1}{NT}\left\|\left(\mathrm{E}'-\frac{1}{T}\mathrm{E}'1_T1_T'-\frac{1}{N}1_N1_N'\mathrm{E}'+\frac{1}{NT}1_N1_N'\mathrm{E}'1_T1_T'\right)'\left(F-\frac{1}{T}1_T1_T{}'F\right)\right\|^2 \\ &= \frac{1}{NT}\left\|\mathrm{E}'F-\frac{1}{T}\mathrm{E}'1_T1_T'F-\frac{1}{N}1_N1_N'\mathrm{E}'F+\frac{1}{NT}1_N1_N'\mathrm{E}'1_T1_T'F\right\|^2 \\ &= \frac{1}{NT}\left\|\mathrm{E}'F\right\|^2+\frac{1}{NT}\left\|\frac{1}{T}\mathrm{E}'1_T1_T'F\right\|^2+\frac{1}{NT}\left\|\frac{1}{N}1_N1_N'\right\|^2\left\|\mathrm{E}'F\right\|^2+\frac{1}{NT}\left\|\frac{1}{N}1_N1_N'\right\|^2\left\|\frac{1}{T}\mathrm{E}'1_T1_T'F\right\|^2 \\ &\le \frac{2}{NT}\left\|\mathrm{E}'F\right\|^2+\frac{2}{NT}\left\|\frac{1}{T}\mathrm{E}'1_T1_T'F\right\|^2 \\ &\le O_p(1)+\frac{2}{NT}\left\|\mathrm{E}'1_T\right\|^2\left\|\frac{1}{T}1_T'F\right\|^2 = O_p(1)+2\left(\frac{1}{NT}\sum_{i=1}^{N}\left\|\sum_{t=1}^{T}\varepsilon_{it}\right\|^2\right)\left\|\bar{f}\right\|^2 = O_p(1),\end{aligned}$$

where 1_T is a $T\times 1$ vector of ones. Thus, $N^{-1/2}\left\|\mathrm{B}^{d}\right\| = O_p(1)$ and $(NT)^{-1/2}\left\|\ddot{\mathrm{E}}'\dot{F}\right\| = O_p(1)$.

Then, we have (i), because

$$\frac{1}{NT}\left|tr(A'\mathrm{B}^{d\prime}\ddot{\mathrm{E}}'\dot{F}G)\right| \le \frac{1}{\sqrt{T}}\left\|A\right\|\left\|G\right\|\frac{\left\|\mathrm{B}^{d}\right\|}{\sqrt{N}}\frac{\left\|\ddot{\mathrm{E}}'\dot{F}\right\|}{\sqrt{NT}} = O_p\left(\frac{1}{\sqrt{T}}\right)\times O_p(1) = O_p\left(\frac{1}{\sqrt{T}}\right).$$

We obtain (ii), because

$$\frac{1}{NT^2}\left|tr(A'\dot{F}'\ddot{\mathrm{E}}\ddot{\mathrm{E}}'F^d A)\right| \leq \frac{1}{NT^2}\|AA'\|\left\|\dot{F}'\ddot{\mathrm{E}}\ddot{\mathrm{E}}'\dot{F}\right\| \leq \frac{1}{T}\|A\|^2\left[\frac{1}{NT}\left\|\ddot{\mathrm{E}}'\dot{F}\right\|^2\right] = O_p\left(\frac{1}{T}\right).$$

Lemma 2: Suppose that two matrices *A* and *B* are symmetric of order *p*. Then,

$$\psi_{j+k-1}(A+B) \leq \psi_j(A) + \psi_k(B)\,,\ \ j+k \leq p+1.$$

Proof: See Onatski (2006) or Rao (1973, p. 68).

Proof of Theorem **1:** Observe that

$$\frac{1}{NT^2}\dot{F}'\ddot{X}\ddot{X}'\dot{F} = \frac{1}{NT^2}\dot{F}'(\dot{F}\mathrm{B}^{d\prime} + \ddot{\mathrm{E}})(\mathrm{B}^d\dot{F}' + \ddot{\mathrm{E}}')\dot{F}$$
$$= \left(\frac{\dot{F}'\dot{F}}{T}\right)\frac{\mathrm{B}^{d\prime}\mathrm{B}^d}{N}\left(\frac{\dot{F}'\dot{F}}{T}\right) + \left(\frac{\dot{F}'\dot{F}}{T}\right)\frac{\mathrm{B}^{d\prime}\ddot{\mathrm{E}}'\dot{F}}{NT} + \frac{\dot{F}'\ddot{\mathrm{E}}\mathrm{B}^d}{NT}\left(\frac{\dot{F}'\dot{F}}{T}\right) + \frac{\dot{F}'\ddot{\mathrm{E}}\ddot{\mathrm{E}}'\dot{F}}{NT^2}.$$

Thus, we have

$$\frac{\hat{\mathrm{B}}^{d\prime}\mathrm{B}^d}{N} = \frac{\mathrm{B}^{d\prime}\mathrm{B}^d}{N} + \frac{\mathrm{B}^{d\prime}\ddot{\mathrm{E}}'\dot{F}}{NT}A_T + A_T\frac{\dot{F}'\ddot{\mathrm{E}}\mathrm{B}^d}{NT} + A_T\left(\frac{\dot{F}'\ddot{\mathrm{E}}\ddot{\mathrm{E}}'\dot{F}}{NT^2}\right)A_T\,,$$

where $A_T = (\dot{F}'\dot{F}/T)^{-1}$ and $\|A_T\| = O_p(1)$ by Assumption A.

Now, let $\hat{\Xi}^l$ be the matrix of the eigenvectors corresponding to the first $l(\leq k)$ largest eigenvalues $\hat{\mu}_{NT,1} \geq \hat{\mu}_{NT,2} \geq \cdots \geq \hat{\mu}_{NT,l}$ of $\hat{\mathrm{B}}^{d\prime}\mathrm{B}^d/N$. Similarly, define Ξ^l for the matrix $\mathrm{B}^{d\prime}\mathrm{B}^d/N$. For any $l \leq r$, we have

$$\Sigma_{j=1}^l\hat{\mu}_{NT,j} = tr\left(\frac{1}{N}\hat{\Xi}^{l\prime}\hat{\mathrm{B}}^{d\prime}\mathrm{B}^d\Xi^l\right)$$
$$= tr\left(\frac{1}{N}\hat{\Xi}^{l\prime}\hat{\mathrm{B}}^{d\prime}\mathrm{B}^d\Xi^l\right) + 2tr\left(\frac{1}{NT}\Xi^{l\prime}\mathrm{B}^{d\prime}\ddot{\mathrm{E}}'\dot{F}A_T\Xi^l\right)$$
$$+tr\left(\frac{1}{NT^2}\hat{\Xi}^{l\prime}A_T\dot{F}'\ddot{\mathrm{E}}\ddot{\mathrm{E}}'\dot{F}'A_T\Xi^l\right)$$
$$\leq tr\left(\frac{1}{N}\Xi^{l\prime}\mathrm{B}^{d\prime}\mathrm{B}^d\Xi^l\right) + O_p\left(\frac{1}{\sqrt{T}}\right) + O_p\left(\frac{1}{T}\right)$$
$$= \Sigma_{j=1}^l\lambda_j\left(\frac{\mathrm{B}^{d\prime}\mathrm{B}^d}{N}\right) + O_p\left(\frac{1}{\sqrt{T}}\right) + O_p\left(\frac{1}{T}\right),$$

by Lemma 1, because $\left\|\hat{\Xi}^l\right\| = O_p(1)$ and $\left\|A_T\hat{\Xi}^l\right\| \leq \|A_T\|\left\|\Xi^l\right\| = O_p(1)$. In addition,

$$\begin{aligned}\Sigma_{j=1}^{l}\hat{\mu}_{NT,j} &= tr\left(\frac{1}{N}\widehat{\Xi^{l\prime}\mathrm{B}^{d\prime}}\mathrm{B}^{d}\Xi^{l}\right) \geq tr\left(\frac{1}{N}\Xi^{l\prime}\mathrm{B}^{d\prime}\mathrm{B}^{d}\Xi^{l}\right) \\ &= tr\left(\frac{1}{N}\Xi^{l\prime}\mathrm{B}^{d\prime}\mathrm{B}^{d}\Xi^{l}\right) + 2tr\left(\frac{1}{NT}\Xi^{l\prime}\mathrm{B}^{d\prime}\ddot{\mathrm{E}}'\dot{F}A_{T}\Xi^{l}\right) \\ &\quad + tr\left(\frac{1}{NT^{2}}\Xi^{l\prime}A_{T}\dot{F}'\ddot{\mathrm{E}}\ddot{\mathrm{E}}'\dot{F}A_{T}\Xi^{l}\right) \\ &= \Sigma_{j=1}^{l}\lambda_{j}\left(\frac{\mathrm{B}^{d\prime}\mathrm{B}^{d}}{N}\right) + O_{p}\left(\frac{1}{\sqrt{T}}\right) + O_{p}\left(\frac{1}{T}\right)\end{aligned},$$

Since these two results hold for any $l \leq r$, we have

$$\hat{\mu}_{NT,j} = \lambda_{j}\left(\frac{\mathrm{B}^{d\prime}\mathrm{B}^{d}}{N}\right) + O_{p}\left(\frac{1}{\sqrt{T}}\right) + O_{p}\left(\frac{1}{T}\right).$$

Thus, for $1 \leq j \leq r$, we have

$$p\lim_{T\to\infty}\hat{\mu}_{NT,j} = \lambda_{j}\left(\frac{\mathrm{B}^{d\prime}\mathrm{B}^{d}}{N}\right) > 0\,.$$

Next, since we have $rank(\mathrm{B}^{d}) = r$, we can rewrite $\mathrm{B}^{d} = AC'$, where A and C are $N \times r$ and $k \times r$ matrices, respectively, and $rank(A) = rank(C) = r$. Let $P(A) = A(A'A)^{-1}A'$, and $Q(A)$ = $1 - P(A)$. Using the fact that $P(A)\mathrm{B}^{d} = \mathrm{B}^{d}$ and $Q(A)\mathrm{B}^{d} = 0$, we can easily show

$$\frac{\widehat{\mathrm{B}^{d\prime}\mathrm{B}^{d}}}{N} = \frac{\mathrm{B}^{d\prime}[P(A)+Q(A)]\mathrm{B}^{d}}{N} = \frac{\mathrm{B}^{d\prime}P(A)\mathrm{B}^{d}}{N} + A_{T}\left(\frac{\dot{F}'\ddot{\mathrm{E}}Q(A)\ddot{\mathrm{E}}'\dot{F}}{NT^{2}}\right)A_{T}\,,$$

Thus, for $j = 1,\cdots,k-r$, we have

$$\begin{aligned}\lambda_{r+j}\left(\frac{\widehat{\mathrm{B}^{d\prime}\mathrm{B}^{d}}}{N}\right) &\leq \lambda_{r+1}\left(\frac{\mathrm{B}^{d\prime}P(A)\mathrm{B}^{d}}{N}\right) + \lambda_{j}\left(A_{T}\left(\frac{\dot{F}'\ddot{\mathrm{E}}Q(A)\ddot{\mathrm{E}}'\dot{F}}{NT^{2}}\right)A_{T}\right) \\ &\leq 0 + \lambda_{j}\left(A_{T}\left(\frac{\dot{F}'\ddot{\mathrm{E}}\ddot{\mathrm{E}}'\dot{F}}{NT^{2}}\right)A_{T}\right) \leq \frac{1}{NT^{2}}\left|tr(A_{T}'\dot{F}'\ddot{\mathrm{E}}\ddot{\mathrm{E}}'FA_{T})\right| = O_{p}(T^{-1}),\end{aligned}$$

where the first inequality is due to Lemma 2. Thus, for any $1 \leq r+1 \leq j' \leq k$, $\hat{\mu}_{NT,j'} = O_{p}\left(1/T\right)$. Notice that the second part holds even for $r = 0$, which is the “no-factor” case.

***Proof of Theorem* 2:** For $1 \leq j \leq r$, $p\lim_{T\to\infty}\hat{\mu}_{NT,j} > 0$, because $rank(\mathrm{B}^{d\prime}\mathrm{B}^{d}/N) = r$. Since $g(T) \to 0$, $\lim_{T\to\infty}\Pr[\hat{\mu}_{NT,j} > g(T) \mid j \leq r] = 1$. For $0 \leq r < j \leq k$, $p\lim_{T\to\infty}T\hat{\mu}_{NT,j} < \infty$.

Thus, $\lim_{T\to\infty} \Pr(\hat{\mu}_{NT,j} < g(T) \,|\, 0 \le r < j \le k) = \lim_{T\to\infty} \Pr(T\hat{\mu}_{NT,j} < Tg(T) \,|\, 0 \le r < j \le k) = 1$, because $Tg(T) \to \infty$ and $T\hat{\mu}_{NT,j} = O_p(1)$.

References

Ang, Andrew, Jun Liu, and Krista Schwarz, 2010, Using Stocks or Portfolios in Tests of Factor Models, Working paper.

Ahn, Seung C., and Alex R. Horenstein, 2011, Eigenvalue ratio test for the number of factors, working paper, Arizona State University.

Bai, Jushan, and Serena Ng, 2002, Determining the number of factors in approximate factor models. Econometrica, 70, 191 – 221.

Campbell, John Y., Andrew W. Lo and A. Craig MacKinlay, 1997, The Econometrics of Financial Markets (Princeton University Press, Princeton, New Jersey).

Chamberlain, Gary, and Michael Rothschild, 1983, Arbitrage, factor structure, and mean-variance analysis on large asset markets, Econometrica, 51, 1281 – 1304.

Chen, Long, Robert Novy-Marx and Lu Zhang, 2010, An Alternative Three Factor Model, working paper.

Chen, Nai-Fu, Richard Roll and Stephen A., Ross, 1986, Economic forces and the stock market, Journal of Business, 59, 383 – 403.

Connor, Gregory, and Robert A. Korajzcyk, 1993, A test for the number of factors in an approximate factor model, the Journal of Finance, 48, 1263 – 1291.

Cragg, John G., and Stephen G. Donald, 1997, Inferring the rank of a matrix, Journal of Econometrics, 76, 223 – 250.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, Journal of Financial Economics, 33, 3 – 56.

Fama, Eugene F., and James D. MacBeth, 1973, Risk, return and equilibrium: empirical tests, Journal of Political Economy, 81, 607 – 636.

Kan, Raymond, and Chu Zhang, 1999, Two-pass tests of asset pricing models with useless factors, the Journal of Finance, 54, 203 – 235.

Kleibergen, Frank, and Richard Paap, 2006, Generalized reduced rank tests using the singular value decomposition, Journal of Econometrics, 133, 97-126.

Lintner, John, 1965, The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, Review of Economic and Statistics, 47, 13 – 37.

Liu, Laura X., and Lu Zhang, 2008, Momentum profits, factor pricing, and macroeconomic risk, Review of Financial Studies, 21, 2417 – 48.

Merton, Robert, 1972, An intertemporal capital asset pricing model, Econometrica, 41, 867 – 87.

Mossin, Jan, 1966, Equilibrium in a capital asset market, Econometrica, 35, 768 – 783.

Onatski, A., 2010, Determining the number of factors from empirical distribution of eigenvalues, Review of Economic and Statistics, 92, 1004-1016.

Rao, Calyampudi R., 1973, Linear Statistical Inference and Its Applications, 2nd eds., John Wiley & Sons (New York, New York).

Robin, Jean-Marc., and Richard J. Smith, 2000, Tests of rank, Econometric Theory, 16, 151 - 175.

Ross, Stephen A., 1976, The arbitrage theory of capital asset pricing, Journal of Economic Theory, 13, 341 – 360.

Sharpe, William, 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, The Journal of Finance, 19, 425 – 442.

Treynor, Jack, 1999, Toward a theory of market value and risky assets, Unpublished manuscript, Final version in Asset Pricing and Portfolio Performance; Robert A. Korajczyk, ed., London: Risk Books, 15 – 22.

Zhou, Guofu, 1995, Small sample rank tests with applications to asset pricing, Journal of Empirical Finance, 2, 71 – 93.

Graph 1: The value of *g(d,T)* with different R_square and T

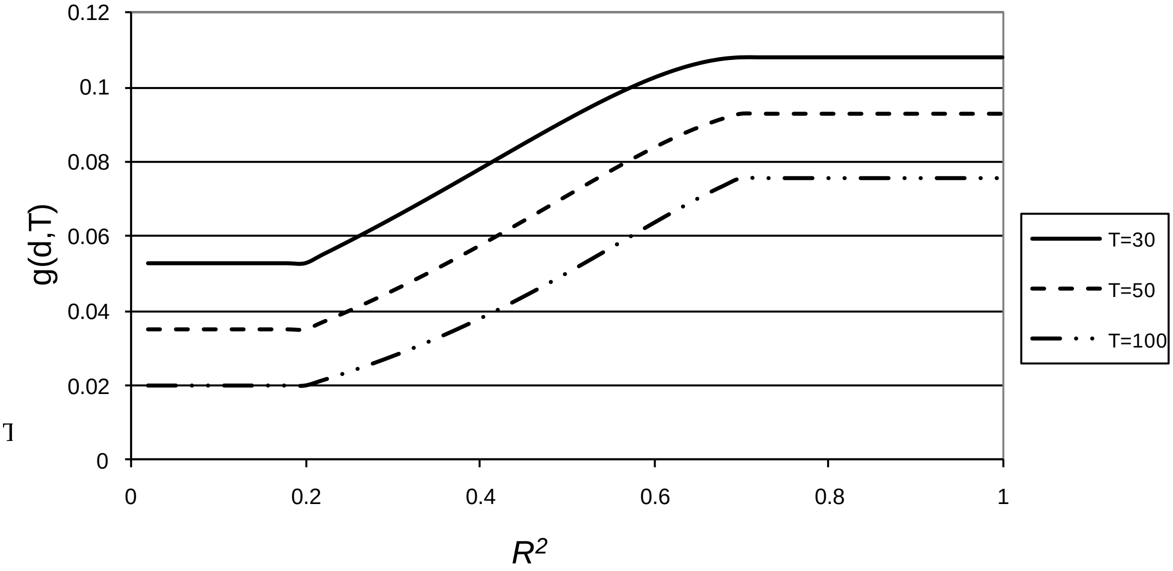

Table 1: Cases with *I.I.D.* and Both Serially and Cross-sectional Correlated Errors

T	N	$\rho = J = \delta = 0$						$\delta = 0.2, \rho = 0.5, J = 8$					
		$k=3$			$k=5$			$k=3$			$k=5$		
		$r=1$	$r=2$	$r=3$	$r=1$	$r=3$	$r=5$	$r=1$	$r=2$	$r=3$	$r=1$	$r=3$	$r=5$
50	50	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.04]	3.00 [0.00]	5.00 [0.06]	1.01 [0.10]	2.00 [0.03]	3.00 [0.00]	1.06 [0.24]	3.00 [0.06]	5.00 [0.04]
	100	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.05]	2.00 [0.00]	3.00 [0.00]	1.03 [0.17]	3.00 [0.03]	5.00 [0.00]
	200	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.03]	2.00 [0.00]	3.00 [0.00]	1.00 [0.05]	3.00 [0.03]	5.00 [0.00]
	500	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]
	1000	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]
100	50	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.05]	3.00 [0.00]	5.00 [0.00]
	100	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]
	200	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]
	500	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]
	1000	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]
200	50	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]
	100	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]
	200	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]
	500	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]
	1000	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]
500	50	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]
	100	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]
	200	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]
	500	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]
	1000	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]
1000	50	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]
	100	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]
	200	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]
	500	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]
	1000	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]

Data are generated with $\lambda_j = 1/r$ for $1 \le j \le r$. The value reported in each cell is the mean of the rank estimates from 1,000 simulations, and the value in the bracket is the standard deviation of the estimates.

Table 2: Cases with Cross-Sectional Correlation Only and Serial Correlation Only.

T	N	$\delta = 0.2, \rho = 0, J = 8$						$\delta = 0, \rho = 0.5$					
		k=3			k=5			k=3			k=5		
		r=1	r=2	r=3	r=1	r=3	r=5	r=1	r=2	r=3	r=1	r=3	r=5
50	50	1.00 [0.08]	2.00 [0.03]	3.00 [0.00]	1.08 [0.27]	3.00 [0.08]	5.00 [0.03]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.01 [0.08]	3.00 [0.00]	5.00 [0.06]
	100	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.02 [0.13]	3.00 [0.03]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]
	200	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.05]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]
	500	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]
	1000	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]
100	50	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]
	100	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]
	200	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]
	500	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]
	1000	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]
200	50	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.03]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]
	100	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]
	200	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]
	500	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]
	1000	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]
500	50	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]
	100	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]
	200	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]
	500	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]
	1000	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]
1000	50	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]
	100	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]
	200	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]
	500	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]
	1000	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]

Data are generated with $\lambda_j = 1/r$ for $1 \le j \le r$. The value reported in each cell is the mean of the rank estimates from 1,000 simulations, and the value in the bracket is the standard deviation of the estimates.

Table 3: Cases with Weak Factors

T	SNR	$\delta = \rho = J = 0$						$\delta = 0.2, \rho = 0.5, J = 8$					
		k=3			k=5			k=3			k=5		
		r=1	r=2	r=3	r=1	r=3	r=5	r=1	r=2	r=3	r=1	r=3	r=5
60	0.025	1.00 [0.15]	1.93 [0.26]	2..74 [0.44]	1.22 [0.42]	2.94 [0.30]	4.45 [0.55]	1.13 [0.35]	1.96 [0.28]	2.79 [0.42]	1.56 [0.56]	2.99 [0.39]	4.45 [0.57]
	0.05	1.01 [0.12]	2.00 [0.06]	3.00 [0.00]	1.26 [0.44]	3.02 [0.15]	4.99 [0.08]	1.17 [0.38]	2.04 [0.19]	3.00 [0.00]	1.64 [0.55]	3.13 [0.34]	4.99 [0.08]
	0.1	1.02 [0.12]	2.00 [0.04]	3.00 [0.00]	1.28 [0.45]	3.00 [0.04]	5.00 [0.00]	1.18 0.38	2.03 [0.18]	3.00 [0.00]	1.67 [0.55]	3.03 [0.18]	5.00 [0.00]
	0.2	1.01 [0.10]	2.00 [0.00]	3.00 [0.00]	1.19 [0.39]	3.00 [0.00]	5.00 [0.00]	1.14 [0.35]	2.01 [0.08]	3.00 [0.00]	1.50 [0.52]	3.00 [0.05]	5.00 [0.00]
	0.3	1.00 [0.05]	2.00 [0.00]	3.00 [0.00]	1.06 [0.23]	3.00 [0.00]	5.00 [0.00]	1.06 [0.25]	2.00 [0.04]	3.00 [0.00]	1.28 [0.45]	3.00 [0.00]	5.00 [0.00]
	0.5	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.07]	3.00 [0.00]	5.00 [0.00]	1.02 [0.12]	2.00 [0.00]	3.00 [0.00]	1.08 [0.26]	3.00 [0.00]	5.00 [0.00]
100	0.025	1.00 [0.00]	1.99 [0.06]	2.98 [0.13]	1.01 [0.11]	2.99 [0.08]	4.93 [0.24]	1.04 [0.20]	2.01 [0.11]	2.99 [0.08]	1.34 [0.48]	3.03 [0.18]	4.96 [0.20]
	0.05	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.01 [0.11]	3.00 [0.03]	5.00 [0.00]	1.05 [0.22]	2.01 [0.12]	3.00 [0.00]	1.37 [0.49]	3.06 [0.24]	5.00 [0.00]
	0.1	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.01 [0.11]	3.00 [0.00]	5.00 [0.00]	1.05 [0.22]	2.01 [0.11]	3.00 [0.00]	1.38 [0.49]	3.01 [0.11]	5.00 [0.00]
	0.2	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.01 [0.11]	3.00 [0.00]	5.00 [0.00]	1.04 [0.20]	2.00 [0.04]	3.00 [0.00]	1.30 [0.46]	3.00 [0.00]	5.00 [0.00]
	0.3	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.05]	3.00 [0.00]	5.00 [0.00]	1.02 [0.13]	2.00 [0.00]	3.00 [0.00]	1.12 [0.33]	3.00 [0.00]	5.00 [0.00]
	0.5	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.00 [0.00]	3.00 [0.00]	5.00 [0.00]	1.00 [0.00]	2.00 [0.00]	3.00 [0.00]	1.01 [0.10]	3.00 [0.00]	5.00 [0.00]

All simulated data are drawn with $N = 100$. The value reported in each cell is the mean of the estimated ranks from 1,000 simulations, and the value in the bracket is the standard deviation of the estimates.

Table 4: Cases with Both Strong and Weak Factors

N	T	$\delta=\rho=J=0$						$\delta=0.2, \rho=0.5, J=8$					
		$r=2, \lambda_1=1$			$r=3, \lambda_1=\lambda_2=1$			$r=2, \lambda_1=1$			$r=3, \lambda_1=\lambda_2=1$		
		λ_2=0.1	λ_2=0.2	λ_2=0.3	λ_3=0.1	λ_3=0.2	λ_3=0.3	λ_2=0.1	λ_2=0.2	λ_2=0.3	λ_3=0.1	λ_3=0.2	λ_3=0.3
100	40	1.99 [0.10]	2.00 [0.00]	2.00 [0.00]	2.91 [0.30]	3.00 [0.00]	3.00 [0.00]	1.99 [0.08]	2.00 [0.04]	2.00 [0.04]	2.96 [0.20]	3.00 [0.00]	3.00 [0.00]
	60	2.00 [0.04]	2.00 [0.00]	2.00 [0.00]	2.93 [0.25]	3.00 [0.00]	3.00 [0.00]	1.99 [0.03]	2.00 [0.00]	2.00 [0.00]	2.98 [0.15]	3.00 [0.00]	3.00 [0.00]
	100	2.00 [0.00]	2.00 [0.00]	2.00 [0.00]	2.98 [0.14]	3.00 [0.00]	3.00 [0.00]	2.00 [0.00]	2.00 [0.00]	2.00 [0.00]	2.99 [0.05]	3.00 [0.00]	3.00 [0.00]
	150	2.00 [0.00]	2.00 [0.00]	2.00 [0.00]	3.00 [0.04]	3.00 [0.00]	3.00 [0.00]	2.00 [0.00]	2.00 [0.00]	2.00 [0.00]	3.00 [0.00]	3.00 [0.00]	3.00 [0.00]
	200	2.00 [0.00]	2.00 [0.00]	2.00 [0.00]	3.00 [0.00]	3.00 [0.00]	3.00 [0.00]	2.00 [0.00]	2.00 [0.00]	2.00 [0.00]	3.00 [0.00]	3.00 [0.00]	3.00 [0.00]
	300	2.00 [0.00]	2.00 [0.00]	2.00 [0.00]	3.00 [0.00]	3.00 [0.00]	3.00 [0.00]	2.00 [0.00]	2.00 [0.00]	2.00 [0.00]	3.00 [0.00]	3.00 [0.00]	3.00 [0.00]
	500	2.00 [0.00]	2.00 [0.00]	2.00 [0.00]	3.00 [0.00]	3.00 [0.00]	3.00 [0.00]	2.00 [0.00]	2.00 [0.00]	2.00 [0.00]	3.00 [0.00]	3.00 [0.00]	3.00 [0.00]
	1000	2.00 [0.00]	2.00 [0.00]	2.00 [0.00]	3.00 [0.00]	3.00 [0.00]	3.00 [0.00]	2.00 [0.00]	2.00 [0.00]	2.00 [0.00]	3.00 [0.00]	3.00 [0.00]	3.00 [0.00]

Data are generated with three factor candidate variables ($k=3$). The value reported in each cell is the mean of the rank estimates from 1,000 simulations, and the value in the bracket is the standard deviation of the estimates.

Table 5: Cases with No Commom Factors

N	T	$\delta = \rho = J = 0$			$\delta = 0.2, \rho = 0.5, J = 8$		
		k=1	k=3	k=5	k=1	k=3	k=5
100	50	0.00 [0.00]	0.14 [0.35]	0.81 [0.51]	0.05 [0.21]	0.48 [0.52]	1.12 [0.60]
	100	0.00 [0.00]	0.01 [0.04]	0.05 [0.21]	0.01 [0.10]	0.16 [0.39]	0.59 [0.54]
	200	0.00 [0.00]	0.00 [0.00]	0.00 [0.00]	0.00 [0.05]	0.05 [0.22]	0.23 [0.43]
	500	0.00 [0.00]	0.00 [0.00]	0.00 [0.00]	0.00 [0.00]	0.01 [0.10]	0.03 [0.18]
	1000	0.00 [0.00]	0.00 [0.00]	0.00 [0.00]	0.00 [0.00]	0.00 [0.03]	0.01 [0.08]
200	50	0.00 [0.00]	0.09 [0.29]	0.72 [0.51]	0.02 [0.15]	0.40 [0.50]	1.07 [0.55]
	100	0.00 [0.00]	0.00 [0.03]	0.01 [0.10]	0.00 [0.04]	0.08 [0.28]	0.38 [0.50]
	200	0.00 [0.00]	0.00 [0.00]	0.00 [0.00]	0.00 [0.00]	0.01 [0.08]	0.05 [0.22]
	500	0.00 [0.00]	0.00 [0.00]	0.00 [0.00]	0.00 [0.00]	0.00 [0.00]	0.00 [0.00]
	1000	0.00 [0.00]	0.00 [0.00]	0.00 [0.00]	0.00 [0.00]	0.00 [0.00]	0.00 [0.00]

Data are generated with the factors with the SNRs of zero ($r = 0$). The value reported in each cell is the mean of the rank estimates from 1,000 simulations, and the value in the bracket is the standard deviation of the estimates.

Table 6: Estimation Results from Different Factor Models Using Individual Stock Returns

Sample Period	*(T,N)*	Jan. 1972 - Dec. 2004 *(396,313)*		Jan. 1972 - Dec. 1987 *(192,816)*		Jan. 1988 - Dec. 2004 *(204,1288)*		Jan. 1972 - Dec. 1978 *(84,1384)*		Jan. 1979 - Dec. 1992 *(168,1640)*		Jan. 1993 - Dec. 2004 *(144,1855)*	
	k	$\hat{r}$	R^2	$\hat{r}$	R^2	$\hat{r}$	R^2	$\hat{r}$	R^2	$\hat{r}$	R^2	$\hat{r}$	R^2
1) FF Factors	**3**	3	*0.073*	3	*0.097*	3	*0.078*	3	*0.155*	3	*0.065*	3	*0.104*
2) CRR Factors	**5**	1	*0.018*	1	*0.035*	1	*0.026*	2	*0.068*	1	*0.038*	2	*0.037*
3) FF and CRR factors	**8**	4	*0.089*	4	*0.127*	4	*0.103*	5	*0.209*	4	*0.100*	5	*0.139*
4) FF, UI, DEI factors	**5**	4	*0.081*	4	*0.112*	4	*0.089*	4	*0.180*	4	*0.080*	4	*0.120*
5) FF, MOM factors	**6**	4	*0.090*	4	*0.118*	3	*0.103*	5	*0.198*	4	*0.089*	4	*0.135*
6) FF, CNZ factors	**5**	4	*0.087*	4	*0.110*	4	*0.098*	4	*0.183*	3	*0.081*	4	*0.127*
7) FF, MOM, UI, DEI, CNZ factors	**10**	6	*0.110*	6	*0.146*	6	*0.131*	7	*0.250*	6	*0.119*	7	*0.174*

This table reports the estimation of the rank of the beta matrix for U.S individual stock returns. Each line of the table represents a different estimated model. For each model we report the number of factor candidates used (*k*), the estimated number of factors among the factor candidates ($\hat{r}$) and the average R^2 of the regressing the response variables on the factor-candidates.

Table 7: Estimation Results from Different Factor Models Using Stock Portfolio Returns

		100 Size and B/M Portfolios		25 Size and B/M Portfolios		25 MOM and Size Portfolios		49 Industrial Portfolios		30 Industrial Portfolios	
	k	$\hat{r}$	R^2	$\hat{r}$	R^2	$\hat{r}$	R^2	$\hat{r}$	R^2	$\hat{r}$	R^2
1) FF Factors	**3**	3	*0.42*	2	*0.66*	3	*0.33*	3	*0.12*	3	*0.12*
2) CRR Factors	**5**	1	*0.015*	1	*0.018*	2	*0.023*	1	*0.017*	1	*0.019*
3) FF, CRR factors	**8**	3	*0.43*	2	*0.67*	3	*0.34*	4	*0.14*	4	*0.14*
4) FF, MOM factors	**6**	3	*0.43*	2	*0.67*	3	*0.67*	4	*0.14*	4	*0.14*
5) FF, CNZ factors	**5**	4	*0.44*	3	*0.69*	3	*0.41*	4	*0.16*	4	*0.17*
6) FF, MOM, CRR,CNZ factors	**13**	4	*0.46*	3	*0.70*	3	*0.70*	6	*0.19*	7	*0.20*

This table reports the estimation of the rank of the beta matrix for U.S stock portfolio returns. For every portfolio set the time span is January 1972 - December 2004 (T=396). Each line of the table represents a different estimated model. For each model we report the number of factor candidates used (k), the estimated number of factors among the factor candidates ($\hat{r}$) and the average R^2 of the regressing the response variables on the factor-candidates.