Beta Matrix and Common Factors in Stock Returns*

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

Alex R. Horenstein[b]
University of Miami

Na Wang[c]
Hofstra University

This Version: June 2015

We consider the estimation methods for the rank of a beta matrix generated by risk factors and study which method would be appropriate for data with a large number (*N*) of risky assets. We find that a restricted version of Cragg and Donald's (1997) Bayesian Information Criterion (BIC) estimator is quite accurate even if *N* is large. Using twenty-six empirical factors for U.S. stock returns, we show that beta matrices from many multifactor asset-pricing models fail to have full column rank, suggesting that risk premiums in these models are under-identified and that many empirical factors are redundant.

Key Words: factor models, beta matrix, rank, risk factors, and asset prices.
JEL Classification: C01, C23, C31, G12

----

[*] We are grateful to Raymond Kan, George Korniotis, Alok Kumar, Richard Roll, Aurelio Vazquez, Todd Pronno, the participants at Columbia University, Instituto Tecnológico Autónomo de México, Universidad Autónoma de Nuevo León, Sogang University, Hofstra University, Arizona State University, University of Miami, Hitotsubashi University, and numerous conferences at which we presented this paper. We also thank Laura Xiaolei Liu, Jonathan Lewellen, and Stefan Nagel for sharing data with us.

[a] Department of Economics, W.P. Carey School of Business, Arizona State University, Tempe, AZ, 85287, USA; School of Economics, Sogang University, 35 Baekbeom-ro, Mapo-gu, Seoul 121-742. South Korea. Email: miniahn@asu.edu.

[b] Department of Economics, University of Miami, Coral Gables, FL, 33124. Email: horenstein@bus.miami.edu.

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

1. Introduction

Treynor (1962), Sharpe (1964), Lintner (1965) and Mossin (1966) developed the Capital Asset Pricing Model (CAPM), which predicts that a single market factor drives co-movement in asset returns. In later work, Merton's (1972) Intertemporal CAPM and Ross's (1976) Arbitrage Pricing Theory (APT) suggested that investors may consider multiple risk sources when making their investment decisions, providing the foundations for multifactor asset-pricing models. Since the advent of these two theories, many variables have been proposed as proxies for the true common risk factors that drive co-movement in asset returns. We refer to these variables as empirical factors. Some well-known examples are the three factors of Fama and French (1993) and the five macroeconomic factors of Chen, Roll and Ross (1986).[1]

Multifactor models typically predict that return on a risky asset is determined by the amount of undiversifiable risk from holding the asset (betas) and the necessary rewards (premiums) to induce investors to bear the risks. Under an asset-pricing model with K empirical factors, the K betas of an asset are simply the coefficients in a regression model, in which the dependent variable is the asset's (excess) return and the independent variables are the K factors. The $N \times K$ matrix of betas is the matrix of K betas from each of the N different asset returns. In order to examine the empirical relevance of an asset-pricing model, researchers need to estimate the beta matrix and the risk premiums related to each of the K factors. However, they can obtain consistent (asymptotically unbiased) risk premium estimators only if the $N \times K$ *true* (but unobservable) beta matrix has full column rank; that is, all of the columns in the true beta matrix should be linearly independent. It is not sufficient that the *estimated* beta matrix has full column rank. Thus, it is important to estimate the rank of the true beta matrix.[2]

The rank of a beta matrix is important not only for identifying the factors' risk premiums, but also for providing inferences on the number of latent common factors in asset returns. As we show, the rank of the beta matrix corresponding to a set of empirical factors equals the number of true latent factors that are correlated with the empirical factors. Thus, the beta matrix with a rank of two or more indicates the presence of multiple latent factors, given that the beta matrix must

[1] Harvey, Liu, and Zhu (2013) categorize 314 empirical factors from 311 different papers published in top-tier finance journals and current working papers since 1967.

[2] The literature has reported some special cases in which beta matrices fail to have full column rank. Kan and Zhang (1999a, 1999b) consider a case in which all betas corresponding to an empirical factor equal zero. Ahn, Perez and Gadarowski (APG, 2013) find that the estimated market betas have very small limited cross-sectional variations and that some betas of the three factors of Fama and French (1993) are highly multicollinear.

have a rank equal to one if the empirical factors are correlated with only one latent true factor. In addition, if some empirical factors are not correlated with latent true factors at all, the beta matrix cannot have full column rank. The rank of the beta matrix remains the same when such empirical factors are removed from the estimation.[3]

Many methods are available to estimate the rank of a matrix,[4] such as those in Anderson (1951), Cragg and Donald (1997), Robin and Smith (2000), and Kleibergen and Paap (2006). These methods are useful for cases in which both N (the number of assets analyzed) and K (the number of common factors) are sufficiently small compared to the sample size (T, the number of time series observations on each return and common factor). For example, Burnside (2010) uses a test proposed by Cragg and Donald (1997) to check whether the beta matrices from several multifactor models have full column rank. Conducting Monte Carlo simulations, he found that the test method has good power to detect rank deficiency of a beta matrix for data with $N = 25$ and $T = 240$. Applying the method to actual data, he also found that beta matrices from some models failed to have full column rank.

The principal focus of this paper is the identification of methods that accurately estimate the rank of a beta matrix even if N is large. A recent study by Lewellen, Nagel and Shanken (2010) suggests that the relevance of an asset-pricing model can be better tested by analyzing a large number of asset returns. This finding raises the following two questions: (i) Are the available rank estimators accurate even if data with large N are used (e.g., $N \geq 25$)? (ii) Is the rank of a beta matrix more accurately estimated when data with larger N are used? This paper addresses each of these questions.

Our estimator of interest is Cragg and Donald's (1997) Bayesian Information Criterion (BIC) estimator, computed under the assumption that return data are independently and identically distributed (*i.i.d.*) conditionally on empirical factors. We refer to this estimator as the restrictive BIC (RBIC) estimator. The main findings from our asymptotic analysis and simulations are the following. First, the RBIC estimator is consistent (asymptotically unbiased) even if data are, in fact, conditionally heteroskedastic and/or autocorrelated. Second, the RBIC

[3] However, the rank of a beta matrix corresponding to a set of empirical factors is not necessarily equal to the total number of true latent factors. For example, if a beta matrix corresponding to five empirical factors is found to have a rank of two, then only two latent factors are correlated with the five empirical factors. The total number of true latent factors can be greater than two if some latent factors are not correlated with the empirical factors used in the regressions.

[4] Al-Sadoon (2015) provides an excellent survey of the available rank estimation methods.

estimator performs well in the data comparable to monthly data of 20 or more years ($T \geq 240$), and the number of assets analyzed (N) is not greater than half of the number of time series observations ($N \leq T/2$). However, for the data with N close to T, the estimator is not reliable at all. Third, for data with $N \leq T/2$, the finite-sample performance of the RBIC estimator dominates that of other estimators by a large margin.

Empirically, many multifactor models have been proposed in the literature. Most of these models explain the cross-section of returns better than the CAPM does. However, with the richness of empirical factors, we need to address some important questions. Our paper aims to answer three of these questions: (i) Are these empirical factors capturing different common risk factors? (ii) How many common risk factors are correlated with the proposed empirical factors? (iii) Which empirical factors are truly important? To answer these questions, we apply the RBIC estimation method to the monthly and quarterly returns of U.S. stock portfolios and individual stocks over the period from 1972 to 2011. We consider 26 empirical factors proposed by previous studies. We analyze both monthly and quarterly returns using the five factors of Chen, Roll, and Ross (1986, CRR); the three factors of Fama and French (1993, FF); the three factor-model of Jagannathan and Wang (1996, JW); the four factors of Carhart (1997); the five factors of Fama and French (2015, FF5); the four factors of Hou, Xue and Zhang (2015, HXZ); and the short-term and long-term reversal factors. For quarterly returns, we also consider the macroeconomic factors used by six additional asset-pricing models: the consumption CAPM and the models of Lettau and Ludvigson (2001); Lustig and Van Nieuwerburgh (2004); Li, Vassalou and Xing (2006); Yogo (2006); and Santos and Veronesi (2006).

The key results from our actual data analysis are summarized as follows. First, for both portfolio and individual stock returns, many models fail to produce full rank beta matrices. This result is consistent with the findings of Burnside (2010) and Kleibergen and Paap (2006). Second, when all of the 26 empirical factors are regressed together on portfolio returns, the rank of the beta matrix is, at most, six. This result suggests that the empirical factors are correlated with, at most, six different latent risk factors. Third, the estimated ranks of the beta matrix using individual stock returns as response variables are smaller than the ranks obtained using portfolio returns. Overall, many of the proposed empirical factors appear to be correlated with the same latent risk factors, and thus, many empirical asset-pricing models' risk premiums are under-identified.

The rest of this paper is organized as follows. Section 2 introduces the factor model that we investigate and the BIC estimators. We also examine the asymptotic and finite-sample properties of the RBIC estimator. Section 3 reports our Monte Carlo simulation results, and Section 4 discusses the results from actual data analysis. Some concluding remarks follow in Section 5. All of the proofs of our theoretical results are given in the appendix.

2. Rank Estimation

We begin with the approximate factor model that Chamberlain and Rothschild (1983) consider. Let x_{it} be the response variable for the i^{th} cross-section unit at time t, where $i = 1, 2, \ldots, N$, and $t = 1, 2, \ldots, T$. Explicitly, x_{it} can be the (excess) return on asset i at time t. The response variables x_{it} depend on the J latent factors $g_t = (g_{1t}, \ldots, g_{Jt})'$. That is,

$$x_t = \eta + \mathrm{B}_g g_t + u_t , \tag{1}$$

where $x_t = (x_{1t}, \ldots, x_{Nt})'$; η is the N-vector of individual intercepts; B_g is the $N \times J$ matrix of factor loadings; and u_t is the N-vector of idiosyncratic components of individual returns with $E(g_t u_t') = 0_{J\times N}$. The matrix B_g is assumed to have full column rank ($rank(\mathrm{B}_g) = J$) because, otherwise, the model (1) can reduce to a model with $(J-1)$ or fewer factors. Both g_t and u_t are unobservable.

Observables are K empirical factors, $f_t = (f_{1t}, \ldots, f_{Kt})'$, which are correlated with r ($\leqslant J$) latent factors in g_t but not with the idiosyncratic errors in u_t. This assumption implies that

$$g_t = \theta + \Xi f_t + v_t , \tag{2}$$

where θ is the J-vector of intercepts and Ξ is a $J \times K$ matrix of coefficients with $rank(\Xi) = r$, $E(f_t u_t') = 0_{K\times N}$, $E(f_t v_t') = 0_{K\times J}$, and $E(v_t u_t') = 0_{J\times N}$. The error vector v_t is the vector of the components of g_t which is not correlated with f_t. If we substitute (2) into (1), we obtain

$$x_t = (\eta + \mathrm{B}_g \theta) + \mathrm{B}_g \Xi f_t + (\mathrm{B}_g v_t + u_t) \equiv \alpha + \mathrm{B} f_t + \varepsilon_t , \tag{3}$$

where we denote the i^{th} row of α and B by α_i and $\beta_i' = (\beta_{i1}, \beta_{i2}, \ldots, \beta_{iK})$, respectively. The focus of this paper is to estimate the rank of the beta matrix B, which we denote by r.

Some remarks follow on the linear factor model (3). First, the rank of $\mathrm{B} = \mathrm{B}_g \Xi$ is determined by that of Ξ because B_g is a full column rank matrix. That is, $rank(\mathrm{B}) = r$, which is the number of latent factors or their linear combinations that are correlated with the empirical factors f_t. Thus, the rank of the beta matrix B equals the maximum number of true latent factors that can be explained by the empirical factors f_t, and the rank can be smaller than the total number of true factors, J. Second, even if individual returns are generated by an exact factor model (in which the idiosyncratic errors in u_t in (1) are mutually independent), the errors in ε_t in (3) could be cross-sectionally correlated through $\mathrm{B}_g v_t$ unless the variables in f_t are perfectly correlated with g_t (so that $v_t = 0_{J\times 1}$). Accordingly, the rank of the beta matrix B needs to be estimated allowing for possible cross-sectional correlations in ε_t. Third, if $r = J$, the beta matrix can perfectly explain expected individual returns. For example, if $r = J = K$ -- that is, if the number of empirical factors f_t equals the number of true latent factors g_t ($K = J$) and if the former variables are correlated with all of the latter variables ($r = J$) -- the beta matrix B has the full column rank and can explain expected individual returns perfectly. Specifically, there exists a unique K-vector γ satisfying the pricing restriction, $E(x_t) = \mathrm{B}\gamma$,[5] where x_t contains excess returns (see Lewellen, Nagel and Shanken, 2010). However, if $r = J < K$ -- that is, if too many empirical factors are used compared to the number of true latent factors -- the beta matrix does not have full column rank. As a consequence, there are an infinite number of K-vectors γ satisfying the pricing restriction. The beta matrix B may still perfectly explain the expected excess returns $E(x_t)$, but the risk prices (γ) are not unique. For this case, Burnside (2010) shows that the two-pass estimator of γ is not asymptotically normal. This problem arises even if all of the empirical factors f_t are correlated with (linear combinations of) the true latent factors g_t. To explain assets' returns, the use of too many empirical factors is not harmful as long as each of them is correlated with true latent factors. However, the resulting two-pass estimates of factor prices would not provide reliable statistical inferences if the number of empirical factors used is greater than the number of true latent factors captured by them.

In order to discuss how to estimate the rank of the beta matrix B in (3), let us introduce some notation. Let

$$\hat{\Sigma}_{xf} = T^{-1}\Sigma_{t=1}^{T}(x_t - \bar{x})(f_t - \bar{f})'; \hat{\Sigma}_{ff} = T^{-1}\Sigma_{t=1}^{T}(f_t - \bar{f})(f_t - \bar{f})',$$

[5] If x_t contains raw returns, $E(x_t) = [1_N, \mathrm{B}]\gamma$, where 1_N is an N-vector of ones.

where $\overline{f} = T^{-1}\Sigma_{t=1}^{T} f_t$ and $\overline{x} = T^{-1}\Sigma_{t=1}^{T} x_t$. Then, the Ordinary Least Squares (OLS) estimator of B is given by $\hat{\mathrm{B}} = [\hat{\beta}_1, \hat{\beta}_2, ..., \hat{\beta}_N]' = \hat{\Sigma}_{xf}\hat{\Sigma}_{ff}^{-1}$. Under suitable conditions detailed in the appendix and the assumption of fixed *N*, we can show that as $T \to \infty$,

$$\sqrt{T}vec(\hat{\mathrm{B}}' - \mathrm{B}') \to_d N\left(0, \Omega\right), \tag{4}$$

where $vec(\bullet)$ is a matrix operator stacking all the columns in a matrix into a column vector, Ω is a finite positive definite matrix; and "$\to_d$" means "converges in distribution." Let $\hat{\Omega}$ be a consistent estimator of Ω as $T \to \infty$, and let us use A_p and M_p to denote $K\times p$ and $N\times p$ ($0 \le p \le K$) matrices of full column rank. Finally, let $\hat{A}_p$ and $\hat{M}_p$ be the minimizers of the objective function

$$\Pi_T(A_p, M_p, \hat{\Omega}) = Tvec(\hat{\mathrm{B}} - A_p M_p')'\hat{\Omega}^{-1}vec(\hat{\mathrm{B}} - A_p M_p'). \tag{5}$$

Cragg and Donald (1997) show the following:

$$\Pi_T(\hat{A}_p, \hat{M}_p, \hat{\Omega}) \to_p \infty, \text{ for } p < r \tag{6}$$

$$\Pi_T(\hat{A}_r, \hat{M}_r, \hat{\Omega}) \to_p \chi^2_{(N-r)(K-r)}, \text{ for } p = r, \tag{7}$$

where r is the true rank of B, and "$\to_p$" means "converges in probability." Based on these findings, they develop two different rank estimation methods. One estimator, which they refer to as the Testing Criterion (TC) method, is obtained by repeatedly testing the null hypotheses of $r = p$ (p = 0, 1, 2, ..., K – 1, where K is the number of empirical factors used) against the alternative hypothesis of full column rank. Each hypothesis is tested by using $\Pi_T(\hat{A}_p, \hat{M}_p, \hat{\Omega})$ as a $\chi^2_{(N-p)(K-p)}$ statistic. The TC estimate is the minimum value of *p* that does not reject the hypothesis of $r = p$. If all of the null hypotheses are rejected, the TC estimate equals *K*.

The other estimator, which Cragg and Donald (1997) refer to as the Bayesian Information Criterion (BIC) estimator, is obtained by finding the value of *p*, which minimizes the criterion function

$$C(p) = \Pi_T(\hat{A}_p, \hat{M}_p, \hat{\Omega}) - \ln(T) \times (N - p)(K - p), \tag{8}$$

where p = 0, 1, ... , K. For (8), $\ln(T)$ can be replaced by any $w(T)$ function such that $w(T) \to \infty$ and $w(T)/T \to 0$ as $T \to \infty$. Clearly, ln(T) is a possible function to use. The BIC estimator computed with any $w(T)$ is a consistent estimator.

While both the TC and BIC estimators have desirable large-sample properties, they are computationally burdensome to use in practice, especially for the cases with large *N*. This is so because the matrices A_p and M_p contain a large number of unknown parameters to be estimated, especially for the cases with large *N* and *p*. In unreported experiments, we attempt to compute the TC and BIC estimators using the same simulated data that we use for the results reported in Section 3. We observe that standard minimization algorithms failed to find $\hat{A}_p$ and $\hat{M}_p$ too often.

This computational problem can be resolved if we impose some restrictions on the covariance structure of the error terms. For example, suppose that the idiosyncratic error vectors ε_t are independently and identically distributed (*i.i.d.*) conditionally on the empirical factors f_t with the conditional variance-covariance matrix, $Var(\varepsilon_t \mid f_t) = \Sigma_{\varepsilon\varepsilon}$. The individual errors ε_{it} are still allowed to be cross-sectionally correlated; that is, the off-diagonal elements of $\Sigma_{\varepsilon\varepsilon}$ need not be zero. For this case, the computation procedures for the TC and BIC estimators are considerably simplified. When the error vectors are *i.i.d.* over time, $\hat{\Omega}_R = \hat{\Sigma}_{\varepsilon\varepsilon} \otimes \hat{\Sigma}_{ff}^{-1}$ is a consistent estimator of Ω, where "$\otimes$" means the Kronecker product and $\hat{\Sigma}_{\varepsilon\varepsilon}$ is a consistent estimator of $\Sigma_{\varepsilon\varepsilon}$; *i.e.*,

$$\hat{\Sigma}_{\varepsilon\varepsilon} = \frac{1}{T-K} \Sigma_{t=1}^{T} [(x_t - \bar{x}) - \hat{\mathrm{B}}(f_t - \bar{f})][(x_t - \bar{x}) - \hat{\mathrm{B}}(f_t - \bar{f})]'.$$

Cragg and Donald (1997) show that when $\hat{\Omega}_R$ is used for $\hat{\Omega}$,

$$\Pi_T(\hat{A}_K, \hat{M}_K, \hat{\Omega}_R) = 0; \; \Pi_T(\hat{A}_p, \hat{M}_p, \hat{\Omega}_R) = T \times \Sigma_{j=1}^{K-p} \hat{\psi}_j, \tag{9}$$

where $p = 0, 1, \dots, K-1$, $\hat{\psi}_j = \psi_j\left(\hat{\Sigma}_{ff} \hat{\mathrm{B}}' \hat{\Sigma}_{\varepsilon\varepsilon}^{-1} \hat{\mathrm{B}}\right)$, and $\psi_j(\bullet)$ denotes the j^{th} smallest eigenvalue of the matrix in the parenthesis. We refer to the TC estimator using (9) as the "restricted" TC (RTC) estimator.

Many previous studies have considered the statistic $\Pi_T(\hat{A}_p, \hat{M}_p, \hat{\Omega}_R)$. For example, if we replace $\hat{\psi}_j$ by $\ln(1+\hat{\psi}_j)$, $\Pi_T(\hat{A}_p, \hat{M}_p, \hat{\Omega}_R)$ becomes Anderson's (1951) Likelihood Ratio (LR) rank test statistic. When $p = r$ and the error vector ε_t is *i.i.d.* over time, $\Pi_T(\hat{A}_p, \hat{M}_p, \hat{\Omega}_R)$ and the LR statistic are asymptotically identical. Robin and Smith (2000) show that $\Pi_T(\hat{A}_p, \hat{M}_p, \hat{\Omega}_R)$ and the LR statistics with $p = r$ are asymptotically a weighted sum of independent χ_1^2 random variables, when the error vector is conditionally heteroskedastic and/or autocorrelated. A difficulty in using $\Pi_T(\hat{A}_p, \hat{M}_p, \hat{\Omega}_R)$ for the TC estimation is that its asymptotic distribution should be simulated using some parameter estimates. As a treatment of this problem, Kleibergen and Paap (2006) propose an alternative statistic that is asymptotically chi-squared even if the error vector is conditionally heteroskedastic and/or autocorrelated. Their statistic is also asymptotically identical to $\Pi_T(\hat{A}_p, \hat{M}_p, \hat{\Omega}_R)$ when $p = r$ and the error vector ε_t is *i.i.d.* over time.

The BIC estimator computed with $\hat{\Omega}_R$ for $\hat{\Omega}$ is another way to consistently estimate the rank of the beta matrix, even if the error vector is conditionally heteroskedastic and/or autocorrelated. We refer to this BIC estimator as the restricted BIC (RBIC) estimator, which is the minimizer of the criterion function (8) with (9):

$$C_R(p) = T \times \Sigma_{j=1}^{K-p} \hat{\psi}_j - \ln(T) \times (N-p)(K-p) .$$

An intuition for the consistency of the RBIC estimator follows. While Cragg and Donald (1997) have shown the consistency of the BIC estimator (based on (8)) using results (6) and (7), the same consistency result is obtained with $\hat{\Omega}$ in (5) replaced by any (asymptotically) positive definite matrix as long as (6) holds and $\Pi_T(\hat{A}_r, \hat{M}_r, \hat{\Omega})$ is $O_p(1)$ (an asymptotically bounded random variable). The statistic $\Pi_T(\hat{A}_r, \hat{M}_r, \hat{\Omega})$ need not be an asymptotically chi-squared random variable, as in (7). According to Robin and Smith's (2000) results, $\Pi_T(\hat{A}_r, \hat{M}_r, \hat{\Omega}_R)$ is $O_p(1)$ regardless of whether the error vectors are *i.i.d.* over time. Thus, although the RBIC estimator is computed under the *i.i.d.* assumption, it is consistent under much more general conditions.[6]

Furthermore, we can test whether or not some betas are cross-sectionally constant by comparing the ranks of two matrices: the beta matrix B and its demeaned version, $Q_N\mathrm{B}$ = $(\dot{\beta}_1, \dot{\beta}_2, ..., \dot{\beta}_N)'$, where Q_N = $I_N - N^{-1}1_N 1_N'$, 1_N is an N-vector of ones, $\dot{\beta}_i = \beta_i - \overline{\beta}$, and $\overline{\beta} = N^{-1}\Sigma_{i=1}^N \beta_i$. If a column of B (or a linear combination of the columns of B) is proportional to a vector of ones, the corresponding column of the demeaned beta matrix ($Q_N\mathrm{B}$) becomes a zero vector. Thus, $rank(Q_N\mathrm{B}) = r-1$. For the same reason, if two columns of B are proportional

[6] The proof of the consistency under more general conditions is in the Appendix.

to a vector of ones, $rank(Q_N \mathrm{B}) = r-2$. If no column of B has constant betas, the two matrices B and $Q_N\mathrm{B}$ must have the same ranks. Therefore, comparing the estimated ranks of the beta matrix (B) and the demeaned beta matrix ($Q_N\mathrm{B}$), we can determine whether a constant-beta factor exists in f_t.

The RBIC estimator can be easily modified to estimate the rank of the demeaned beta matrix. Define the following criterion function:

$$D_R(p) = T \times \Sigma_{j=1}^{K-p} \psi_j \left(\hat{\Sigma}_{ff} \hat{\mathrm{B}}' Q_N (Q_N \hat{\Sigma}_{\varepsilon\varepsilon} Q_N)^{+} Q_N \hat{\mathrm{B}} \right) - \ln(T) \times (N-1-p)(K-p), \tag{10}$$

where p = 1, ... , K–1, and $D_R(K) = 0$. Then, the RBIC estimator of the demeaned beta matrix ($Q_N\mathrm{B}$) equals the minimizer of $D_R(p)$. We refer to this estimator as the RBICD estimator. Note that even for the cases in which $\hat{\Sigma}_{\varepsilon\varepsilon}$ has full column rank, $Q_N \hat{\Sigma}_{\varepsilon\varepsilon} Q_N$ may not. That is why we use the Moore-Penrose generalized inverse of $Q_N \hat{\Sigma}_{\varepsilon\varepsilon} Q_N$ in $D_R(p)$.

Because the RBIC (RBICD) estimator is consistent as $T \to \infty$ with fixed N, we can expect the estimator to have good finite-sample properties for the data with large T and relatively small N. However, it is unknown whether the estimator would remain consistent as both N and T grow infinitely. One immediate problem with using the RBIC (RBICD) estimator for the data with large N is that $\hat{\Sigma}_{\varepsilon\varepsilon}$ is not invertible if $N > T$. This numerical problem can be resolved if we use the Moore-Penrose generalized inverse matrix of $\hat{\Sigma}_{\varepsilon\varepsilon}$ ($\hat{\Sigma}_{\varepsilon\varepsilon}^{+}$) instead of $\hat{\Sigma}_{\varepsilon\varepsilon}^{-1}$. However, it is still difficult to determine whether the RBIC (RBICD) estimator computed with $\hat{\Sigma}_{\varepsilon\varepsilon}^{+}$ would be consistent with the data with both large N and T. We investigate this question using Monte Carlo simulations in the next section.

3. Finite-Sample Properties of RTC and RBIC Estimators

3.1. Simulation Setup

The foundation of our simulation exercises is the following data-generating process:

$$x_{it} = \alpha_i + \Sigma_{j=1}^{K} \beta_{ij} f_{jt} + \varepsilon_{it};\ \varepsilon_{it} = \phi \frac{(\xi_{i1} h_{1t} + \xi_{i2} h_{2t})}{\sqrt{\xi_{i1}^2 + \xi_{i2}^2}} + \sqrt{1-\phi^2} v_{it},$$

where the empirical factors f_{jt} and the ξ_{i1}, ξ_{i2}, h_{1t}, h_{2t}, and v_{it} in ε_{it} are all randomly drawn from $N(0,1)$. For simplicity, we set $\alpha_i = 0$ for all i. Under this setup, the variance of error ε_{it} is equal to one for all *i* and *t*. The factor components h_{1t} and h_{2t} can be viewed as common latent factors that are not correlated with the empirical factors f_{jt}. The errors ε_{it} are cross-sectionally correlated through h_{1t} and h_{2t} if $\phi \neq 0$. We have also considered the cases in which the errors are serially correlated. We do not report the results because they are not materially different from the results reported below. For the reported simulations, we set ϕ = 0.2. The use of greater values for ϕ does not change the estimation results substantially.

We generate the beta matrix B by the following three steps. First, we draw an $N \times r$ random matrix B_g such that its first column equals the vector of ones, and the entries in the other columns are drawn from $N(0,1)$. Second, we 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.[7] Finally, we set $\mathrm{B} = \mathrm{B}_g \Lambda^{1/2} C'$, where $\Lambda = diag(\lambda_1, ..., \lambda_r)$. This setup is equivalent to the case in which individual returns are generated by r true factors $g_t = (g_{1t}, ..., g_{rt})' = \Lambda^{1/2} C' f_t$ with $Var(g_t) = \Lambda$. The factor loading matrix corresponding to g_t is B_g. By construction, the factors in g_t are mutually independent, and the betas corresponding to g_{1t} are constant over different individual response variables. Under this setup, $rank(\mathrm{B}) = r$ and $rank(Q_N \mathrm{B}) = r - 1$. We use this setup to investigate the finite-sample performances of the RBIC and RBICD estimators. We also consider the RTC estimator because it is consistent under our data-generating process. Robin and Smith (2000) and Kleibergen and Paap (2006) suggest alterative TC estimators. We do not consider the performances of these alternative TC estimators because they are numerically the same as the RTC estimator if they are computed with $\hat{\Omega}_R$ for $\hat{\Omega}$.

Under our data-generating setup, each of the empirical factors f_t can have non-zero explanatory power for individual response variables x_{it}, even if the beta matrix B does not have full column rank. The parameter λ_j equals the variance of the j^{th} true factor, g_{jt}. Given that B_g is drawn from $N(0,1)$, the λ_j equals the signal-to-noise ratio (SNR) of g_{jt} (e.g., the ratio of the return variations caused by the true factor g_{jt} and by idiosyncratic errors ε_{it}). The population

[7] The random matrix is of the form $M'M$, where the entries of the $K \times K$ matrix *M* are drawn from *N*(0,1).

average *R*-square (average explanatory power of the empirical factors f_t for individual response variables x_{it}) equals $(\Sigma_{j=1}^{r}\lambda_j)/(1+\Sigma_{j=1}^{r}\lambda_j)$.

We try four different values of *T*: *T* = 60, 120, 240, and 480. For each *T*, we generate six different sets of response variables *N*: *N* = 25, 50, 75, 100, 200 and 400. For each combination of *N* and *T*, we also consider two cases: one with five empirical factors (*K* = 5) and the other with fifteen factors (*K* = 15). For both cases, we use two different beta ranks, $r = 1$ and 4, and three different values, $\lambda =$ 0.03, 0.1 and 0.5, for the SNR of a true factor. When $r = 4$, we use the same value of λ for all four factors: $\lambda_j = \lambda$ for $j = 1$, 2, 3 and 4. For each combination of *N, T, K*, *r* and λ, we generate 1,000 samples.

Our simulation setup may not represent the true data-generating processes of asset returns. However, we choose parameter values such that the simulated data have properties similar to those of actual data. First, empirical studies of asset-pricing models routinely use monthly data over five, ten, twenty, or forty years. The values of *T* are chosen to be consistent with this practice. Second, the empirical factors proposed in the literature sometimes have low explanatory power for stock returns. To investigate the cases in which empirical factors have limited explanatory power for response variables, we generate data with latent factors with very low SNRs, $\lambda = 0.03$. Third, the idiosyncratic error components of actual returns are likely to be cross-sectionally correlated. Under our simulation setup, the error terms are cross-sectionally correlated through the unobserved factor components h_{1t} and h_{2t}. We could have generated cross-sectionally correlated errors using the estimated variance-covariance matrix of the errors from actual data; however, for the actual data with N close to or greater than *T*, we could not consistently estimate the variance matrix of the error vector $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t}, ..., \varepsilon_{Nt})'$. For example, the estimated variance matrix is not invertible if *N* > *T*, although the true variance matrix would be invertible. Thus, the errors generated based on an estimated variance matrix from actual data are likely to have a different cross-sectional covariance structure from that of the idiosyncratic error components of actual returns.

Finally, the empirical factors proposed in the literature (f_t in our notation) are unlikely to be perfectly correlated with true latent factors (g_t in our notation). When this is the case, the finite-sample performances of the rank estimators can depend on the degrees of correlation between f_t and g_t. Our simulations can provide useful information for such general cases.

When empirical factors are imperfect proxy variables, the errors ε_{it} should be cross-sectionally correlated. In addition, for the cases of imperfect correlation, the SNR of a latent factor (λ_j) in our simulations can be interpreted as the SNR of a linear combination of empirical factors that is maximally correlated with the factor. For example, if a latent factor has an SNR of 0.03, it can be interpreted as the linear combination of the empirical factors maximally correlated with the latent factor having an SNR of 0.03 and explaining approximately 3% (=0.03/1.03) of the total variation in response variables.

3.2. Simulation Results

Given that the RTC and RBIC estimators are designed for data with relatively large *T* and small *N*, our simulation aims to address the data size that is required to obtain reliable inferences from the estimators.

We begin by considering the finite-sample performance of the RTC estimator. Table 1 reports the RTC estimation results from our simulations with five and fifteen empirical factors (*K* = 5 and *K* = 15). We consider two cases: $r = 1$ when *K* = 5, and $r = 4$ when *K* = 15. Data are generated such that the true latent factors (or linear combinations of the five empirical factors) have SNRs of 0.03, 0.1, and 0.5. Using the combination of low and high values of SNRs, our simulation results provide better guidance for the analysis of actual return data. Table 1 reports the percentages (%) of correct estimation by the RTC estimator. The percentages of underestimation and overestimation are reported below in parentheses.

Table 1 shows that the RTC estimator performs rather poorly in most of the *N* and *T* combinations. For the cases with *N* = 25, the accuracy of the RTC estimator is not greater than 50% when *T* is less than or equal to 120. The estimator's performance deteriorates as *T* decreases or *N* decreases. For example, for the cases with *N* = 25 and *T* = 480, the accuracy of the RTC estimator is between 87% and 90.1%. However, the accuracy of the estimator is not better than 28% for the cases with $N \geq 50$ and $T \leq 240$, and not better than 9% for the cases with $N \geq 100$ and $T \leq 480$. As *N* increases to about $T/2$, the accuracy of the RTC estimator drops to near zero. Similar patterns are observed in the simulations with any combination of *K* and *r*, whether latent factors are weak ($\lambda = 0.03$) or strong ($\lambda = 0.5$). Furthermore, in all of the cases considered in Table 1, the performance of the RBIC estimator dominates that of the RTC estimator, as Tables 2 and 3 show.

Table 2 reports the performance of the RBIC estimator for the cases with five empirical factors (K = 5). The data-generating process is the same as the one described in the beginning of this section. The accuracy of the RBIC estimator appears to have a non-monotonic relationship with the number of response variables (N). For the cases with $r=1$, the accuracy of the estimator increases with N when $N \le T/2$ and decreases with N when $T/2 < N \le T$. The estimator overestimates the rank of the beta matrix when $T/2 < N \le T$. When $N > T$, the accuracy of the estimator increases with N up to some points (e.g., (T, N) = (60, 100) and (120, 200)). However, as N increases further, the estimator begins to underestimate the beta rank, and its accuracy drops sharply. In order to investigate this irregular behavior of the RBIC estimator further, we conducted some additional experiments using data with $N > T/2$. To save space, we summarize the results here. For a given T, the degree of overestimation by the RBIC estimator increases as N increases from $N = T/2$ to $N = T$. However, the accuracy of the estimator improves as N increases from $N = T$ to $N = 2T$ (or near 2T). Then, as N increases further from $N = 2T$ (or near 2T), the estimator starts to underestimate the beta rank. In general, the RBIC estimator tends to overestimate the beta rank if $T/2 < N < 2T$, while it severely underestimates the rank when $N > 2T$. The tendency to overestimate reverses to a tendency to underestimate at some point in $T < N < 2T$. Thus, the RBIC estimator occasionally performs well at some points when $T < N < 2T$.

We find a similar pattern for the cases with $r = 4$. The accuracy of the RBIC estimator increases with N up to some points where $T/2 < N \le 2T/3$. However, the accuracy drops quickly with N after the points. The results reported in Table 2 suggest that the RBIC estimator should be used with caution for data with $N > T/2$. In particular, the estimator appears to be inappropriate for data with $N > 2T/3$.[8]

Furthermore, for the cases with $T \ge 240$ and $N \le T/2$, the accuracy of the RBIC estimator is near 100% except for the cases with $\lambda = 0.03$. Not surprisingly, the RBIC estimator tends to underestimate the rank of the beta matrix in the presence of factors with very low SNRs.[9] However, the accuracy of the RBIC estimator improves with N as long as $N \le T/2$. For

[8] From some unreported simulations, we find that when $T \ge 960$, the RBIC estimator is accurate as long as $N \le 0.8 \times T$. It appears that range of the ratio N/T in which the RBIC estimator remains reliable increases with T. However, for the data with $T \le 480$, the condition $N \le T/2$ seems to be important to secure the reliability of the RBIC estimator.

[9] For the cases with r = 4, we use the same value for all of the 4 SNRs ($\lambda_1 = ... = \lambda_4$). In unreported experiments, we also try different SNR values. Not surprisingly, the accuracy of the RBIC estimator depends on the smaller SNR

example, as shown in panel (b) of Table 2, for the cases with $T = 240$ and $\lambda = 0.03$, the RBIC estimator's accuracy improves from 33.3% to 100% as *N* increases from 25 to 100. We observe a similar pattern in the cases with $T = 480$ and $\lambda = 0.03$.

The accuracy of the RBIC estimator may depend on the number of empirical factors (*K*). Table 3 reports the estimation results from the cases with *K* = 15. Compared to the results in Table 2, the accuracy of the RBIC estimator generally falls as more empirical factors are used, while the explanatory power of empirical factors (*R*-squared) remains the same. The performance pattern of the RBIC estimator regarding the data size is identical to that obtained in Table 2. Even when the explanatory power of empirical factors is very low ($\lambda_j = 0.03$), the RBIC estimator is quite accurate if *T* or *N* (when $N < T/2$) is large. For example, when $\lambda = 0.03$ and $r = 4$, the accuracy of the RBIC estimator is 99.9% when $(T, N) = (240,100)$ and 98.4% when $(T, N) = (480,50)$.

Finally, we consider the performance of the RBICD estimator. Under our simulation setup, the rank of the demeaned beta matrix equals *r* – 1. We use the simulated data with *r* = 4. The results, reported in Table 4, are similar to those reported in Tables 2 and 3. The accuracy of the RBICD estimator increases with *N* when $N \leq T/2$ and also increases monotonically with *T*.

The three main results from our simulation exercises are the following. First, the RTC estimator is generally inaccurate unless *T* is very large (e.g., *T* = 480) and *N* is sufficiently small compared to *T* (e.g., *T* = 480 and *N* ≤ 25).[10] Second, the RBIC estimator outperforms the RTC estimator in most of the cases we consider. Third, the accuracy of the RBIC estimator has a non-monotonic relationship with the number of response variables (*N*). The power of the estimator initially increases with *N* but falls as *N* increases further from some point (e.g, $N = T/2$). For the data with *N* close to *T*, we cannot expect a reliable inference from the RBIC estimator. The RBICD estimator shows the same pattern. Both the RBIC and RBICD estimators produce reliable inferences in the simulated data whose sizes are compatible with those of the actual data

values. For example, when we use 0.03 for two and 0.1 for two other SNRs, the performance of the RBIC estimator is similar to that reported for the cases with $\lambda_j = 0.03$ ($j = 1, \ldots, 4$) in panel (b) of Table 2.

[10] In our simulations we have not tried unrestricted TC estimators based on Cragg and Donald (1997), Robin and Smith (2000), or Kleibergen and Paap (2006) because they are asymptotically identical to the RTC estimator under our data-generating process. In addition, the unrestricted TC estimators (computed allowing conditional heteroskedasticity and autocorrelation) are not expected to perform better than the RTC estimator in our simulated data. When the idiosyncratic components of returns are heteroskedastic and/or autocorrelated, unrestricted TC estimators would perform better than the RTC estimator. However, the accuracy of the unrestricted TC estimators is unlikely to be better than the numbers reported in Table 1.

we analyze in the next section -- even for the cases in which some empirical factors have low explanatory power for returns.

4. Application

In this section, we estimate the ranks of different beta matrices using various combinations of empirical factors. We conduct our estimation with monthly and quarterly data from January 1972 to December 2011 and analyze both portfolio and individual stock return data. For the analysis of monthly data, we consider a total of sixteen non-repetitive empirical factors from six different multifactor models and two reversal factors proposed in the literature: the three factors of Fama and French (1993, FF3); the four factors of Carhart (1997, Carhart); the five factors of Fama and French (2015, FF5); the four factors of Hou, Xue and Zhang (2015, HXZ); the five factors of Chen, Roll and Ross (1986, CRR); and the three factors of Jagannathan and Wang (1996, JW). The FF3 factors are the CRSP value-weighted portfolio return minus the return on the one-month Treasury bill (VW), Small Minus Big (SMB) and High Minus Low (HML) factors.[11] The Carhart factors are the FF3 plus the momentum factor (MOM, selling losers and buying winners six to twelve months ago). The FF5 factors are the FF3, Robust Minus Weak (RMW) and Conservative Minus Aggressive (CMA) factors. The HXZ factors are the VW, SMB, Investment-to-Asset (I/A) and Return on Equity (ROE) factors. The CRR factors are five macroeconomic variables: industrial production (MP), unexpected inflation (UI), change in expected inflation (DEI), the term premium (UTS), and the default premium (UPR).[12] The JW factors are the VW, LAB (growth rate of labor income) and UPR factors.[13] Finally, we include the two reversal factors (REV) to address the different momentum effects discussed in Jegadeesh and Titman (1993) (REV_S: selling winners and buying losers one month ago; and REV_L: selling winners and buying losers thirteen to sixty months ago).

[11] All of the FF factors are available from Kenneth French's website (www.dartmouth.edu/~kfrench/).

[12] The CRR factors are available from Laura Xiaolei Liu's webpage (http://www.bm.ust.hk/~fnliu/research.html). For detailed information on how these factors have been constructed, see Liu and Zhang (2008). The UPR factor (default premium) equals the yield spread between BAA- and AAA-rated bonds.

[13] The LAB factor is constructed using the NIPA 2.1 and NIPA 2.6 tables for quarterly and monthly data, respectively. The tables are available at the Bureau of Economic Analysis webpage: http://www.bea.gov/iTable. Specifically, the factor is the growth rate of total personal income minus personal dividend income divided by total population.

4.1 Results from Portfolio Returns

Six sets of portfolio returns are used for regressions. Five of them consist of 25 Size and Book-to-Market (B/M) portfolios, 30 Industrial portfolios, 25 Size and Momentum portfolios, 25 Operating Profitability and Investment portfolios, and the 32 Size, Operating Profitability and Investment portfolios. In addition, following the suggestion of Lewellen, Nagel, and Shanken (2010), we also consider the combined set of the 25 Size and B/M and 30 Industrial portfolios.[14] The excess return on each portfolio is computed using the one-month Treasury bill rate as the risk-free rate.

For sensitivity analysis, we also estimate the above factor models using quarterly observations. Analyzing quarterly portfolio returns, we can examine five additional factor models that are discussed in Lewellen, Nagel, and Shanken (2010): the consumption CAPM (CCAPM); the two conditional CCAPMs of Lettau and Ludvigson (2001, LL); the durable-consumption CAPM of Yogo (2006, Yogo); the conditional CAPM of Santos and Veronesi (2006, SV); and the investment-based CAPM of Li, Vassalou and Xing (2006, LVX).[15] The empirical factors used by these models are CG (aggregate consumption growth rate) for the CCAPM; CG, CAY (aggregate consumption-to-wealth ratio) and CG×CAY for the LL model; VW, DCG (durable-consumption growth rate) and NDCG (nondurable-consumption growth rate) for the Yogo model; VW and VW×LC (labor income-to-consumption ratio) for the SV model; and DHH (change in the gross private investment for households), DCORP (change in gross private investment for non-financial corporate firms), and DNCORP (change in gross private investment for non-financial non-corporate firms) for the LVX model.[16] Note that most of these models added at a quarterly frequency, together with the CRR and the JW models,

[14]According to Lewellen, Nagel and Shanken (2010), the 25 Size and B/M portfolios have a strong factor structure favoring the FF model, and, thus, model specification tests can produce more-reliable inferences when the tests are done with additional portfolios that are not strongly correlated with the SMB and HML factors.

[15]We do not include the factors of Lustig and Van Nieuwerburgh (2003) because the time series data on the housing collateral ratio (MYMO) factor are available only up to the first quarter of 2005.

[16] We are grateful to Jonathan Lewellen and Stefan Nagel for sharing their data with us. The CG, CAY, and LC factors can be directly downloaded or constructed using the data available from Sydney Ludvigson's website, http://www.econ.nyu.edu/user/ludvigsons. The DCG and NDCG factors are constructed using data from the *NIPA* 2.3.3 and *NIPA* 2.3.5 tables. We also use the *Consumer-Durables Goods: Chain-Type Quantity Indexes for Net Stock* table for constructing DCG. All of these tables are available at the Bureau of Economic Analysis webpage: http://www.bea.gov/iTable. For the DHH, DCORP, and DNCORP factors we use the *Flow of Funds Accounts* tables available at the Federal Reserve Board's webpage: http://www.federalreserve.gov. Specifically, we use the table FA155019005 for the DHH factor, the tables FA105019005 and FA105020005 for the DCORP factor, and the tables FA115019005 and FA115020005 for the DNCORP factor.

include non-traded factors. As is well known -- and as we observe in our empirical results -- non-traded factors have a significantly lower signal-to noise-ratio than traded ones.

4.1.1 Results from Monthly Stock Portfolio Returns

In this subsection, we report the estimation results obtained using the six sets of monthly portfolio returns as response variables. The sample period is from January 1972 to December 2011 (T = 480). The cross-sectional dimension N equals the number of portfolios used to estimate the beta matrix. The estimation results from the entire sample period are reported in panel (a) of Table 5. Panels (b) and (c) report results using the first and second half of the sample period, respectively. For each combination of portfolio sets and empirical factors, we report the adjusted R-square ($\bar{R}^2$, explanatory power of empirical factors) and the estimated rank of the beta matrix using the RBIC estimator. The RBICD estimation results are reported in parentheses. Our simulation results, reported in Section 3.2, indicate that the RBIC estimator produces reliable inferences when using data with $T \geq 240$ and $N \leq T/2$. The data used for Table 5 satisfy all these conditions.

The main observations from panel (a) of Table 5 are the following. First, for any of the six portfolio sets, the RBIC estimator predicts that the beta matrix corresponding to the FF3 factors has the full column rank ($r = 3$). These results are consistent with the notion that the FF3 factors are correlated with three linearly independent latent risk factors. We obtain a similar result from the estimation of the HXZ model. For five of the six portfolio sets, the HXZ model produces full column beta matrices ($r = 4$). One exception is the 25 Operating Profitability and Investment portfolios, for which the beta matrix has the rank of three. Overall, the four HXZ factors appear to be correlated with four linearly independent latent risk factors.

Second, two other models occasionally generate full column rank beta matrices. The Carhart model generates full column rank beta matrices for two of the six portfolio sets. So, depending on the portfolio returns analyzed, the MOM factor sometimes captures one latent factor that cannot be explained by the FF3 factors alone. The FF5 model produces a full column rank beta matrix only for the 32 Size, Operating Profitability and Investment portfolios. For other sets of portfolios, the beta matrices corresponding to the FF5 factors are all rank-deficient (the estimated ranks are usually four). This result indicates that the risk premiums corresponding to the FF5 factors would be underidentified.

Third, the beta matrices of the MOM plus REV, CRR and JW models are all rank-deficient, regardless of which portfolio returns are analyzed.[17] The explanatory power of these models for returns is not as strong as that of the FF3, Carhart, HXZ or FF5 models.[18] The explanatory power of the CRR factors for portfolio returns is particularly low, explaining no more than 2% of the average total variation in the portfolio returns analyzed.

Fourth and finally, when all of the sixteen factors are used in the regression, the estimated beta rank is four or five, depending on the portfolio returns analyzed. In other words, the sixteen empirical factors are only correlated with up to five latent factors. One or two additional latent factors appear to exist in returns that are identified by the sixteen empirical factors, but not by the FF3 factors alone. These additional latent factors appear to be correlated with the MOM factor (in Carhart), the RMW and CMA factors (in FF5), or the I/A and ROE factors (in HXZ). Table 5 shows that when these five empirical factors are added to the FF3 model (an eight-factor model), for any of the six sets of portfolio returns, the beta matrix has exactly the same rank as the beta matrix from all of the sixteen empirical factors.[19]

In unreported experiments, we examine many other larger sets of portfolio returns obtained by merging two different portfolio sets. The estimation results from the larger data are similar to those reported in Table 5. For example, the rank estimation results from the 25 Size and Momentum plus 30 Industry Portfolios are similar to those from the 25 Size and B/M plus 30 Industry Portfolios. The combination of the 25 Operating Profitability and Investment portfolios with another set of portfolios yields results similar to those from the 32 Size, Operating Profitability and Investment portfolios. In some cases, a five-factor model combining the FF3 and HZX factors generates a full column rank beta matrix, but only occasionally.

The RBICD estimation results are reported (in parentheses) in panel (a) of Table 5. Nearly 80% of the rank estimation results using the RBICD estimates are consistent with those using the RBIC estimates. For instance, for the 32 Size, Operating Profitability and Investment Portfolios, the rank of the demeaned beta matrix has the same rank as the original beta matrix for any factor model. In contrast, for the 25 Size and Momentum Portfolios, the RBICD estimates

[17] Similar to our analysis, Kleibergen and Paap (2006) also find that the beta matrix from JW models does not have full column rank.

[18] Among the FF3 factors, the market factor has the strongest explanatory power for portfolio returns.

[19] We also investigate whether some of the eight factors can be omitted without reducing the rank of the beta matrix. We find that the estimated rank of the beta matrix remains unchanged only when we omit the I/A factor, and the rank always drops for some sets of portfolios as any other empirical factor is omitted from regression.

are quite often smaller than the RBIC estimates by one, suggesting that some betas are constant over different portfolios. However, the RBIC and RBICD estimation results are generally similar.

In panels (b) and (c) of Table 5, we execute the same exercises as before, using two subsample periods: 1972-1991 and 1992-2011. In each of these subsamples, we have T=240. As shown in our simulations in Section 3.2, using smaller T often leads to underestimation of rank if the proposed factors have weak SNR. Panel (b) of Table 5 corresponds to the first half of the sample period and demonstrates that the estimated ranks of many factor models are smaller than those using the entire sample period. For example, the HXZ model generates a rank of three (instead of four) in every portfolio set. Putting all of the sixteen factors together, the estimated rank is reduced by one in five of the six portfolios, compared to those in panel (a) of Table 5.

In contrast, panel (c) of Table 5, which corresponds to the second half of the sample period, shows results very similar to those using the entire sample period. Given that the reported average R^2s in the second period are smaller than those in the first period, it seems surprising that the estimated ranks in panel (c) are generally larger than those reported in panel (b). Table 6 explains this apparent puzzle.

In Table 6, we report the average R^2s of each individual factor using the two subsample periods: 1972-1991 and 1992-2011. All empirical factors have been orthogonalized with respect to the market portfolio (VW). This means that each factor's average R^2 represents the amount of variation it captures in each sample, beyond what VW has already captured. We observe several patterns from the average R^2 during these two subsample periods. First, the market portfolio's SNR clearly dominates those of the other factors, which additionally capture less than 10% of the total variation. Second, the non-traded factors (MP, UI, DEI, UTS, UPR and LAB) have average R^2s of 1% or less. Hence, the RBIC and RBICD might find them redundant even if they are correlated with a latent source of risk.

In Table 6, we further calculate the ratio of the average R^2 in the second subsample divided by that of the first subsample (Ratio). We observe that the market portfolio's (VW) explanatory power decreases between 30% and 15% from the first period to the second, depending on which portfolio set is used as the response variable. This result explains why the average R^2 of every model that includes VW is usually smaller in the second period than those reported in panel (a) and panel (b) of Table 5. At the same time, most traded factors' average R^2 are much higher in the second period than in the first. For example, when using the 25 Size and Book to Market Portfolios as response variables, the average R^2 of RMW is 25 times larger in the second period than in the first one. This explains why the rank of the beta matrix estimated for the FF5 model is generally larger in the second period than in the first one. The other traded factors (except REV_L) show similar patterns, though less pronounced. Hence, we should expect to see smaller ranks in the first period than in the second, since the weak factors are weaker in the first period. All of these results are in line with what we observe in the simulation section.

4.1.2 Results from Quarterly Stock Portfolio Returns

Using quarterly returns, we re-estimate the FF3, Carhart, FF5 and HXZ models, which most often produce full column rank beta matrices for monthly portfolio return data. We compute the quarterly returns for the same six sets of portfolios as in Table 5, using the data from the first quarter of 1972 to the fourth quarter of 2011 (T = 160). The estimation results are presented in Table 7.[20]

The estimation results from the four models are very similar to those in panel (a) of Table 5. We find again that the beta matrix of the FF3 factors has full column rank for all of the six portfolio sets. The RBICD estimation results suggest that the demeaned beta matrix of the FF factors also has a rank of three. In fact, the FF3 model is the only model that has the full column rank beta matrix for all six sets of portfolios. The HXZ model also produces a full column rank beta matrix for five sets of portfolios, but not for the 25 Operating Profitability and Investment Portfolios. For the Carhart model, which adds the MOM factor to the FF3 model, the beta matrix has a rank of three in most cases. The beta matrix has full column rank only for the 25 Size and Momentum portfolios. The FF5 factors produce a full rank beta matrix for two sets: the 25 Operating Profitability and Investment portfolios, and with the 32 Size, the Operating Profitability and Investment portfolios. In contrast, the eight-factor model with the FF5, MOM, I/A and ROE factors now has the beta matrix rank equal to five for all six portfolio sets.

Next, we consider the five macroeconomic factor models discussed in Lewellen, Nagel and Shanken (2010). The models are the unconditional CCAPM and the models of Lettau and Ludvigson (2001, LL); Yogo (2006, Yogo); Santos and Veronesi (2006, SV); and Li, Vassalou and Xing (2006, LVX). We observe all of the empirical factors used by these models only quarterly and refer to all of these factors as *quarterly macroeconomic factors*. These macroeconomic factors, like those in the CRR and JW models, are mostly non-traded factors. Hence, their signal-to-noise ratios are significantly lower than those of the traded factors.

[20] We do not include the analysis in the subsample periods since we have only 160 quarterly observations for each asset.

The results from the RBIC and RBICD estimations are reported in Table 7. The CCAPM factor has a rank of zero for four of the six portfolio sets. The beta matrices corresponding to LL and LVX factors have a rank of one for all six portfolio sets, while the demeaned beta matrices all have a rank of zero. The RBIC and RBICD rank estimates from Yogo's three-factor model are all ones for all six portfolio sets. These results imply that none of the LL, LVX and Yogo models can identify risk premiums. The SV model is the only one that generates a full column rank beta matrix ($r = 2$) consistently for all six portfolio sets. However, the demeaned beta matrix has a rank of one for five portfolio sets.

When we add all of the quarterly macroeconomic factors to the eight-factor models of the FF5, MOM, I/A and ROE factors, the rank of the beta matrix increases for the 30 Industrial Portfolios and the 25 Size and B/M plus 30 Industrial Portfolios. For other sets of portfolios, adding the quarterly macroeconomic factors does not increase the rank of the beta matrix. For industry-based portfolios, the quarterly macroeconomic factors appear to be correlated with one additional latent factor that the eight factors do not explain. We also run some unreported tests to identify which of the macroeconomic factors or which linear combinations among them can increase the rank of the beta matrix. None of the single macroeconomic factors add significant explanatory power for quarterly returns. We found that at least six macroeconomic factors are needed to increase the rank of the beta matrix by one.

4.2 Results from Individual Stock Returns

In this subsection, we use the monthly individual stock excess returns over the one-month Treasury bill as response variables. Individual stock returns include dividends. The data are downloaded from CRSP and exclude REITs (Real Estate Investment Trusts) and ADRs (American Depositary Receipts). We also exclude the observations in which the stocks show more than 300% excess returns in a given month, because such huge variations are unlikely due to changes in common factors. Excessively high or low returns are most likely to be driven by idiosyncratic shocks.

Table 8 shows the results with individual stock returns. For each dataset, we sample 10,000 random sets of 50 stocks (N=50) and 10,000 random sets of 200 stocks (N=200). This exercise gives us further information using real data regarding the performance of the RBIC (RBICD) estimator when the signal-to-noise ratio of the factors is low and the number of cross-section variables N is smaller or bigger than $T/2$. The table reports the average result of the RBIC, RBICD, and Adjusted R^2 together with their standard deviation over the 10,000 random samples in parentheses.

Panel (a) of Table 8 shows the case in which $N<T/2$. The rank of the beta matrices is, on average, less than three, even if we use the sixteen factor candidates together. The Fama-French three-factor model is the only one in which the average rank of beta matrices is close to the number of factors used. The RBIC and RBICD rank estimators are not able to capture the non-traded factors (CRR and two of the three factors in JW). This is expected since the SNR is, on average, less than one third of that in the tables using portfolio returns as response variables, where non-traded factors were already not captured. In this setup, RBIC and RBICD seem to miss some traded factors, too.

Panel (b) of Table 8 shows the effect of increasing N to the point at which $N>T/2$. For the entire sample period (1972-2011), we still have $N<T/2$, while for the two subsample periods, we have $N>T/2$. The results in the entire sample period remain similar to those in panel (a) of Table 8. However, in both of the subsample periods, the ranks of the beta matrices estimated using RBIC and RBICD are close to the number of empirical factors used. As we illustrate in the simulation section, for samples similar in size to the two subsamples used in panel (b) of Table 8, the estimators tend to overestimate the rank of the beta matrices when $N>T/2$. Therefore, it seems prudent in finite samples to maintain the $N<T/2$ rule.

4.3. Summary of Empirical Results

The main results from our analysis of actual return data can be summarized as follows. First, the FF3 factors appear to be correlated with three linearly independent latent factors using both portfolio and individual stock returns and different sample periods. The HXZ four factors produce full rank beta matrices quite often when using portfolio returns as response variables, and, thus, they appear to be correlated with four different latent sources of risk. The FF5 factors also seem correlated with four latent factors in most of the cases in which portfolio returns are used as response variables, generating rank-deficient beta matrices too often. The Carhart model also generates rank-deficient beta matrices quite often, except when tested using portfolios sorted by momentum. The remaining factor models that we have considered fail to generate full column rank beta matrices across assets and time periods.

Second, among the twenty-six empirical factors, we can identify up to five or six latent factors in the U.S. stock returns when we use portfolios as response variables. When using individual stock returns, the same set of empirical factors seems to capture a smaller number of factors. Based on our simulations, the difference is driven mainly by the facts that individual stock returns contain more idiosyncratic risk, and the factor structure has a smaller SNR ratio.

Third, the rank of the beta matrix obtained using the longer period is greater than or equal to those obtained using shorter periods when keeping the same response variables and empirical factors.

Fourth, except for the market portfolio (VW), most traded factors have larger SNR in the later subsample period, 1992-2011, than in the earlier subsample period, 1972-1991. Hence, the estimated ranks of the beta matrices generated by these factors in the second subsample are usually larger than or equal to those estimated in the first subsample.

Fifth, and last, if the time series dimension is not large enough, we should use datasets with $T<N/2$. At the same time, to maximize SNR, we should apply the estimators proposed in this paper, using portfolio returns as response variables and traded factors as independent variables.

5. Concluding Remarks

In this paper, we estimate the ranks of the beta matrices of many empirical factors that have been used in the literature. We conduct Monte Carlo simulations to determine which of the existing rank estimators produces the most-reliable inferences, particularly when the number of assets to be analyzed is large. Our simulation results indicate that the Testing Criterion estimators are reliable only if the number of time series observation (T) is very large and the number of assets (N) is substantially small. In contrast, the restricted BIC estimator proposed by Cragg and Donald (1997) is quite accurate for data with $T \geq 240$ and $N \leq T/2$.

Our empirical analysis shows that the Fama and French (1993) three-factor model appears to be correlated with three linearly independent latent factors across portfolios and time periods. The Hou, Xue and Zhang (2015) four-factor model generates full rank beta matrices when using portfolio returns as response variables and samples including recent periods. Other factor models in our analysis often fail to produce beta matrices with full column rank. This finding suggests that the risk premiums corresponding to many proposed factor models are under-identified. In addition, the twenty-six empirical factors we have considered identify, at most, five or six different latent risk factors using stock portfolio returns. However, their explanatory power for individual stock returns is limited, with the estimated rank often less than three.

Bai and Ng (2002) and Onatski (2010) develop estimation methods that can estimate the number of latent factors without using empirical factors. Using these methods, they find evidence that two latent factors drive U.S. individual stock returns. Our exercise using individual stock returns yields similar results. In contrast, though, our results using portfolio returns suggest up to four additional latent factors. This difference can be explained by the lower signal-to-noise ratios of the empirical factors on individual stock returns, suggesting the use of portfolio returns as dependent variables when testing asset-pricing models.

Appendix

In this appendix, we show that the RBIC estimator is consistent. A similar proof can be used to show the consistency of the RBICD estimator. We begin with the following assumptions.

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

Assumption B (idiosyncratic errors): For given fixed *N*, $E(\varepsilon_t) = 0_{N\times 1}$ and

$$\frac{1}{T}\Sigma_{t=1}^{T}\varepsilon_t\varepsilon_t' \to \Sigma_{\varepsilon\varepsilon},$$

where $\Sigma_{\varepsilon\varepsilon}$ is a *N*×*N* finite and positive definite matrix.

Assumption C (factors and idiosyncratic errors): For given fixed *N*,

$$\frac{1}{\sqrt{T}}\Sigma_{t=1}^{T}\varepsilon_t \otimes \begin{pmatrix} 1 \\ f_t \end{pmatrix} \to_d N(0_{N(K+1)\times 1}, \Xi),$$

where Ξ is a finite and positive definite matrix.

Assumption D (betas): For fixed $N > K$, $\mathrm{rank}(\mathrm{B}) = r$ and $rank(Q_N \mathrm{B})$ = r^d ($\leq r$), where 0 ≤ r^d ≤ r ≤ *K*.

Assumptions A-C are the standard assumptions under which the OLS estimator of B is consistent and asymptotically normal. Chapters 3 and 4 of White (1984) provide detailed conditions under which the three assumptions hold. We note that the assumptions allow both factors and idiosyncratic errors to be heteroskedastic and/or autocorrelated over time. In addition, under the assumptions, it is straightforward to show that result (4) in Section 2 holds with Ω = $(I_N \otimes \Sigma_{ff}^{-1})\Xi(I_N \otimes \Sigma_{ff}^{-1})$ and that $\hat{\Sigma}_{\varepsilon\varepsilon}$ is a consistent estimator of $\Sigma_{\varepsilon\varepsilon}$ for fixed *N*:

$$\hat{\Sigma}_{\varepsilon\varepsilon} \to_p \Sigma_{\varepsilon\varepsilon}. \tag{A1}$$

Lemma 1: Under Assumptions A-D, for some positive number a_p,

$$\Pi_T(\hat{A}_p,\hat{M}_p,\hat{\Omega}_R)/T \to_p a_p \text{, for } p<r; \quad \text{(A1)}$$

$$\Pi_T(\hat{A}_r,\hat{M}_r,\hat{\Omega}_R) = T\times\Sigma_{j=1}^{K-r}\hat{\psi}_j = O_p(1). \quad \text{(A2)}$$

***Proof of Lemma* 1:** When $p<r$, for any A_p and M_p matrices of full column rank, $\text{rank}(A_pM'_p)\le p<r$. Thus, for any A_p and M_p, $\mathrm{B}\neq A_pM'_p$. Consequently,

$$vec(\hat{\mathrm{B}}'-\hat{A}_p\hat{M}'_p)\to_p \xi_p,$$

where ξ_p is a non-zero constant vector. Thus, by Cragg and Donald (1997),

$$\begin{aligned}\Pi_T(\hat{A}_p,\hat{M}_p,\hat{\Omega}_R)/T &= vec(\hat{\mathrm{B}}'-\hat{A}_pM'_p)'(\hat{\Sigma}_{\varepsilon\varepsilon}^{-1}\otimes\hat{\Sigma}_{ff})vec(\hat{\mathrm{B}}'-\hat{A}_pM'_p)' \\ &= \xi'_p(\Sigma_{\varepsilon\varepsilon}^{-1}\otimes\Sigma_{ff})\xi_p+o_p(1)\equiv a_p+o_p(1),\end{aligned}$$

which shows (A1). Smith and Robin (2000) show that $T\Sigma_{j=1}^{K-r}\psi_j(\hat{\Sigma}_{ff}\hat{\mathrm{B}}'\Sigma_{\varepsilon\varepsilon}^{-1}\hat{\mathrm{B}})$ is asymptotically a weighted sum of $(N-r)(K-r)$ independent chi-squared random variables even if $\hat{\Sigma}_{ff}$ and $\hat{\Sigma}_{\varepsilon\varepsilon}$ are replaced by any conformable finite positive definite matrices. Thus, (A2) holds.

Theorem 1: Let $\hat{r}$ be the RBIC estimator. Then, under Assumptions A-D, $\lim_{T\to\infty}\Pr(\hat{r}=r)$ = 1. That is, $\hat{r}$ is a consistent estimator of r.

***Proof of Theorem* 1:** Note that

$$m(\hat{r})-m(r)=(1-\hat{r})(K-\hat{r})-(1-r)(K-r)<(>)0, \quad \text{(A3)}$$

if $\hat{r}>(<)r$. If $\hat{r}\neq r$, $C_R(r)-C_R(\hat{r})>0$. Thus, $\Pr(\hat{r}\neq r)\le\Pr[C_R(r)-C_R(\hat{r})>0]$. Thus, we can complete the proof by showing that as $T\to\infty$, $\Pr[C_R(r)-C_R(\hat{r})>0]\to 0$. Suppose that $\hat{r}>r$. Then,

$$\begin{aligned}\Pr\left[C_R(r)-C_R(\hat{r})>0\right] &= \Pr\left[T\Sigma_{j=1}^{K-r}\hat{\psi}_j - T\Sigma_{j=1}^{K-\hat{r}}\hat{\psi}_j + \ln(T)(m(\hat{r})-m(r))>0\right] \\ &= \Pr\left[T\Sigma_{j=1}^{K-r}\hat{\psi}_j+\ln(T)(m(\hat{r})-m(r))>0\right]\to 0\end{aligned} \quad \text{(A4)}$$

because $T\times\Sigma_{j=1}^{K-r}\hat{\psi}_j=O_p(1)$ by Lemma 1 and $\ln(T)(m(\hat{r})-m(r))\to-\infty$ by (A3). Therefore, $\Pr(\hat{r}>r)\to 0$ as $T\to\infty$. Similarly, if $\hat{r}<r$,

$$\Pr[C_R(r)-C_R(\hat{r})>0]=\Pr\left[\Sigma_{j=1}^{K-r}\hat{\psi}_j-\Sigma_{j=1}^{K-\hat{r}}\hat{\psi}_j+\frac{\ln(T)}{T}(m(\hat{r})-m(r))>0\right]\to 0 \quad \text{(A5)}$$

because $\ln(T)/T \to 0$, and $\Sigma_{j=1}^{K-\hat{r}}\hat{\psi}_j - \Sigma_{j=1}^{K-r}\hat{\psi}_j \to_p -a_r < 0$ by Lemma 1. Thus, by (A4) and (A5), as $T \to \infty$, $\Pr[C_R(r) - C_R(\hat{r}) > 0] \to 0$ whenever $\hat{r} \neq r$. This completes the proof.

Remark: We note that the above proof is almost identical to the proof of Theorem 3 in Cragg and Donald (1997). The only difference is that we use (A2) for the proof. For their proof, Cragg and Donald use the fact that $\Pi_T(\hat{A}_r, \hat{M}_r, \hat{\Omega})$ is an asymptotically chi-squared random variable with the degrees of freedom equal to $m(p) = (N-p)(K-p)$. We add this proof to clarify that the proof requires $\Pi_T(\hat{A}_r, \hat{M}_r, \hat{\Omega})$ to be only an asymptotically bounded random variable, not to be a random variable of a particular distribution.

References

Ahn, Seung C., and Alex R. Horenstein, 2013, Eigenvalue ratio test for the number of factors, *Econometrica* 81, 1203 – 1227.

Ahn, Seung C., Marcos F. Perez and C. Gadarowski, 2013, Two-pass estimation of risk premiums with multicollinear and near-invariant betas, *Journal of Empirical Finance* 20, 1 – 17.

Al-Sadoon, M.M., 2015, A general theory of rank testing, Barcelona GSE Working paper 750.

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

Burnside, Craig, 2010, Identification and inference in linear stochastic discount factor models, NBER Working Paper 16634.

Carhart, Mark M., 1997, On persistence in mutual fund performance, *The Journal of Finance* 52, 57 – 82.

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

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

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.

Fama, E. and K. French, 2015, A Five-Factor Asset Pricing Model, *Journal of Financial Economics* 116, 1-22.

Harvey, Campbell, Yan Liu and Heqing Zhu, 2013, . . . and the Cross-Section of Expected Returns, working paper.

Hou K., Xue C. and L. Zhang, 2015, Digesting Anomalies: An Investment Approach, *Review of Financial Studies*, 28, 650-705.

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

Kan, Raymond, and Chu Zhang, 1999b, GMM tests of stochastic discount factor models with useless factor, *Journal of Financial Economics* 54, 103 – 127.

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

Lewellen, Jonathan, Stephan Nagel, and Jay Shanken, 2010, A skeptical appraisal of asset pricing tests, *Journal of Financial Economics* 96, 175 – 194.

Lettau, Martin, and Sydney Ludvigson, 2001, Resurrecting the (C)CAPM: A cross-sectional test when risk premia are time-varying, *Journal of Political Economy* 109, 1238 – 1287.

Li, Qing, Maria Vassalou, and Yuhang Xing, 2006, Sector investment growth rates and the cross section of equity returns, *Journal of Business* 79, 1637 – 1665.

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 – 2448.

Lustig, Hanno N., and Stijn Van Nieuwerburgh, 2005, Housing collateral, consumption insurance, and risk premia: An empirical perspective, *Journal of Finance* 60, 1167 – 1219.

Merton, Robert, 1972, An intertemporal capital asset pricing model, *Econometrica* 41, 867 – 887.

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

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

Pastor, Lubos, and Robert F. Stambaugh, 2003, Liquidity Risk and Expected Stock Returns, *Journal of Political Economy* 111, 642 – 685.

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

Roll, Richard, 1977, A critique of the asset pricing theory's tests, Part I: On past and potential testability of the theory, *Journal of Financial Economics* 4, 129 – 176.

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

Santos, Tano, and Pietro Veronesi, 2006, Labor income and predictable stock returns, *Review of Financial Studies* 19, 1 – 44.

Sharpe, William, 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, *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.

White, Halbert, 1984, Asymptotic Theory for Econometricians (Academic Press, INC, San Diego, California).

Yogo, Motohiro, 2006, A consumption-based explanation of expected stock returns, *Journal of Finance* 61, 539 – 580.

Table 1: RTC Estimation of Beta Matrix from Five Empirical Factors

Reported are the percentages (%) of correct estimation by the RTC estimator from 1,000 simulated datasets. The percentages (%) of under- and overestimation by each estimator are reported in parentheses (•,•). Data are drawn by $x_{it} = \Sigma_{j=1}^{K} \beta_{ij} f_{jt} + \varepsilon_{it}$; $\varepsilon_{it} = \phi(\xi_{1i} h_{1t} + \xi_{2i} h_{2t}) / (\xi_{1i}^2 + \xi_{2i}^2)^{1/2} + (1-\phi^2)^{1/2} v_{it}$, where K = 5 and 15, $\phi = 0.2$, and f_{jt} (j = 1, ... , K), $h_{j't}$, $\xi_{j'i}$ ($j' = 1,2$) and v_{it} are all randomly drawn from $N(0,1)$. For the beta matrix B, we draw an $N\times r$ random matrix B_g such that its first column equals the vector of ones and the entries in the other columns are drawn from $N(0,1)$. 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. $\mathrm{B} = \mathrm{B}_g \Lambda^{1/2} C'$, where $\Lambda = diag(\lambda_1, ..., \lambda_r)$. The rank of the beta matrix equals one or four (r = 1, 4) for K equals to 5 and 15, respectively. RTC refers to the Testing Criterion estimator of Cragg and Donald (1997), which is computed under the assumption that the ε_{it} are *i.i.d.* over time (but cross-sectionally correlated).

RTC		(a) K = 5, r = 1, $\lambda_1 = \lambda$			(b) K = 15, r = 4, $\lambda_i = \lambda$ for i = 1,...,4		
		λ = .03 R^2=.029	λ = .10 R^2=.091	λ = .50 R^2=.333	λ = .03 R^2=.107	λ = .10 R^2=.286	λ = .50 R^2=.667
T=60	N =25	4.1% (0.0, 95.9)	2.9% (0.0, 97.1)	2.3% (0.0, 97.7)	1.9% (0.0, 98.1)	0.1% (0.0, 99.9)	0.0% (0.0, 100)
	N =50	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)
T=120	N =25	48.5% (0.0, 51.5)	46.4% (0.0, 53.6)	45.2% (0.0, 54.8)	45.9% (3.9, 50.2)	26.1% (0.0, 73.9)	20.7% (0.0, 79.3)
	N =50	0.1% (0.0, 99.9)	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)
	N=100	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)
T=240	N =25	79.2% (0.0, 20.8)	77.5% (0.0, 22.5)	76.6% (0.0, 23.4)	74.3% (2.0, 23.7)	68.1% (0.0, 31.9)	66.1% (0.0, 33.9)
	N =50	27.5% (0.0, 72.5)	27.0% (0.0, 73.0)	27.2% (0.0, 72.8)	4.6% (0.0, 95.4)	3.8% (0.0, 96.2)	3.6% (0.0, 96.4)
	N =100	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)
	N=200	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)
T=480	N =25	90.1% (0.0, 9.9)	89.9% (0.0, 10.1)	90.0% (0.0, 10.0)	90.1% (0.1, 9.8)	87.8% (0.0, 12.2)	87.0% (0.0, 13.0)
	N =50	70.8% (0.0, 29.2)	70.5% (0.0, 29.5)	70.2% (0.0, 29.8)	51.6% (0.0, 48.4)	49.3% (0.0, 80.7)	48.1% (0.0, 51.9)
	N=100	8.4% (0.0, 91.6)	8.5% (0.0, 91.5)	8.4% (0.0, 91.6)	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)
	N=200	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)
	N=400	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)

Table 2: RBIC Rank Estimation of Beta Matrix from Five Empirical Factors

Reported are the percentages (%) of correct estimation by the RBIC estimator from 1,000 simulated datasets. The percentages (%) of under- and overestimation by each estimator are reported in parentheses (•,•). Data are drawn by $x_{it} = \Sigma_{j=1}^{K} \beta_{ij} f_{jt} + \varepsilon_{it}$; $\varepsilon_{it} = \phi(\xi_{1i} h_{1t} + \xi_{2i} h_{2t}) / (\xi_{1i}^2 + \xi_{2i}^2)^{1/2} + (1-\phi^2)^{1/2} v_{it}$, where K = 5, $\phi = 0.2$, and f_{jt} (j = 1, ... , K), $h_{j't}$, $\xi_{j'i}$ ($j' = 1, 2$) and v_{it} are all randomly drawn from N(0,1). For the beta matrix B, we draw an $N \times r$ random matrix B_g such that its first column equals the vector of ones and the entries in the other columns are drawn from N(0,1). We also draw a random K×K positive definite matrix; compute the first r orthonormalized eigenvectors of the matrix; and set a K×r matrix C using the eigenvectors. $\mathrm{B} = \mathrm{B}_g \Lambda^{1/2} C'$, where $\Lambda = diag(\lambda_1, ..., \lambda_r)$. The rank of the beta matrix equals one or three (r = 1, 4). RBIC refers to the Bayesian Information Criterion (BIC) estimator of Cragg and Donald (1997), which is computed under the assumption that the ε_{it} are *i.i.d.* over time (but cross-sectionally correlated).

RBIC		(a) K = 5, r = 1, $\lambda_1 = \lambda$			(b) K = 5, r = 4, $\lambda_i = \lambda$ for i = 1,...,4.		
		λ = .03 R^2=.029	λ = .10 R^2=.091	λ = .50 R^2=.333	λ = .03 R^2=.107	λ = .10 R^2=.286	λ = .50 R^2=.667
T =60	N =25	72.8% (22.3, 4.9)	89.6% (0.10, 10.3)	88.3% (0.0, 11.7)	2.1% (97.9, 0.0)	79.9% (19.8, 0.3)	99.3% (0.0, 0.7)
	N =50	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)	37.3% (0.4, 62.3)	23.7% (0.0, 76.3)	20.6% (0.0, 79.4)
	N =75	78.6% (14.9, 6.5)	86.9% (0.0, 13.1)	85.5% (0.0, 14.5)	1.7% (98.3, 0.0)	87.7% (12.3, 0.0)	99.7% (0.3, 0.0)
	N =100	0.6% (99.4, 0.0)	57.3% (42.7, 0.0)	100% (0.0, 0.0)	0.0% (100, 0.0)	0.2% (99.8, 0.0)	99.8% (0.2, 0.0)
	N=200	0.0% (100, 0.0)	0.0% (100, 0.0)	11.3% (88.7, 0.0)	0.4% (99.6, 0.0)	0.0% (100, 0.0)	0.0% (100, 0.0)
	N=400	0.0% (100, 0.0)	0.0% (100, 0.0)	1.4% (97.6, 1.0)	0.0% (100, 0.0)	0.0% (100, 0.0)	0.0% (74.7, 25.3)
T =120	N =25	65.4% (34.6, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	2.4% (97.6, 0.0)	3.4% (96.6, 0.0)	100% (0.0, 0.0)
	N =50	97.6% (2.1, 0.3)	99.6% (0.0, 0.4)	99.5% (0.0, 0.5)	43.9% (56.1, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)
	N =75	73.6% (0.0, 23.4)	70.7% (0.0, 29.3)	70.6% (0.0, 29.4)	97.9% (1.7, 0.4)	99.1% (0.0, 0.9)	98.9% (0.0, 1.1)
	N =100	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)	28.8% (0.0, 71.2)	25.3% (0.0, 74.7)	23.9% (0.0, 76.1)
	N=200	6.4% (93.6, 0.0)	99.8% (0.2, 0.0)	100% (0.0, 0.0)	0.0% (100, 0.0)	90.5% (9.5, 0.0)	100% (0.0, 0.0)
	N=400	0.0% (100, 0.0)	0.0% (100, 0.0)	97.6% (2.1, 0.3)	0.0% (100, 0.0)	0.0% (100, 0.0)	33.5% (45.6, 20.9)

Table 2 continued...

RBIC		(a) K = 5, r = 1, $\lambda_1 = \lambda$			(b) K = 5, r = 4, $\lambda_i = \lambda$ for i = 1,...,4		
		λ = .03 R^2=.029	λ = .10 R^2=.091	λ = .50 R^2=.333	λ = .03 R^2=.107	λ = .10 R^2=.286	λ = .50 R^2=.667
T =240	N =25	97.7% (2.3, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	33.3% (66.7, 0.0)	99.9% (0.1, 0.0)	100% (0.0, 0.0)
	N =50	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	81.8% (18.2, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)
	N =75	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	98.5% (1.5, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)
	N =100	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)
	N=200	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)	36.7% (0.0, 63.3)	34.9% (0.0, 65.1)	34.5% (0.0, 65.5)
	N=400	93.7% (6.3, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	21.3% (78.7, 0.0)	100% (0.0, 0.0)	96.4% (0.0, 3.6)
T =480	N =25	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	90.0% (10.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)
	N =50	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)
	N =75	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)
	N =100	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)
	N=200	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)
	N=400	0.4% (0.0, 99.6)	0.4% (0.0, 99.6)	0.4% (0.0, 99.6)	53.6% (0.0, 43.7)	55.6% (0.0, 44.4)	55.1% (0.0, 44.9)

Table 3: RBIC Estimation of Beta Matrix from Fifteen Empirical Factors

Reported are the percentages (%) of correct estimation by the RBIC estimator from 1,000 simulated datasets. The percentages (%) of under- and overestimation by each estimator are reported in parentheses (•,•). Data are drawn by $x_{it} = \Sigma_{j=1}^{K} \beta_{ij} f_{jt} + \varepsilon_{it}$; $\varepsilon_{it} = \phi(\xi_{1i} h_{1t} + \xi_{2i} h_{2t}) / (\xi_{1i}^2 + \xi_{2i}^2)^{1/2} + (1-\phi^2)^{1/2} v_{it}$, where K = 15, $\phi = 0.2$, and f_{jt} (j = 1, ... , K), $h_{j't}$, $\xi_{j'i}$ ($j' = 1, 2$) and v_{it} are all randomly drawn from N(0,1). For the beta matrix B, we draw an $N \times r$ random matrix B_g such that its first column equals the vector of ones and the entries in the other columns are drawn from 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. $\mathrm{B} = \mathrm{B}_g \Lambda^{1/2} C'$, where $\Lambda = diag(\lambda_1, ..., \lambda_r)$. The rank of the beta matrix equals one or four (r = 1, 4). RBIC refers to the Bayesian Information Criterion (BIC) estimators of Cragg and Donald (1997), which is computed under the assumption that the ε_{it} are *i.i.d.* over time (but cross-sectionally correlated).

RBIC		(a) K = 15, r = 1, $\lambda_1 = \lambda$			(b) K = 15, r = 4, $\lambda_i = \lambda$ for i = 1,...,4		
		λ = .03 R^2=.029	λ = .10 R^2=.091	λ = .50 R^2=.333	λ = .03 R^2=.107	λ = .10 R^2=.286	λ = .50 R^2=.667
T =60	N =25	18.7% (0.0, 80.9)	9.3% (0.0, 90.7)	6.9% (0.0, 93.1)	24.9% (72.2, 2.9)	59.6% (8.7, 31.7)	40.5% (0.0, 59.5)
	N =50	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)	0.6% (0.0, 99.4)	0.0% (0.0, 100)	0.0% (0.0, 100)
	N =75	12.6% (87.4, 0.0)	82.4% (17.5, 0.1)	99.2% (0.0, 0.8)	0.0% (100, 0.0)	2.2% (97.8, 0.0)	100% (0.0, 0.0)
	N =100	0.0% (100, 0.0)	2.0% (98.0, 0.0)	99.1% (0.9, 0.0)	0.0% (100, 0.0)	0.0% (100, 0.0)	71.1% (28.9, 0.0)
	N=200	0.0% (100, 0.0)	0.0% (100, 0.0)	0.0% (100, 0.0)	0.0% (100, 0.0)	0.0% (100, 0.0)	0.0% (100, 0.0)
	N=400	0.0% (100, 0.0)	0.0% (100, 0.0)	1.3% (96.6, 2.1)	0.0% (100, 0.0)	0.0% (100, 0.0)	0.0% (68.0, 32.0)
T =120	N =25	45.1% (54.9, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	0.2% (99.8, 0.0)	85.7% (14.3, 0.0)	100% (0.0, 0.0)
	N =50	83.6% (1.0, 15.4)	80.7% (0.0, 19.3)	79.2% (0.0, 20.8)	51.4% (48.2, 0.0)	94.0% (0.0, 6.0)	92.3% (0.0, 7.7)
	N =75	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)	3.5% (0.0, 96.5)	1.7% (0.0, 98.3)	1.1% (0.0, 98.9)
	N =100	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)
	N=200	0.1% (99.9, 0.0)	96.1% (3.9, 0.0)	100% (0.0, 0.0)	0.0% (100, 0.0)	34.0% (66.0, 0.0)	100% (0.0, 0.0)
	N=400	0.0% (100, 0.0)	0.0% (100, 0.0)	74.7% (25.1, 0.2)	0.0% (100, 0.0)	0.0% (100, 0.0)	2.2% (91.2, 6.6)

Table 3 continued...

RBIC		(a) K = 15, r = 1, $\lambda_1 = \lambda$			(b) K = 15, r = 4, $\lambda_i = \lambda$ for i = 1,...,4		
		λ = .03 R^2=.029	λ = .10 R^2=.091	λ = .50 R^2=.333	λ = .03 R^2=.107	λ = .10 R^2=.286	λ = .50 R^2=.667
T =240	N =25	80.4% (19.6, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	4.9% (95.1, 0.0)	99.2% (0.8, 0.0)	100% (0.0, 0.0)
	N =50	99.9% (0.1, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	62.7% (37.3, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)
	N =75	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	97.4% (2.6, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)
	N =100	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	99.9% (0.1, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)
	N=200	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)
	N=400	68.6% (31.4, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	98.5% (1.5, 0.0)	100% (0.0, 0.0)	99.6% (0.0, 0.4)
T =480	N =25	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	64.9% (35.1, 0.0)	99.9% (0.1, 0.0)	100% (0.0, 0.0)
	N =50	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	98.4% (1.6, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)
	N =75	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)
	N =100	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)
	N=200	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)
	N=400	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)

Table 4: RBIC Rank Estimation of Demeaned Beta Matrix

Reported are the percentages (%) of correct estimation by the RBICD estimators from 1,000 simulated datasets. The percentages (%) of under- and overestimation by each estimator are reported in parentheses (•,•). Data are drawn by $x_{it} = \Sigma_{j=1}^{K} \beta_{ij} f_{jt} + \varepsilon_{it}$; $\varepsilon_{it} = \phi(\xi_{1i} h_{1t} + \xi_{2i} h_{2t}) / (\xi_{1i}^2 + \xi_{2i}^2)^{1/2} + (1-\phi^2)^{1/2} v_{it}$, where $\phi = 0.2$, and f_{jt} (j = 1, ... , K), $h_{j't}$, $\xi_{j'i}$ ($j' = 1, 2$) and v_{it} are all randomly drawn from N(0,1). For the beta matrix B, we draw an $N \times r$ random matrix B_g such that its first column equals the vector of ones and the entries in the other columns are drawn from N(0,1). We also draw a random K×K positive definite matrix; compute the first r orthonormalized eigenvectors of the matrix; and set a K×r matrix C using the eigenvectors. $\mathrm{B} = \mathrm{B}_g \Lambda^{1/2} C'$, where $\Lambda = diag(\lambda_1, ..., \lambda_r)$. The rank of the beta matrix equals four (r = 4). RBICD refers to the Bayesian Information Criterion (BIC) estimators of Cragg and Donald (1997) for the demeaned beta matrix, which is computed under the assumption that the ε_{it} are *i.i.d.* over time.

RBICD		(a) K = 5, r = 4, $\lambda_i = \lambda$ for i = 1,...,4			(b) K = 15, r = 4, $\lambda_i = \lambda$ for i = 1,...,4		
		λ = .03 R^2=.107	λ = .10 R^2=.286	λ = .50 R^2=.667	λ = .03 R^2=.107	λ = .10 R^2=.286	λ = .50 R^2=.667
T =60	N =25	4.0% (96.0, 0.0)	87.9% (11.2, 0.9)	98.7% (0.0, 1.3)	44.7% (47.3, 8.0)	53.1% (3.7, 43.2)	35.4% (0.0, 64.6)
	N =50	8.0% (0.0, 92.0)	3.7% (0.0, 96.3)	3.2% (0.0, 96.8)	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)
	N =75	23.4% (76.5, 0.1)	96.3% (1.8, 1.9)	96.5% (0.0, 3.5)	0.0% (100, 0.0)	24.1% (75.9, 0.0)	99.6% (0.0, 0.4)
	N =100	0.0% (100, 0.0)	4.7% (95.3, 0.0)	100% (0.0, 0.0)	0.0% (100, 0.0)	0.0% (100, 0.0)	89.0% (11.0, 0.0)
	N=200	0.0% (100, 0.0)	0.0% (100, 0.0)	0.0% (100, 0.0)	0.0% (100, 0.0)	0.0% (100, 0.0)	0.0% (100, 0.0)
	N=400	0.0% (100, 0.0)	0.0% (100, 0.0)	0.0% (74.7, 25.3)	0.0% (100, 0.0)	0.0% (100, 0.0)	0.0% (80.6, 19.4)
T =120	N =25	8.3% (91.7, 0.0)	96.6% (3.4, 0.0)	100% (0.0, 0.0)	0.2% (99.8, 0.0)	89.9% (10.1, 0.0)	100% (0.0, 0.0)
	N =50	60.2% (39.8, 0.0)	100% (0.0, 0.0)	94.6% (0.0, 5.4)	67.4% (30.8, 1.8)	92.0% (0.0, 8.0)	90.8% (0.0, 9.2)
	N =75	96.2% (0.5, 3.3)	95.7% (0.0, 4.3)	94.6% (0.0, 5.4)	1.5% (0.0, 98.5)	0.8% (0.0, 99.2)	0.5% (0.0, 99.5)
	N =100	3.3% (0.0, 96.7)	2.6% (0.0, 97.4)	2.3% (0.0, 97.7)	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)
	N=200	0.0% (100, 0.0)	97.2% (2.8, 0.0)	100% (0.0, 0.0)	0.0% (100, 0.0)	62.7% (37.3, 0.0)	100% (0.0, 0.0)
	N=400	0.0% (100, 0.0)	0.0% (100, 0.0)	33.5% (45.6, 20.9)	0.0% (100, 0.0)	0.0% (100, 0.0)	13.1% (83.8, 3.1)

Table 4 continued...

RBICD		(a) K = 5, r = 4, $\lambda_i = \lambda$ for i = 1,...,4			(b) K = 15, r = 4, $\lambda_i = \lambda$ for i = 1,...,4		
		λ = .03 R^2=.107	λ = .10 R^2=.286	λ = .50 R^2=.667	λ = .03 R^2=.107	λ = .10 R^2=.286	λ = .50 R^2=.667
T =240	N =25	46.1% (53.9, 0.0)	99.9% (0.1, 0.0)	100% (0.0, 0.0)	11.7% (88.3, 0.0)	99.5% (0.5, 0.0)	100% (0.0, 0.0)
	N =50	88.1% (11.9, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	73.9% (26.1, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)
	N =75	98.9% (1.1, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	98.2% (1.8, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)
	N =100	99.9% (0.1, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)
	N=200	8.5% (0.0, 91.5)	7.6% (0.0, 92.4)	7.2% (0.0, 92.8)	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)
	N=400	21.3% (78.7, 0.0)	100% (0.0, 0.0)	96.4% (0.0, 3.6)	8.6% (91.4, 0.0)	100% (0.0, 0.0)	99.7% (0.0, 0.3)
T =480	N =25	92.9% (7.1, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	73.3% (26.7, 0.0)	99.9% (0.1, 0.0)	100% (0.0, 0.0)
	N =50	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	98.9% (1.1, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)
	N =75	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)
	N =100	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)
	N=200	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)	100% (0.0, 0.0)
	N=400	23.7% (0.0, 76.3)	23.4% (0.0, 73.6)	23.4% (0.0, 73.6)	0.0% (0.0, 100)	0.0% (0.0, 100)	0.0% (0.0, 100)

Table 5: Results from Estimation with Six Different Sets of Monthly Portfolio Returns

Reported are the RBIC estimates of the ranks of beta matrices from six different sets of U.S. stock portfolio returns. The RBICD estimates of the ranks of beta matrices are in parentheses. The individual rows of the table report RBIC and RBICD estimation results and the adjusted *R*-squares from portfolio-by-portfolio time series regressions obtained using different sets of empirical factors (with the numbers of empirical factors used in parentheses). FF3, Carhart, FF5, HXZ, MOM, REV, CRR , and JW, respectively, refer to the three Fama-French factors (FF3: VW, SMB, and HML); the Carhart four factors (Carhart: FF3 and MOM); the five Fama-French factors (FF5: FF3, RMW, and CMA); the Hou-Xue-Zhang four factors (HXZ: VW, SMB, I/A, and ROE); the momentum factor (MOM); the short-term and long-term reversal factors (REV); the five Chen-Roll-Ross macroeconomic factors (CRR: MP, UI, DEI, UTS, and UPR); and the three Jagannathan and Wang factors (JW: VW, UPR, and LAB). The sample period is from January 1972 to December 2011 (T = 480).

(a) Whole sample period during January 1972 - December 2011 (T = 480)

Empirical Factors (K)	25 Size and Book to Market		30 Industrial Portfolios		25 Size and B/M + 30 Industrial		25 Size and Momentum		25 Op. Prof. and Investment		32 Size, Op. Prof. and Inv.	
	$\bar{R}^2$	RBIC (RBICD)	$\bar{R}^2$	RBIC (RBICD)	$\bar{R}^2$	RBIC (RBICD)	$\bar{R}^2$	RBIC (RDBIC)	$\bar{R}^2$	RBIC (RBICD)	$\bar{R}^2$	RBIC (RBICD)
FF3 (3)	91.8%	3 (3)	60.7%	3 (3)	73.3%	3 (3)	82.4%	3 (2)	81.1%	3 (2)	87.6%	3 (3)
Carhart (4)	91.9%	3 (3)	61.4%	3 (3)	73.8%	4 (3)	91.4%	4 (3)	81.3%	3 (3)	87.7%	3 (3)
FF5 (5)	91.8%	4 (4)	61.9%	4 (3)	74.0%	4 (4)	82.5%	3 (2)	83.5%	4 (4)	89.6%	5 (5)
HXZ (4)	89.0%	4 (4)	60.6%	4 (4)	72.1%	4 (4)	84.3%	4 (3)	82.2%	3 (2)	88.2%	4 (4)
MOM and REV (3)	12.2%	1 (1)	7.9%	1 (1)	9.6%	2 (2)	22.2%	2 (2)	8.6%	1 (1)	11.1%	1 (1)
CRR (5)	1.3%	0 (0)	1.1%	0 (0)	1.2%	0 (0)	2.0%	0 (0)	0.7%	0 (0)	1.3%	0 (0)
JW (3)	73.4%	1 (1)	57.1%	1 (1)	63.7%	1 (1)	71.7%	1 (1)	78.1%	1 (1)	74.8%	1 (1)
All together (16)	92.9%	4 (4)	64.0%	5 (4)	75.4%	5 (4)	91.9%	4 (3)	84.2%	4 (4)	90.3%	5 (5)
FF5+ HXZ+MOM (8)	92.3%	4 (4)	64.1%	5 (5)	75.5%	5 (4)	92.0%	4 (3)	84.3%	4 (4)	90.4%	5 (5)

(b) Sample period during January 1972 - December 1991 (T = 240)

Empirical Factors (K)	25 Size and Book to Market		30 Industrial Portfolios		25 Size and B/M + 30 Industrial		25 Size and Momentum		25 Op. Prof. and Investment		32 Size, Op. Prof. and Inv.	
	$\bar{R}^2$	RBIC (RBICD)	$\bar{R}^2$	RBIC (RBICD)	$\bar{R}^2$	RBIC (RBICD)	$\bar{R}^2$	RBIC (RDBIC)	$\bar{R}^2$	RBIC (RBICD)	$\bar{R}^2$	RBIC (RBICD)
FF3 (3)	94.1%	3 (3)	71.7%	3 (2)	81.1%	3 (3)	88.5%	2 (2)	85.8%	3 (2)	90.9%	3 (2)
Carhart (4)	94.2%	3 (3)	72.0%	3 (2)	81.3%	3 (3)	93.8%	3 (3)	85.9%	3 (2)	91.0%	3 (2)
FF5 (5)	93.9%	3 (3)	72.4%	4 (3)	81.4%	4 (4)	88.4%	2 (2)	87.6%	4 (3)	91.9%	4 (4)
HXZ (4)	92.6%	3 (3)	71.8%	3 (2)	80.5%	3 (3)	89.8%	3 (3)	86.4%	3 (2)	91.2%	3 (3)
MOM and REV (3)	15.0%	1 (1)	9.6%	1 (1)	11.9%	1 (1)	19.8%	1 (1)	10.8%	1 (1)	13.7%	1 (1)
CRR (5)	8.3%	0 (0)	6.6%	0 (0)	7.3%	0 (0)	8.6%	0 (0)	6.7%	0 (0)	7.5%	0 (0)
JW (3)	79.6%	1 (1)	68.8%	1 (1)	73.3%	1 (1)	77.6%	1 (1)	83.6%	1 (1)	80.4%	1 (1)
All together (16)	94.2%	3 (3)	74.1%	3 (2)	82.5%	4 (4)	93.9%	3 (3)	87.9%	4 (3)	92.3%	4 (4)
FF5+ HXZ+MOM (8)	94.0%	3 (3)	73.0%	4 (3)	81.8%	4 (4)	93.9%	3 (3)	87.8%	4 (3)	92.2%	4 (4)

(c) Sample period during January 1992 - December 2011 (T = 240)

Empirical Factors (K)	25 Size and Book to Market		30 Industrial Portfolios		25 Size and B/M + 30 Industrial		25 Size and Momentum		25 Op. Prof. and Investment		32 Size, Op. Prof. and Inv.	
	$\bar{R}^2$	RBIC (RBICD)	$\bar{R}^2$	RBIC (RBICD)	$\bar{R}^2$	RBIC (RBICD)	$\bar{R}^2$	RBIC (RDBIC)	$\bar{R}^2$	RBIC (RBICD)	$\bar{R}^2$	RBIC (RBICD)
FF3 (3)	90.0%	3 (3)	54.8%	3 (3)	68.6%	3 (3)	68.6%	3 (2)	77.7%	3 (3)	85.8%	3 (3)
Carhart (4)	90.2%	3 (3)	56.2%	3 (3)	69.5%	4 (3)	69.5%	4 (3)	78.2%	4 (3)	86.0%	4 (3)
FF5 (5)	90.1%	4 (3)	55.4%	4 (3)	69.0%	4 (4)	69.0%	3 (2)	80.3%	3 (3)	87.9%	4 (3)
HXZ (4)	87.0%	4 (3)	53.5%	2 (2)	66.6%	3 (3)	66.6%	4 (3)	78.9%	3 (2)	86.1%	4 (3)
MOM and REV (3)	15.2%	2 (2)	11.5%	1 (1)	12.9%	2 (2)	12.9%	2 (1)	12.6%	2 (2)	14.5%	2 (2)
CRR (5)	0.5%	0 (0)	0.1%	0 (0)	0.3%	0 (0)	0.3%	0 (0)	0.3%	0 (0)	0.6%	0 (0)
JW (3)	67.1%	1 (1)	48.6%	1 (1)	55.8%	1 (1)	55.8%	1 (1)	73.0%	1 (1)	69.5%	1 (1)
All together (16)	91.2%	4 (4)	58.7%	4 (4)	71.4%	5 (4)	71.4%	4 (3)	81.7%	4 (3)	88.9%	5 (4)
FF5+ HXZ+MOM (8)	90.9%	4 (4)	57.7%	4 (4)	70.7%	5 (4)	70.7%	4 (3)	81.4%	4 (3)	88.8%	5 (4)

Table 6: Signal to Noise Ratio of Factors over Different Subsample Periods

Reported are the average R^2 of each factor using different sets of portfolios over the subsample periods, 1972-1991 (First) and 1992-2011 (Second), and the Ratio (Second/First) of the average R^2 in the Second Subsample divided by that in the First Subsample. The factors in this analysis include the three Fama-French factors (FF3: VW, SMB, and HML); the Carhart four factors (Carhart: FF3 and MOM); the five Fama-French factors (FF5: FF3, RMW, and CMA); the Hou-Xue-Zhang four factors (HXZ: VW, SMB, I/A, and ROE); the momentum factor (MOM); the short-term and long-term reversal factors (REV_S and REV_L); the five Chen-Roll-Ross macroeconomic factors (CRR: MP, UI, DEI, UTS, and UPR); and the three Jagannathan and Wang factors (JW: VW, UPR, and LAB). All empirical factors have been orthogonalized with respect to the market portfolio (VW).

	25 Size and Book to Market			30 Industrial Portfolios			25 Size and B/M + 30 Industrial			25 Size and Momentum			25 Op. Prof. and Investment			32 Size, Op. Prof. and Inv.		
Sample	First	Second	Ratio	First	Second	Ratio	First	Second	Ratio	First	Second	Ratio	First	Second	Ratio	First	Second	Ratio
VW	0.80	0.68	0.85	0.69	0.49	0.71	0.74	0.58	0.78	0.79	0.68	0.86	0.83	0.72	0.87	0.81	0.70	0.86
SMB	0.10	0.12	1.28	0.03	0.01	0.52	0.06	0.06	1.09	0.10	0.09	0.99	0.01	0.01	1.40	0.08	0.10	1.29
HML	0.04	0.08	2.19	0.01	0.06	5.80	0.02	0.07	3.09	0.01	0.03	3.63	0.01	0.03	3.10	0.01	0.03	3.40
RMW	0.00	0.05	25.00	0.01	0.04	7.00	0.00	0.05	11.50	0.00	0.03	9.00	0.01	0.03	2.70	0.01	0.03	4.71
CMA	0.00	0.02	10.50	0.00	0.02	5.33	0.00	0.02	6.33	0.00	0.01	NA	0.01	0.03	2.64	0.01	0.02	2.75
I/A	0.01	0.05	3.62	0.01	0.04	7.40	0.01	0.04	4.67	0.00	0.02	8.50	0.01	0.04	2.71	0.01	0.03	3.00
ROE	0.04	0.03	0.70	0.01	0.02	1.42	0.02	0.02	0.91	0.04	0.04	0.95	0.01	0.01	1.18	0.02	0.02	0.91
MOM	0.01	0.01	0.70	0.01	0.02	2.57	0.01	0.01	1.63	0.06	0.09	1.51	0.00	0.01	2.33	0.01	0.01	0.83
REV_L	0.01	0.00	0.10	0.01	0.00	0.40	0.01	0.00	0.29	0.03	0.01	0.20	0.00	0.00	0.33	0.01	0.00	0.25
REV_S	0.04	0.04	1.11	0.01	0.01	0.46	0.02	0.02	0.91	0.03	0.03	0.93	0.01	0.02	1.36	0.02	0.04	1.73
MP	0.00	0.00	NA	0.00	0.00	0.67	0.00	0.00	1.00	0.00	0.00	NA	0.00	0.00	2.00	0.00	0.00	NA
UI	0.00	0.00	1.00	0.00	0.00	2.00	0.00	0.00	1.50	0.00	0.00	2.00	0.00	0.00	2.00	0.00	0.00	1.00
DEI	0.00	0.00	2.00	0.00	0.00	1.50	0.00	0.00	1.50	0.00	0.00	1.00	0.00	0.00	2.00	0.00	0.00	2.00
UTS	0.00	0.00	2.00	0.00	0.00	1.50	0.00	0.00	3.00	0.00	0.00	4.00	0.00	0.00	1.00	0.00	0.00	2.00
UPR	0.00	0.00	1.00	0.00	0.00	0.67	0.00	0.00	0.67	0.00	0.01	4.00	0.00	0.00	1.00	0.00	0.00	0.50
LAB	0.00	0.00	1.00	0.01	0.00	0.50	0.00	0.00	0.67	0.00	0.00	2.00	0.00	0.00	1.00	0.00	0.00	1.00

Table 7: Results from Estimation with Six Different Sets of Quarterly Portfolio Returns

Reported are the RBIC estimates of the ranks of beta matrices from six different sets of U.S. stock portfolio returns. The RBICD estimates of the ranks of beta matrices are in parentheses. The individual rows of the table report the rank estimation results and the average adjusted *R*-squares from portfolio-by-portfolio time series regressions obtained using different sets of empirical factors (with the numbers of empirical factors used in parentheses). FF3, Carhart, FF5, HXZ, respectively, refer to the three Fama-French factors (FF3: VW, SMB, and HML); the Carhart four factors (Carhart: FF3 and MOM); the five Fama-French factors (FF5: FF3, RMW, and CMA); and the Hou-Xue-Zhang four factors (HXZ: VW, SMB, I/A, and ROE). The additional models we consider and their empirical factors are the CCAPM with CG; the Lettau-Ludvigson (LL) model with CAY, CG, and CAY×CG; the Yogo (Yogo) model with VW, DCG, and NDCG; the Santos-Veronesi (SV) model with VW and VW×LC; and the Li-Vassalou-Xing (LVX) model with DHH, DCORP, and DHCORP. The sample period is from the first quarter of 1972 to the fourth quarter of 2011 (T = 160).

Empirical Factors (K)	25 Size and Book to Market		30 Industrial Portfolios		25 Size and B/M + 30 Industrial		25 Size and Momentum		25 Op. Prof. and Investment		32 Size, Op. Prof. and Inv.	
	$\bar{R}^2$	RBIC (RBICD)	$\bar{R}^2$	RBIC (RBICD)	$\bar{R}^2$	RBIC (RBICD)	$\bar{R}^2$	RBIC (RBICD)	$\bar{R}^2$	RBIC (RBICD)	$\bar{R}^2$	RBIC (RBICD)
FF3 (3)	93.0%	3 (3)	63.7%	3 (3)	76.2%	3 (3)	85.5%	3 (3)	83.1%	3 (3)	89.2%	3 (3)
Carhart (4)	93.0%	3 (3)	64.0%	3 (3)	76.4%	3 (3)	92.3%	4 (4)	83.0%	3 (3)	89.3%	3 (3)
FF5 (5)	93.1%	4 (4)	65.2%	4 (4)	77.2%	4 (4)	85.3%	3 (3)	85.5%	5 (4)	91.4%	5 (4)
HXZ (4)	90.4%	4 (4)	64.6%	4 (4)	75.6%	4 (4)	87.5%	4 (3)	84.2%	3 (3)	90.0%	4 (4)
FF5+ HXZ+MOM (8)	93.6%	5 (4)	66.5%	5 (4)	78.1%	5 (5)	93.2%	5 (4)	85.8%	5 (4)	91.8%	5 (4)
CCAPM (1)	1.3%	0 (0)	0.9%	0 (0)	1.1%	0 (0)	1.2%	0 (0)	1.8%	1 (0)	1.7%	1 (0)
LL (3)	0.8%	1 (0)	1.2%	1 (0)	1.1%	1 (0)	0.9%	1 (0)	1.6%	1 (0)	1.4%	1 (0)
Yogo (3)	75.8%	1 (1)	59.8%	1 (1)	66.6%	1 (1)	74.0%	1 (1)	80.3%	1 (1)	77.6%	1 (1)
SV (2)	76.3%	2 (1)	60.9%	2 (2)	67.5%	2 (1)	74.7%	2 (1)	80.9%	2 (1)	78.0%	2 (1)
LVX (3)	7.0%	1 (0)	3.2%	1 (0)	4.8%	1 (0)	7.0%	1 (0)	2.9%	1 (0)	5.9%	1 (0)
All together (18)	93.8%	5 (4)	69.1%	6 (4)	79.6%	6 (6)	93.5%	5 (4)	86.2%	5 (3)	92.1%	5 (4)

Table 8: Results from Estimation with Different Sets of Monthly Individual Stock Returns

Reported are the average RBIC and RBICD estimates over 10,000 random samples of the ranks of beta matrices from individual U.S. stock returns. The standard deviations from the estimates are in parentheses. The individual rows of the table report RBIC and RBICD estimation results and the adjusted *R*-squares from asset-by-asset time series regressions. We report in parenthesis the standard deviation of the estimators over 10,000 different samples. The results are obtained using different sets of empirical factors (with the numbers of empirical factors used in parentheses). FF3, Carhart, FF5, HXZ, MOM, REV, CRR , and JW, respectively, refer to the three Fama-French factors (FF3: VW, SMB, and HML); the Carhart four factors (Carhart: FF3 and MOM); the five Fama-French factors (FF5: FF3, RMW, and CMA); the Hou-Xue-Zhang four factors (HXZ: VW, SMB, I/A, and ROE); the momentum factor (MOM); the short-term and long-term reversal factors (REV); the five Chen-Roll-Ross macroeconomic factors (CRR: MP, UI, DEI, UTS, and UPR); and the three Jagannathan and Wang factors (JW: VW, UPR, and LAB).

(a) Results from 10,000 random samples of 50 stocks

Empirical Factors (*K*)	1972-2011 (T=480, N=50)			1972-1991 (T=240, N=50)			1992-2011 (T=240, N=50)		
	RBIC	RBICD	$\bar{R}^2$	RBIC	RBICD	$\bar{R}^2$	RBIC	RBICD	$\bar{R}^2$
FF3 (3)	2.78 (0.41)	2.21 (0.45)	0.28 (0.014)	2.13 (0.34)	1.90 (0.30)	0.31 (0.018)	2.42 (0.51)	2.25 (0.55)	0.19 (0.019)
Carhart (4)	2.79 (0.40)	2.50 (0.52)	0.29 (0.015)	2.16 (0.37)	1.89 (0.31)	0.32 (0.018)	2.44 (0.51)	2.32 (0.53)	0.20 (0.020)
FF5 (5)	2.93 (0.31)	2.24 (0.46)	0.29 (0.014)	2.55 (0.50)	1.93 (0.28)	0.32 (0.018)	2.45 (0.51)	2.23 (0.57)	0.20 (0.020)
HXZ (4)	2.02 (0.13)	1.84 (0.37)	0.27 (0.014)	2.00 (0.07)	1.67 (0.47)	0.31 (0.018)	1.98 (0.26)	1.65 (0.48)	0.19 (0.018)
MOM and REV (3)	0.33 (0.47)	0.32 (0.46)	0.06 (0.006)	0.01 (0.11)	0.01 (0.11)	0.06 (0.006)	0.48 (0.50)	0.34 (0.47)	0.06 (0.009)
CRR (5)	0.00 (0.00)	0.00 (0.00)	0.01 (0.002)	0.00 (0.00)	0.00 (0.00)	0.03 (0.005)	0.00 (0.00)	0.00 (0.00)	0.01 (0.003)
JW (3)	1.00 (0.00)	1.00 (0.00)	0.24 (0.013)	1.00 (0.00)	0.95 (0.22)	0.26 (0.018)	1.00 (0.00)	0.99 (0.06)	0.14 (0.016)
All together (16)	2.96 (0.37)	2.43 (0.54)	0.31 (0.016)	2.65 (0.48)	1.91 (0.31)	0.33 (0.019)	2.46 (0.52)	2.29 (0.53)	0.22 (0.022)

(b) Results from 10,000 random samples of 200 stocks

Empirical Factors (*K*)	1972-2011 (T=480, N=200)			1972-1991 (T=240, N=200)			1992-2011 (T=240, N=200)		
	RBIC	RBICD	$\bar{R}^2$	RBIC	RBICD	$\bar{R}^2$	RBIC	RBICD	$\bar{R}^2$
FF3 (3)	2.94 (0.23)	2.38 (0.48)	0.28 (0.004)	3.00 (0.00)	3.00 (0.01)	0.31 (0.008)	3.00 (0.00)	3.00 (0.01)	0.19 (0.009)
Carhart (4)	2.99 (0.08)	2.99 (0.11)	0.29 (0.004)	3.98 (0.13)	3.96 (0.20)	0.32 (0.008)	4.00 (0.02)	3.99 (0.05)	0.20 (0.009)
FF5 (5)	2.98 (0.14)	2.45 (0.50)	0.29 (0.004)	4.90 (0.30)	4.86 (0.34)	0.32 (0.008)	4.93 (0.25)	4.89 (0.31)	0.20 (0.009)
HXZ (4)	2.00 (0.01)	2.00 (0.01)	0.27 (0.004)	3.92 (0.26)	3.89 (0.30)	0.31 (0.008)	3.99 (0.04)	3.99 (0.07)	0.19 (0.009)
MOM and REV (3)	0.20 (0.39)	0.18 (0.38)	0.06 (0.003)	2.65 (0.49)	2.58 (0.51)	0.06 (0.003)	2.98 (0.12)	2.97 (0.16)	0.06 (0.004)
CRR (5)	0.00 (0.00)	0.00 (0.00)	0.01 (0.000)	3.17 (0.58)	2.99 (0.58)	0.03 (0.002)	3.32 (0.61)	3.17 (0.61)	0.01 (0.001)
JW (3)	1.00 (0.00)	1.00 (0.00)	0.24 (0.004)	2.07 (0.55)	1.96 (0.55)	0.26 (0.008)	2.35 (0.54)	2.42 (0.54)	0.14 (0.007)
All together (16)	3.58 (0.49)	3.53 (0.50)	0.31 (0.005)	11.40 (0.70)	11.10 (0.71)	0.33 (0.008)	12.06 (0.69)	11.80 (0.70)	0.22 (0.010)