10-2018

Shrinking factor dimension: A reduced-rank approach

Dashan HUANG
Singapore Management University, DASHANHUANG@smu.edu.sg

DOI: https://doi.org/10.2139/ssrn.3205697

Shrinking Factor Dimension: A Reduced-Rank Approach*

Dashan Huang
Singapore Management University

Jiaen Li
Washington University in St. Louis

Guofu Zhou
Washington University in St. Louis

First draft: October 2017

Current version: October 2018

*We are grateful to Ravi Jagannathan, Shrihari Santosh, Ngoc-Khanh Tran, Dacheng Xiu, seminar and conference participants at Indiana University, Peking University, Renmin University of China, Shanghai University of Finance and Economics, Washington University in St. Louis, 2018 Booth-EDHEC-RFS conference on "New Methods for the Cross Section of Returns", and 2018 Conference on Financial Predictability and Data Science, and especially to Svetlana Bryzgalova for insightful comments and suggestions that improved the paper substantially.

Send correspondence to Guofu Zhou, Olin School of Business, Washington University in St. Louis, St. Louis, MO 63130; e-mail: zhou@wustl.edu; phone: 314-935-6384.

Shrinking Factor Dimension: A Reduced-Rank Approach

Abstract

We propose a reduced-rank approach (RRA) to reduce a large number of factors to a few parsimonious ones. In contrast to PCA and PLS, the RRA factors are designed to explain the cross section of stock returns, not to maximize factor variations or factor covariances with returns. Out of 70 factor proxies, we find that five RRA factors outperform the Fama-French (2015) five factors for pricing target portfolios, but performs similarly for pricing individual stocks. Our results suggest that existing factor proxies do not provide enough new information at the stock level beyond the Fama-French (2015) five factors.

JEL Classification: G1, G11, G12, G17

Keywords: Reduced rank, PCA, PLS, factors, factor model, cross section

1 Introduction

Explaining why different stocks have different returns is one fundamental question in finance and has received enormous attention over the years. While theories usually identify a few parsimonious factors, there are more than 300 potential factors shown to affect stock returns one way or the other (Harvey, Liu, and Zhu, 2016; Hou, Xue, and Zhang, 2018). Following Cochrane (2011), one can ask two related questions. First, how many factors do we really need based on the existing ones? Second, given a set of well known factors, such as the prominent five factors in Fama and French (FF, 2015), are there other factors that can provide incremental information for explaining the cross section of stock returns?

In this paper, we provide a reduced-rank approach (RRA) that addresses the two questions. Out of 70 factor proxies that we use below including the market factor, we are interested in finding a few out of them. For example, if we restrict to only one factor, the RRA will find a linear combination of the 70 factor proxies that best explains the cross section of stock returns. Interestingly, under the one factor restriction, the factor found by the RRA is almost identical to the market factor (with a correlation of 0.98), indicating that the market factor is the most important one among the 70 proxies. In contrast, the popular principal component analysis (PCA) and partial least squares (PLS) do not do so because they are designed to maximize the variations of the factors and the covariances between the factors and returns, respectively.

Under a five-factor restriction, the five factors chosen by the RRA outperform the FF five factors for pricing the FF (1997) 48 industry portfolios, which are the target portfolios used to select the factors.[1] However, the RRA factors can only perform similarly as the FF five factors when applied to price all individual stocks. The same conclusion is true with alternative target portfolios. Note that this conclusion is unlikely driven by the methodology as the RRA is statistically designed to pick up the best factors to explain the returns. Hence, we interpret our results as evidence supporting the view that the 70 factor proxies do not have enough new information at the stock level beyond the FF five factors.

In the statistical learning literature, the RRA is a dimension reduction tool. It imposes a rank restriction on regression coefficients, so that a lower rank restriction can be effectively used to reduce a large number

[1]Similarly, the popular 25 size and book/market portfolios can be regarded as the target portfolios that are used to select/design the FF (1993) three factors, which are used subsequently to price all stocks.

of regressors/factors into a small number of their composites/linear combinations. Anderson (1951) appears the first such a study in statistics. Reinsel and Velu (1998) provide a book-level analysis on its properties and applications. In finance, Velu and Zhou (1999) apply the RRA to test multi-beta asset pricing models, and Zhou (1994) extends it to the generalized method of moment (GMM) framework of Hansen (1982). Based on Zhou (1994), this paper develops a reduced-rank approach to shrinking factor dimension.

The RRA framework provides a simple GMM test for whether a given set of chosen factors have zero alphas for the target portfolios. In addition, it allows us to impose a given level of mispricing on the target portfolios. This flexibility is interesting since it allows us to price target portfolios at any desired level of accuracy, and then to examine how pricing target portfolios affects pricing individual stocks. Moreover, it also allows us to investigate whether pricing ability on the target portfolios impacts out-of-sample performance for both target portfolios and individual stocks.

This paper also extends the RRA to pre-specify the FF five factors as part of the true factors. Then we ask whether a few composite factors from the remaining 65 proxies can help improve the explanation of the cross section of stock returns. Interestingly, we find that, for both the target portfolios and the cross section of individual stocks, the additional factors add little extra explanatory power, suggesting again not much can be gained out of the factor proxies beyond the FF five factors. In light of Harvey, Liu, and Zhu (2016), it seems there are too many factors in finance. However, our paper indicates that there are too few factors that are useful.

Our paper is related to a growing number of recent studies on factors and firm characteristics. Clarke (2016) identifies level, slope, and curve factors from a few proxies by applying the PCA. Based on firm characteristics, Kelly, Pruitt, and Su (2018) find seven factors that are significant using their instrumented PCA (IPCA). Feng, Giglio, and Xiu (2017) find 14 out of 99 with LASSO, and Freyberger, Neuhierl, and Weber (2018) find significant nonlinear pricing with non-parametric LASSO. Han, He, Rapach, and Zhou (2018) introduce the use of combination forecasts and combination LASSO. Using Bayesian LASSO, Kozak, Nagel, and Santosh (2018) find that the best linear combinations of the proxies in explaining target assets in the stochastic discount factor (SDF) framework and estimate the parameters by numerically solving a dual-penalty problem. In contrast to the PCA methods, our paper extracts factors that are the most useful in explaining the cross section of stock returns. Different from the various LASSO methods, our RRA estimates are analytically done and the asymptotic statistical properties are known.

Another salient feature of the RRA is that it reduces dimension not by reducing the number of factor proxies, but by identifying the dimensionality of the space generated by the target assets. This seems particularly applicable for factor determination. It searches sparsity in the target return space, not sparsity in the factor proxies. In an ideal world where the CAPM is true and a large panel of time series data are available for all stocks, the RRA will recover the CAPM factor, but LASSO cannot because all regressors matter. Theoretically, the RRA is applicable regardless of whether the sparsity is in the factor proxies or in the space generated by them.

The rest of the paper is organized as follows. Section 2 provides the RRA framework. Section 3 extends the RRA to a more general framework allowing for pre-specified or target asset-based factors. Section 4 introduces data and key variables. Section 5 presents empirical results. Section 6 concludes.

2 Methodology

In this section, we provide the RRA and compare it with the PCA and PLS. Then we discuss the performance measures and mispricing restrictions.

2.1 RRA

Following most studies, we assume that target assets that represent the cross section of stock returns are governed by a multi-factor model:

$$R_{it} = \alpha_i + \beta_{i1} f_{1t} + \cdots + \beta_{iK} f_{Kt} + u_{it}, \quad i = 1, \cdots, N;\ t = 1, \cdots, T, \tag{1}$$

where R_{it} is the excess return of asset i in period t, f_{jt} is the realization of the j-th factor in period t $(1 \leq j \leq K)$, u_{it} is the disturbance (i.e., idiosyncratic return) of asset i, K is the number of latent factors, and T is the number of periods.

The case of wide interest is when the true number of factors, K, is typically small, say $K = 5$. The factors are assumed to be related to a number of proxies,

$$f_{kt} = \phi_{k1} g_{1t} + \cdots + \phi_{kL} g_{Lt}, \quad k = 1, \cdots, K, \tag{2}$$

where $g_1, \cdots, g_L$ are L observable variables that can be highly correlated with factor f_k. Typically, L is usually quite large, say $L = 70$. The above equation says that the few true and unknown factors are linear combinations of a set of L observable factor proxies. This assumption is also made for the PCA and PLS, two popular dimension reduction methods, which estimate the combination coefficients differently from our approach. A detailed comparison of the three methods will be provided in Sections 2.2 and 2.3, respectively.

Together with (2), regression (1) can be written in a matrix manner as:

$$R_t = \alpha + \Theta' g_t + U_t, \tag{3}$$

where $R_t = (R_{1t}, \cdots, R_{Nt})'$ is an N-vector of the returns, $g_t = (g_{1t}, \cdots, g_{Lt})'$ is an L-vector of the factor proxies, $U_t = (u_{1t}, \cdots, u_{Nt})'$ is an N-vector of the disturbances, and Θ is an $L \times N$ matrix of the parameters. Equation (2) implies

$$\Theta = \Phi \beta, \tag{4}$$

where Φ is an $L \times K$ matrix of ϕ_{kl} and β is a $K \times N$ matrix of β_{ik}. Then it is clear that

$$H_0: \qquad \operatorname{rank}(\Theta) \leq K. \tag{5}$$

In other words, when the K factors can be expressed as a linear combination of the factor proxies, the rank of the regression coefficients in (3) cannot exceed K. On the other hand, if the regression coefficients have a rank of K, there must be a reduction of L proxies to K factors. Hence, the estimation of Φ and β is to perform a reduced-rank regression in (3), or a regression with a rank restriction on the coefficients.

We solve below analytically by fitting the residual moment condition and therefore cast the problem into the framework of the generalized method of moments (GMM, Hansen, 1982), which allows us to easily obtain the asymptotic distribution of the parameters. Let Z_t be an M-vector of the instruments available at time t. Then, the moment condition is

$$\mathrm{E}[h_t(\alpha,\Theta)] = 0, \qquad h_t(\alpha,\Theta) \equiv U_t(\alpha,\Theta) \otimes Z_t, \tag{6}$$

where $\otimes$ is the Kronecker product that makes h_t an NM-vector function of both the disturbances and the instruments. Let $\boldsymbol{h}_T$ be the sample mean of h_t:

$$\boldsymbol{h}_T(\alpha,\Theta) = \frac{1}{T}\sum_{t=1}^{T} h_t(\alpha,\Theta) = vec(Z'U/T), \tag{7}$$

where Z is a $T \times M$ matrix of the instruments, U is a $T \times N$ matrix of the idiosyncratic returns, and $vec(\cdot)$ is the vectorization operator. Then the GMM estimator solves the following minimization problem:

$$\min_{\alpha,\Theta} Q \equiv \boldsymbol{h}_T(\alpha,\Theta)' W_T \boldsymbol{h}_T(\alpha,\Theta), \tag{8}$$

where W_T is an $NM \times NM$ weighting matrix that is positive definite. The resulting estimator is the GMM estimator of Hansen (1982). In terms of the notation of this paper, $M = L+1$ and

$$Z_t = (1, g_t')'. \tag{9}$$

Since Φ and β enter nonlinearly in the moment condition, the minimization problem has to be solved numerically in general. With nonlinear restrictions, it is very difficult, if not impossible, to find the numerical solution for hundreds of parameters. However, if the weighting matrix is of the following form:

$$W_T \equiv W_1 \otimes W_2, \qquad W_1 : N \times N, \quad W_2 : M \times M, \tag{10}$$

we can analytically solve the problem in two steps. In the first step, conditional on Θ, the estimate of α can be analytically solved as:

$$\hat{\alpha}' = (X'P_0X)^{-1}X'P_0(R - G\Theta), \tag{11}$$

where X is a T-vector of ones, $P_0 = ZW_2Z'$ with Z as a $T \times M$ matrix of the instruments, R is a $T \times N$ matrix of the returns, and G is a $T \times L$ matrix of the factor proxies. The proof, based on Zhou (1994), is provided in Appendix A.

In the second step, to estimate $\Theta = \Phi\beta$, we note that Φ and β are not unique. Given any $K \times K$ non-singular matrix V, the model will be exactly the same as ΦV and $V^{-1}\beta$. In other words, the estimated factors will not be unique, but they differ only up to a rotation. Under a suitable normalization such that the first K columns of β is an identity matrix, the estimates of Φ and β are uniquely given by

$$\hat{\Phi} = (G'PG/T^2)^{-1/2}E, \qquad \hat{\beta} = (G^{*\prime}PG^{*})^{-1}G^{*\prime}PR, \tag{12}$$

where $P = P_0 - P_0X(X'P_0X)^{-1}X'P_0$, $G^* = G\hat{\Phi}$, and E is the $L \times K$ matrix that stacks the 'standardized' eigenvectors ($E'E = I_K$) corresponding to the K largest eigenvalues of the $L \times L$ matrix:

$$(G'PG/T^2)^{-1/2\prime}(G'PR/T^2)W_1(G'PR/T^2)'(G'PG/T^2)^{-1/2}. \tag{13}$$

In summary, the estimators are computed easily in practice. By using the identity weighting matrix, we can compute sequentially from (13) to (12) and to (11), thereby obtaining all the parameter estimates. The estimated factors will then be given by

$$\hat{f}_t = \hat{\Phi}' g_t. \tag{14}$$

In applications, we can standardize $\hat{f}_t$ to make it have an identity covariance matrix.

Theoretically, the estimators are asymptotically consistent, but not necessarily optimal with the minimum covariance matrix. If one is interested in improving the accuracy, it is easy to have a new estimator by using the inverse of

$$\hat{S}_T = \left(\frac{1}{T}\sum \hat{U}_t\hat{U}_t'\right) \otimes \left(\frac{1}{T}\sum Z_tZ_t'\right) = (U'U/T) \otimes (Z'Z/T) \tag{15}$$

as the weighting matrix, where $\hat{U}_t$ is evaluated at the previous estimator with identity weighting matrix. In our applications below, L is large relative to N or the ratio L/N is not small, the inverse of $\hat{S}_T$ is not well behaved. As a result, we simply use only the diagonal elements of $\hat{S}_T$ to obtain a consistent estimator. Of course, the best estimator is obtained by using the optimal weighting matrix of Hansen (1984). But, as pointed out earlier, this is not feasible due to the lack of an analytical solution.

2.2 Comparison with PCA

The principal component analysis (PCA) is a simple and popular dimension reduction tool, which has a long history and has been widely used in all sciences since its introduction by Pearson (1901). It requires a transformation of a set of random variables (factors) into independent principal components, so that the first one has the largest variance, the second one has the second largest, and so on. The PCA uses a few such principal components to represent all the variables (factors). Mathematically, we find $\phi_k^{\text{PCA}} = (\phi_{k1}^{\text{PCA}}, \dots, \phi_{kL}^{\text{PCA}})'$ to solve successively

$$\max_{\phi_k^{\text{PCA}}} \text{Var}(g_t' \phi_k^{\text{PCA}}), \tag{16}$$

such that the later ones are independent from the former. The solution is well known. Empirically, given G, the K eigenvectors, corresponding to the first K largest eigenvalues of the $L \times L$ matrix $G'G$, are the estimates of $\hat{\phi}_1^{\text{PCA}}, \dots, \hat{\phi}_K^{\text{PCA}}$. Then

$$\hat{f}_{kt}^{\text{PCA}} = \hat{\phi}_{k1}^{\text{PCA}} g_{1t} + \cdots + \hat{\phi}_{kL}^{\text{PCA}} g_{Lt}, \quad k = 1, \dots, K, \tag{17}$$

are the PCA factors.

By design, the first K principal components represent the best factors that explain the variations of the L given factor proxies. However, there is no guarantee that they are in any way close to the best factors that explain the cross section of stock returns. Indeed, this is not surprising since no information about the asset returns are used in finding the PCA factors except the factor proxies g_t themselves. In the worst situation, if a factor proxy has the largest variance and little ability to explain stock returns, it will be very likely chosen as the first factor as long as it is uncorrelated with the other factors. Of course, this may not happen in the real data. It just says that one needs to keep in mind that the PCA is designed to explain the factor variations, rather than the returns.

2.3 Comparison with PLS

While the PCA is popular, it does not make use of the target information as it describes g_t simply by a few independent factors of maximal variances. Recognizing this weakness, Wold (1966) introduces the partial least squares (PLS) method to link the factor selection to the target. In our context here, the objective is to search linear combinations of g_t to maximize its covariance with a linear combination of R_t. When $N = 1$, the objective is clearly to maximize the covariance of the extracted factors with R_t. When $N > 1$, it is unclear with which returns the extracted factors should have the maximum covariance. Hence, we choose a linear combination of returns too. Mathematically, the PLS solves

$$\max_{\psi_k^{\text{PLS}}, \phi_k^{\text{PLS}}} \text{Cov}(R_t' \psi_k^{\text{PLS}}, g_t' \phi_k^{\text{PLS}}), \tag{18}$$

where ψ_k^{PLS} and ϕ_k^{PLS} are jointly and successively solved. Following Gu, Kelly, and Xiu (2018), we use the SIMPLS algorithm of de Jong (1993).[2] For our purposes, ϕ_k^{PLS} is what we need. Then the extracted factors are computed similarly as before.

Comparing the PCA and PLS with the RRA, the PCA ignores the target to extract independent factors. In contrast, the PLS uses information of the target to generate independent factors to have maximal covariances with the target returns. However, in asset pricing, we are more interested in explaining the expected returns, rather than the covariances. In this sense, the RRA seems particularly suitable for finance applications. It extracts factors to fit the first moment condition of the model, which is equivalent to finding the best factors to explain the expected returns. In addition, it is a GMM-based approach so that it is flexible in adding instrumental variables and is capable of drawing inferences and testing hypotheses within the popular GMM framework.

[2] We are grateful to Dacheng Xiu for sharing their Python codes with us.

2.4 Performance measures

In this paper, we are interested in how the extracted factors explain individual stock returns. We use two measures to assess the performance of various factor models. The first measure is total adj-R^2 and is defined by Kelly, Pruitt, and Su (2018) as

$$\text{Total adj-}R^2 = 1 - \frac{\sum_{i,t}(R_{it} - \hat{\alpha}_i - \hat{\beta}_{i1}\hat{f}_{1t} - \cdots - \hat{\beta}_{iK}\hat{f}_{Kt})^2 \times \frac{T_i - 1}{T_i - K - 1}}{\sum_{i,t} R_{it}^2}, \tag{19}$$

which is the fraction of return variance explained by the estimated models. Note that the summation on i is over the universe of all stocks and T_i is the number of returns of asset i. We depart from Kelly, Pruitt, and Su (2018) by using the adjusted R^2 so that we can compare the performance of asset pricing models with different number of factors. A subtle aspect on the total adj-R^2, as addressed by Kelly, Pruitt, and Su (2018), is that the denominator is the sum of squared stock returns without demeaning. The reason for this choice is that, at the firm level, the historical mean is so noisy that using the historical mean as the benchmark will lower the bar for "good" pricing performance. Although the PCA, PLS, and RRA are obtained in different ways from the same target assets, the above measure is straightforward to compute from regressions of all excess returns on the PCA, PLS, and RRA factors, respectively.

The second measure is root-mean-squared pricing error (RMSPE),

$$\text{RMSPE} = \frac{1}{N}\sum_{i=1}^{N}\sqrt{\frac{1}{T}\sum_{t}(R_{it} - \hat{\beta}_{i1}\hat{f}_{1t} - \cdots - \hat{\beta}_{iK}\hat{f}_{Kt})^2}, \tag{20}$$

which assesses to what extent the return variations are not attributable to the extracted factors.

2.5 Number of factors

In the RRA framework, given the target returns, one can conduct an analytical GMM test to explore the optimal number of factors. Although the weighting matrix given by (10) is not necessarily optimal under general heteroscedasticity and the usual Hansen over-identification test cannot be directly used, we can compute, based on Zhou (1994), an alternative GMM test as follows,

$$H_z = T\,(M_T \boldsymbol{h}_T)' V_T\,(M_T \boldsymbol{h}_T)\,, \tag{21}$$

where V_T is an $NM \times NM$ diagonal matrix with diagonal elements $(1/v_1, \cdots, 1/v_d, 0, \cdots, 0)$ in which v_j is the j-th largest positive eigenvalue of the $NM \times NM$ semidefinite matrix

$$\Omega_T = [I - D_T(D_T' W_T D_T)^{-1} D_T' W_T] S_T [I - D_T(D_T' W_T D_T)^{-1} D_T' W_T]', \tag{22}$$

M_T is an $NM \times NM$ matrix, of which the ith row is the standardized eigenvector corresponding to the ith largest eigenvalue of Ω_T, D_T is the first order derivative of $\boldsymbol{h}_T$ with respect to α and Θ (the analytical representation of D_T is given in Appendix B), and d is the number of overidentifications. Under the rank K hypothesis, H_z is asymptotically chi-squared distributed with the degree of freedom $d = N(L+1) - q$, where $q = N + LK + KN - K^2$ is the number of parameters.

2.6 RRA with mispricing restrictions

So far asset pricing restrictions have played no roles in choosing the factors. This section shows that how the factors can be chosen conditional on a given degree of mispricing.

For a fixed number of K, say $K = 5$, it is unlikely that the chosen factors have zero alphas for all of the target assets in (1), especially when the number of the target assets. On the other hand, if the chosen factors have large alpahs for some of the target assets, there is no reason to expect them to perform well for pricing individual stocks. Hence, it is of interest to impose a certain constraint on the mispricing of the chosen factors in pricing the target assets.

A simple way is to assume that

$$\alpha = \eta 1_N, \tag{23}$$

where η is a given constant and 1_N is an N-vector of ones. For example, if $\eta = 1\%/12$, then the allowable annual mispricing is 1% across the assets.

However, different assets have different volatilities. It is intuitive that the mispricing is likely greater for more volatile assets. Hence, it seems more sensible to consider a mispricing restriction as the following form,

$$\alpha_i = \eta \sigma_i, \tag{24}$$

where η is a given constant and σ_i is the standard deviation of asset i. For example, if $\eta = 1\%/12$, then the allowable annual mispricing adjusted for volatility is 1% across the assets.

With a mispricing restriction, the estimation and GMM tests are as easy as the standard case in Section 2.1. We let $P = P_0$ and replace R by subtracting α from each of its rows. However, since there are N fewer parameters, the degree of freedom of the chi-squared test becomes $N(L+1) - q_0$, where $q_0 = LK + KN - K^2$. Note that the above GMM is a joint test of a K factor model and a mispricing restriction. This is clearly more stringent than a K factor model. In other words, a greater number of factors may be needed to both explain the returns and to satisfy the asset pricing constraints, than the number of factors merely sufficient for explaining the returns.

3 Extension with Additional Factor Structures

Our paper so far focuses on finding the best few factors from a large given set of L factors. In this section, we extend this analysis to a more general situation, where we have a set of pre-specified or extracted factors already, and ask for a few additional best factors from a given set of proxies.

3.1 Pre-specified factors

Previously, the factors are modeled as linear combinations of proxies, which ignores any pre-specified factors. Based on theory and empirical studies, it is well known that the market factor is one of the most important factors (see, e.g., Harvey and Liu, 2018). Currently, the five factors of FF (2015) are one set of the most studied. It is hence of interest to include the market factor or the five factors or some other factors as the true factors, while searching for additional factors as best linear combination from a large set of proxies.

In this case, we have the following multi-factor model for the asset returns:

$$R_{it} = \alpha_i + \beta_0' F_{0t} + \beta_{i1} f_{1t} + \cdots + \beta_{iK} f_{Kt} + \varepsilon_{it} \quad i = 1, \cdots, N; \;\; t = 1, \cdots, T, \tag{25}$$

where F_{0t} is a K_0-vector of pre-specified factors and β_0 is the factor loading matrix, while the rest are similar as (1). That is, we assume the other factors are related to a number of proxies,

$$f_{kt} = \phi_{k1} g_{1t} + \cdots + \phi_{kL} g_{Lt}, \quad k = 1, \cdots, K, \tag{26}$$

where $g_{1t}, \cdots, g_{Lt}$ are L proxies excluding F_{0t}.

The estimation can be done as easily as before. The only difference is to re-define X in (11) as a $T \times (1 + K_0)$ matrix of ones and pre-specified factors, and expanding α into an $N \times (1 + K_0)$ matrix to include β_0. However, the degree of the GMM test has to be adjusted down by K_0 to reflect the pre-specified K_0 parameters. In the case of using the FF five factors as the the pre-specified factors, we simply let $K_0 = 5$ and estimate the rest from other factor proxies.

It should be pointed out that our procedure is fundamentally different from the case of regressing the returns on the specified factors and then obtaining factors to explain the fitted residuals. This procedure introduces estimation errors in subsequent econometric analysis because the residuals are estimated with errors and the estimation ignores the impact of other factors. In contrast, our procedure provides either the optimal GMM estimator or the efficient maximum likelihood estimator with suitable assumptions on the weighting matrix or return distributions. This is similar to estimating beta pricing models (see, e.g., Shanken and Zhou, 2007), and is the best procedure that estimates all parameters simultaneously, while other procedures estimate one set of parameters irrespective ofp others and is hence subject to estimation errors in addition to the usual sampling variations.

3.2 Target asset-based (TAB) factors

In finding true factors that explain the target asset returns, one view is that the true factors are not functions of some given factor proxies, but rather latent functions of the target asset returns only. In line with this view, we assume that asset returns are now governed by a multi-factor model as follows:

$$R_{it} = \alpha_i + b_{i1} f_{1t}^e + \cdots + b_{iJ} f_{Jt}^e + e_{it}, \quad i = 1, \cdots, N, \tag{27}$$

where R_{it} is the excess return of asset i in period t $(1 \leq i \leq N)$, f_{jt}^e is the realization of the j-th factor in period t $(1 \leq j \leq J)$, e_{it} is the disturbance, J is the number of latent factors, and T is the number of periods.

In contrast with (1), we do not assume (2). Then the J factors need to be extracted from the target asset returns by some statistical procedures with suitable auxiliary assumptions. The common assumption from factor analysis is that the to-be-extracted factors are linear functions of returns,

$$f_{jt}^e = c_{j1} R_{1t} + \cdots + c_{jN} R_{Nt}, \quad j = 1, \cdots, J, \tag{28}$$

where $c_{j1}, \cdots, c_{jN}$ are to be directly estimated from the target asset returns. With this assumption, the model can be written as

$$R_t = \alpha + BC' R_t + e_t, \tag{29}$$

where α is an N-vector of the alphas, B is an $N \times J$ matrix of the factor loadings, C is an $N \times J$ matrix of factor weights on the assets, and e_t is an N-vector of the residuals.

There are many ways to obtain the estimates of C, and hence α and B by running regressions on $f_t^e = C' R_t$. A simple way is to minimize the expected mean-squared errors, $E(e_t' e_t)$. The solution is well known (see, e.g., Bai, 2003; Balvers and Stivers, 2018), and the extracted factors are

$$\hat{f}_t^e = \hat{C}' R_t, \tag{30}$$

where $\hat{C}$ is the first J eigenvectors of the sample covariance matrix of R_t:

$$\hat{\Sigma}_R = (R - 1_T\hat{\mu}')'(R - 1_T\hat{\mu}')/T^2.$$

That is, under the loss function of the expected mean-squared errors, the extracted factors are obtained from the first J principal components of the asset returns. Note that f_t is not unique because any rotation will give rise to new factors. However, α and BC' are unique.

In term of the set-up of this paper, we can estimate α and BC' by using the GMM objective function,

$$\min\ Q_2 \equiv (\boldsymbol{h}_T^e)' W_T \boldsymbol{h}_T^e, \qquad \boldsymbol{h}_T^e = \frac{1}{T}\sum_{t=1}^{T}[(I_N - BC')(R_t - \mu)] \otimes Z_t, \tag{31}$$

where I_N is an identity matrix and μ is the mean of R_t. It can be shown that (see Appendix C) the solution for the extracted factors still have the same form as (30). However, C is now the first J eigenvectors of

$$\hat{\Sigma}_q = W_1^{1/2}(R - 1_T\hat{\mu}')'P_0(R - 1_T\hat{\mu}')W_1^{1/2}/T^2, \qquad P_0 = ZW_2Z',$$

where 1_T is a T-vector of ones. It is clear that the previously extracted factors from returns are only a special case of the GMM extracted factors when $W_1 = I$ and the instrument variable is a constant (i.e., $Z = 1_T$). In other words, the GMM is an extension of the PCA for extracting factors with the use of instruments.

Balvers and Stivers (2018) provide a novel approach to extract factors from returns under a mispricing restriction of the form $\alpha'\alpha = \eta\mu'\mu$ with η as a given constant. They solve the estimator almost analytically by assuming the mean-squared error loss function. For the exact pricing relation $\alpha = 0_N$ with 0_N as an N-vector of zeros, He, Huang, and Zhou (2018) provides a simple analytical expression. Here we extend the latter to the case where instrumental variables are available.

If the exact pricing relation $\alpha = 0_N$ holds, we have $e_t = (I - BQ')R_t$ by taking expectation on both side of (29). This says that μ no longer plays an explicit role in the estimation. As a result, following the earlier derivation, the solution for the extracted factors still has the same form as (30). However, C is the first J eigenvectors of

$$\hat{\Sigma}_{q0} = W_1^{1/2}R'P_0RW_1^{1/2}/T^2, \qquad P_0 = ZW_2Z'.$$

3.3 TAB factors augmented with RRA factors

Now we consider a factor model with both TAB factors and RRA factors. In other words, we are interested in a model in which the first J factors are extracted from target assets and the next K factors are RRA factors obtained from a given set of L observable factor proxies. In short, we extend (1) to include TAB factors.

Mathematically, we have now

$$R_{it} = \alpha_i + b_{i1} f_{1t}^e + \cdots + b_{iJ} f_{Jt}^e + \beta_{i1} f_{1t} + \cdots + \beta_{iK} f_{Kt} + \varepsilon_{it}, \tag{32}$$

where $f_t^e = (f_{1t}^e, \cdots, f_{Jt}^e)'$ are the extracted factors from (30), and $f_t = (f_{1t}, \cdots, f_{Kt})'$ are the extracted factors from (14). In terms of matrix notation, we can write the model as

$$R_t = \alpha + B f_t^e + \beta' f_t + \varepsilon_t. \tag{33}$$

In comparison with (1), there is an extracted factor component, $f_t^e = C' R_t$.

In general, the GMM objective in this case contains interaction terms of C and β, and its analytical solution is intractable. To obtain an analytical solution, we make the assumption that the two components are uncorrelated. This assumption seems intuitive in our context here. We can first identify J factors $C' R_t$, and then identify K factors. For easy identification, we require the later identified K factors are uncorrelated with $C' R_t$. Under zero correlation, we can solve Q, B, and α as before. Then Φ and β are obtained with the same formula as (12), with R replaced by

$$Y = (I - \hat{B}\hat{C}')(R - 1_T \hat{\mu}'),$$

and with the proxies g_t being de-meaned. The intuition is easily seen from the simple case of minimizing $E(\varepsilon_t' \varepsilon_t)$, where ε_t is given in (32) and can be rewritten as

$$\varepsilon_t = (I_N - \hat{B}\hat{C}')(R_t - \mu') - \beta'(f_t - \mu_f).$$

Under the zero correlation assumption, $E(\varepsilon_t' \varepsilon_t)$ has only two terms: $(R_t - \mu')$ and $(f_t - \mu_f)$. As a result, their estimation can be done sequentially and analytically.

4 Data

4.1 Target assets

We explore two sets of target assets to proxy for the cross section of stock returns. The first set of target assets consists of the FF (1997) 48 industry portfolios, and the second set consists of 202 characteristic portfolios used by Giglio and Xiu (2018): 25 portfolios sorted by size and book-to-market ratio, 17 industry portfolios, 25 portfolios sorted by operating profitability and investment, 25 portfolios sorted by size and variance, 35 portfolios sorted by size and net issuance, 25 portfolios sorted by size and accruals, 25 portfolios sorted by size and beta, and 25 portfolio sorted by size and momentum.

The reason to proxy for the cross section of stock returns with portfolios is that portfolios can efficiently reduce idiosyncratic noises in individual stock returns. All portfolios are valued-weighted. The sample period is from January 1974 to December 2016 (516 months).

4.2 Factor proxies

We consider 70 factor proxies, which include the FF (2015) five factors, momentum factor, Pastor and Stambaugh (2003) liquidity factor, Hou, Xue, and Zhang (2015) ROE factor, and the value-weighted decile spread portfolios of 62 anomalies that have significant CAPM alpha (the construction of these anomalies are detailed in Appendix D). We independently replicate about 120 anomalies that are examined by Green, Hand, and Zhang (2017) and Hou, Xue, and Zhang (2018) and restrict the data to the 1974–2016 sample period. For each anomaly, we only consider the holding period of one month. Table 1 reports the average returns and CAPM alphas of the 70 factor proxies. Among these proxies, there are only two pairs of portfolios having a correlation higher than 0.90, between Ssgrow and Egr and between Roaq and Roeq. Untabulated results show that including or excluding those portfolios with correlation higher than 0.9 or 0.8 does not affect the pricing power of different factor models.

4.3 Testing assets

We consider four sets of testing assets to evaluate the pricing performance of different factor models. The first two sets of testing assets are the two sets of target assets: 48 industry portfolios and 202 characteristic portfolios. The third set is the universe of all domestic common stocks listed on the NYSE, Amex, and Nasdaq exchanges (i.e., stocks that have a CRSP share code of 10 or 11), and the third set is all-but-micro stocks, stocks that are larger than the NYSE 20th percentile based on market equity at the end of June each year. Both FF (2015) and Hou, Xue, and Zhang (2015) find that it is micro stocks that plague the failure of existing factor models. If a stock is delisted with missing delisting return, we assume a return of −30% as Shumway (1997). Finally, if a stock is included in the test if it has at least 24 month returns.

5 Empirical Results

In this section, we present all the empirical results and show two main findings: 1) to price all individual stocks, the FF five factors perform similarly as any five composite factors based on 70 factor proxies that include the FF five factors, and 2) including more factors based on existing factor proxies cannot significantly improve the pricing performance relative to the FF five factors.

5.1 In-sample performance

Table 2 reports the total adj-R^2s and root-mean-squared pricing errors, defined in (19) and (20), of different factor models in explaining the testing assets.[3] In this table, the FF 48 industry portfolios are used as the target assets to represent the cross section of stock returns. As benchmarks, we consider a series of factor models and refer them as FF for brevity, where one factor refers to the market factor (i.e., CAPM), three factors to FF (1993), five factors to FF (2015), and six factors to FF (2018) (i.e., five factors plus the momentum factor), respectively. The PCA K factors are the K principal components corresponding to the largest K eigenvalues of the covariance matrix of the 70 factor proxies. The PLS and RRA factors are those that are obtained from the 70 factor proxies with the goal of maximizing the factor-return covariance or

[3] Both Matlab and Python codes for computing the major results of Table 2 are downloadable

explaining the target of the 48 industry portfolios, respectively.

Panel A of Table 2 shows that the FF one-, three-, five-, and six-factor models explain 51.39%, 55.57%, 57.77%, and 58.37% of cross-sectional variations in the 48 industry portfolios, respectively. This result suggests that the market factor is the most important one (Harvey and Liu, 2018), and explains half of variations of the 48 industry portfolios, whereas the rest factors jointly explain less than 10% of variations of the returns. Surprisingly, the momentum factor has little incremental power in explaining the target asset returns: the total adj-R^2 increases by less than 1%.

The second row of Panel A shows results on the PCA. Its one-, three-, five-, and six-factor models explain only 16.74%, 20.49%, 29.92%, and 33.14% of variations of the 48 industry portfolios, which are much worse than the corresponding FF factor models. When we include its first ten factors, the total adj-R^2 is only 40.78% and still much smaller than the total adj-R^2 of using the market factor alone. Consistent with our earlier discussion in Section 2.2, the reason for the subpar performance of the PCA is because the PCA factors are designed to explain the variations of the factor proxies, and not to have the ultimate statistical objective—describing the variation of the target returns.

The third row of Panel A shows that the PLS factors outperform the PCA factors, but still underperform the FF factors when $K < 5$. For example, the total adj-R^2s with the PLS one and three factors are 23.42% and 47.19%, which are higher than 16.74% and 20.49% with the PCA one and three factors, but lower than 51.39% and 55.57% with the FF one and three factors, respectively. The reason why the one-factor model with the PCA and PLS substantially underperform the market factor may be due to the fact that they put a lot of weights on all the proxies, so that they are quite different from the market when constrained to have only one factor. Indeed, the PCA and PLS one factors have low correlations with the market factor: 0.51 and 0.62. When $K \geq 5$, the total adj-R^2 with the PLS factor model slightly outperform the FF factor model.

The last row of Panel A shows that, with the target of explaining the average returns of the 48 industry portfolios, the RRA factor model performs the best. The total adj-R^2s with its one-, three-, five-, and six-factor models are 54.28%, 61.04%, 64.63%, and 65.27%, and are all larger than the corresponding values with the FF, PCA and PLS models. Compared with the first factor in the PCA and PLS, the first RRA factor has a correlation of 0.98 with the market factor, and its total adj-R^2 is even abut 3% greater than that of the market factor. To obtain further insight on why the first RRA factor outperforms the PCA and PLS, Figure 1 plots the weights of the first PCA, PLS, and RRA factors on the 70 factor proxies, respectively. For comparison, we normalize the weights of each factor to have an L_2-norm of 1. The loadings of the three factors on the market portfolio is 0.14, 0.42, and 0.94, respectively, explaining why the market has a low correlation with the PCA and PLS factors and high correlation with the RRA factor. In sum, the RRA performs in the way as it is designed to do.

The right-hand side of Panel A shows that the root-mean-squared pricing errors do not decrease too much when we include more factors in describing the 48 industry portfolios. For example, when the number of factors is set to 5, the root-mean-squared pricing errors with the FF, PCA, PLS, and RRA models are 4.10%, 5.49%, 4.08%, and 3.75%, respectively; when including one more factor, the corresponding values are still 4.07%, 5.34%, 3.95%, and 3.70%, suggesting that existing factor proxies do no help much in reducing the pricing error relative to the FF five-factor model.

In Panel A, the testing assets are the same as the target returns, and the RRA factors are extracted from them. This makes the RRA factors easier to outperform the FF factors. Panels B through D of Table 2 present the total adj-R^2s and root-mean-squared pricing errors when the testing assets are different from the target assets. Panel B considers the 202 characteristic portfolios in Giglio and Xiu (2018) as the testing assets. Since most of the portfolios are finer sorts of the FF five factors, the pricing power should tilt toward the FF five-factor model. As expected, the FF three-, five-, and six-factor models perform better than the corresponding PCA, PLS, and RRA factor models. However, the RRA factor models still perform much better than the PCA and PLS, and the performance is close to the FF models, despite it is designed to explain the 48 industry portfolios. For example, when the number of factors is restricted to 5, the total adj-R^2 is 84.45% with the RRA factors and 86.94% with the FF factors.

Panel C provides some of the major results of this paper. When the testing assets are all individual stocks, the FF, PLS, and RRA models perform similarly when they are set to have the same number of factors. For example, when there is one factor, i.e., $K = 1$, the total adj-R^2s are 9.37%, 10.02, and 10.05% with the FF, PLS, and RRA models, respectively. When we extend to $K = 5$, the total adj-R^2s are 14.70%, 14.35%, and 14.39%, respectively. The same conclusion is true for the mean-squared pricing error. In contrast, the PCA factors underperform in all the cases. For example, the total adj-R^2 is 8.88% when $K = 1$ and 12.39% when $K = 5$. Similar as Panels A and B, when K increases from 5 to 6, the total adj-R^2 increases by less than 1% for all models, once again suggesting that a model with more than five factors provides little incremental power in explaining the cross section of individual stock returns relative to a five-factor model.

Panel D repeats Panel C but excludes micro stocks in evaluating the factor models. The reason for this test is that, as argued by both FF (2015) and Hou, Xue, and Zhang (2015), it is micro stocks that plague extant factor models. The results show that while the pricing power of all factor models improve dramatically, their patterns are similar to Panel C. The total adj-R^2s with the FF and RRA factors are very close to each other in all the cases, and both of them significantly outperform that with the PCA factors. The PLS model underperform the FF and RRA models when $K < 5$ and similarly when $K \geq 5$. The root-mean-squared pricing errors are reported in the right-hand side of Table 2 and are consistent with the total adj-R^2s.

Since both the PLS and RRA use the information in the target assets in obtaining the factors, one natural question is whether their pricing performance is sensitive to the choice of the target. To answer this question, we use the 202 characteristic portfolios in Giglio and Xiu (2018) as the target to proxy for the cross section of stock returns and report the total adj-R^2s and root-mean-squared pricing errors in Table 3. For easy comparison, we repeat those results in Table 2 on the FF and PCA factors.

In Panel A of Table 3, when the testing assets are the 48 industry portfolios, which are not the target assets, the pricing performance of the PLS and RRA factors is slightly weaker than that in Table 2. For instance, when we are restricted to $K = 5$, the total adj-R^2s with the PLS and RRA are 57.99% and 60.46%, smaller than that in Table 2 (58.97% and 64.63%). However, the RRA value is close to the total adj-R^2 with the FF five-factor model (57.77%) and is significantly larger than that with the PCA five factors (29.92%). In Panel B, when the testing assets are the target assets, the RRA factors slightly outperform the FF factors, which outperform the PCA and PLS. For example, when $K = 5$, the total adj-R^2s with the FF, PCA, PLS, and RRA are 86.94%, 50.37%, 83.68%, and 88.97%, respectively.

In Panel C, when the testing assets are individual stocks, the FF, PLS, and RRA perform similarly. Specifically, when $K = 1$, the total R^2s with the FF, PLS, and RRA are 9.37%, 9.89%, and 11.62%; and when $K = 5$, the corresponding values are 14.70%, 14.74%, and 15.91%, respectively. Again, the PCA factors perform the worst, with total adj-R^2s 8.88% when $K = 1$ and 12.39% when $K = 5$. In Panel D, when the micro stocks are excluded, the total adj-R^2 pattern is the same as Panel C, implying that micro stocks do not play a role in comparing different factor models. In sum, when the target assets are portfolios, the FF, PLS, and RRA factor models may perform differently; but when the target assets are individual stocks, the FF, PLS, and RRA perform similarly and outperform the PCA factor models.

Summarizing Tables 2 and 3, we conclude that when the testing assets are portfolios, one model may have better pricing power than another; but when the testing assets are the universe of individual stocks, the FF, PCA, PLS, and RRA models perform similarly, although the FF are based on only five factors whereas the PLS and RRA are based on 70 factor proxies. This suggests that, based on existing factor proxies, little is gained beyond using five factors in pricing the cross section of stock returns.

5.2 Out-of-sample performance

In the previous section, we examine the performances of the PCA, PLS, and RRA factors based on the full sample of data. One natural question is how the performances change over time if the factor weights are constructed ex ante. To answer this question, following Kozak, Nagel, and Santosh (2018), we use the first 30-year data to estimate the weights, assign the weights to the rest 13 years, and investigate the pricing power over the 13-year out-of-sample period, from January 2004 to December 2016.[4]

Table 4 provides the out-of-sample total R^2_{OS}s and root-mean-squared pricing errors when the target returns are the 48 industry portfolios. Specifically, in Panel A when the testing assets are the target assets, the RRA factors consistently outperform the FF factors, although the magnitudes are not large. For example, the out-of-sample total R^2_{OS}s with the FF and RRA are 55.28% and 56.61% when $K = 1$, and are 58.58% and 63.72% when $K = 5$. Again, the PCA model performs the worst in all cases. For the PLS, its out-of-sample R^2 increases dramatically from $K = 1$ to 5, and is comparable with that of the FF and RRA models when $K = 5$. Similar to the in-sample results, for all models, the total R^2_{OS}s do not increase significantly when we extend the number of factors from $K = 5$ to 6. These results have two practical implications. First, to

[4]The results do not change quantitatively when the first 20-year data are used for estimation and the rest 23-year data are used for out-of-sample evaluation.

price the 48 industry portfolios, a five-factor model seems adequate. Second, when the number of factors is suitably specified, the weights on the PLS and RRA factors are stable over time.

In Panel B, when the testing assets are not the target assets, but the 202 characteristic portfolios, the FF and RRA perform similarly, with total R^2_{OS}s close to each other, and both of them outperform the PCA and PLS factors. In Panels C and D when the testing assets are all individual stocks and all-but-micro stocks, the results are more similar to each other. For example, in the case of all-but-micro stocks, the five-factor out-of-sample R^2_{OS}s with the FF, PCA, PLS, and RRA are 29.61%, 23.96%, 29.49%, and 31.44%, respectively. Table 5 shows similar results when the target assets are the 202 characteristic portfolios that are used to obtain the PLS and RRA factors.

In summary, both Tables 4 and 5 show that the hypothesis that the 70 factor proxies do not have enough new information relative to the FF factors continues to hold out-of-sample. This out-of-sample finding is also consistent with Kozak, Nagel, and Santosh (2018) who show that the investment value in 50 anomalies extends to an out-of-sample period if the stochastic discount factor is estimated with a Bayesian LASSO method.

5.3 Number of factors

We perform the GMM test of Section 2.5 to explore how many factors we need to explain the cross section of stock returns. Specifically, to explain the average returns of the 48 industry or 202 characteristics portfolios, we examine the suitable number of factors, respectively.

Untabulated results show that the chi-square statistic H_z in (21) is large and rejects a model with up to 10 RRA factors in explaining the average returns of ether the 48 industry portfolios or the 202 characteristic portfolios. Similar results are also obtained if we use the original Hansen's (1982) GMM test by solving (8) for the parameters with aid of our analytical solution as the starting point for numerical iterations.

Our result that the GMM test rejects a model with as many as 10 factors is consistent with Kozak, Nagel, and Santosh (2018) who show that it is difficult to find a sparse stochastic discount factor with a few factors to explain 50 anomalies examined in their paper. However, the GMM result only says that there are more than 10 factors that are needed to fit well the model moment conditions. This is a model assessment from the statistical perspective. Economically, though, our empirical results show that the number of factors beyond 5 improves little on the pricing errors, implying that the number of factors may be taken as a number around 5 for practical purposes. Note that the factors are extracted from the 70 factor proxies, so the fruitful way to reduce pricing error is not to add more existing factor proxies into the model, but to find new factor proxies that contain incremental information.

5.4 Results for RRA factors with mispricing restrictions

This section explores how a mispricing restriction affects the pricing performance of the RRA factors. Without loss of generality, we consider the mispricing constraint as (24) with four scenarios: $\eta = 0, 0.5/12$, $1/12$, and $1.5/12$, suggesting that the mispricing ranges from zero to 1.5/12 times of the standard deviation. Conceptually, if a K-factor model prices the target assets well, a zero-alpha restriction does not affect the pricing performance, and in contrast, a non-zero alpha restriction will reduce the pricing performance.

Table 6 reports the results when the target returns are the 48 industry portfolios. Note that over our sample period, the monthly average volatility of the 48 industry portfolios is 6.5%. Hence, while a constraint $\eta = 0$ implies that we do not allow any mispricing when extracting the RRA factors, a constraint $\eta = 1/12$ suggests that we allow a mispricing as large as 6.5% per year for each asset when extracting the RRA factors. From Table 6, we have several observations. First, imposing a zero-alpha constraint does not dramatically affect the pricing power of the RRA factors in this paper. For example, in Panel A when the testing assets are the target assets (i.e., 48 industry portfolios), in the case of no mispricing ($\eta = 0$), the total adj-R^2s with the one-, three-, and five-factor RRA models are 54.60%, 60.98%, and 64.73%, respectively, which are almost the same as the total adj-R^2s in the last row of Panel A in Table 2 (i.e., 54.28%, 61.04%, and 64.67%), where no restrictions are imposed on the alphas.

Second, the pricing power of the RRA factors is insensitive to the magnitude of mispricing. For example, in Panel A, even we increase the permitted mispricing to $\eta = 1.5/12$, the pricing power of the one-, three-, and five-factor RRA models is only slightly weaker than the case of $\eta = 0$, with total adj-R^2 52.44%, 58.86%, and 62.50%, respectively. Third, the insensitivity of the RRA pricing power with respect to the mispricing restriction continues to hold when the testing assets are different from the target assets. Specifically, Panels B, C, and D show that all the results display the same pattern as in Panel A when the testing assets are the 202 characteristic portfolios, all individual stocks, and all-but-micro stocks, respectively. Moreover, the results are close to those in Table 2 when there are no mispricing constraints.

Fourth, similar to Table 2, a five-factor model seems suitable for pricing all the four sets of testing assets. When the number of factors increases from five to six, both the total adj-R^2s and root-mean-squared pricing errors do not change very much. For example, when $\eta = 0$, the total adj-R^2 increases from 14.43% to 15.15%, and when $\eta = 1.5/12$, the total adj-R^2 increases from 13.99% to 14.70%. The change in root-mean-squared pricing error is more indistinguishable, from 16.44% to 16.25% when $\eta = 0$ and from 16.51% to 16.32% when $\eta = 1.5/12$. Table 7 confirms this finding when the target assets are the 202 characteristic portfolios.

Finally, it is of interest to see the out-of-sample performance of the RRA factor when the factor weights are estimated in-sample and with various mispricing restrictions. Tables 8 and 9 report the out-of-sample total adj-R^2s and root-mean-squared pricing errors. The results show that the pricing performance of the RRA factors with mispricing restrictions is almost the same as the case without mispricing restrictions. For example, in Panel A of Tables 8 when the testing assets are 48 industry portfolios and are also the target assets, the out-of-sample total adj-R^2s for the one-, three-, and five-factor RRA models are 56.60%, 60.43%, 63.84% with zero-alpha restriction, and are 56.61%, 60.61%, and 63.72% without mispricing restriction (i.e., Panel A of Table 4). When the testing assets are all stocks, the corresponding values are 13.26%, 14.65%, and 16.19% with zero-alpha restriction, which are also quantitatively the same as the cases without mispricing restriction (13.37%, 14.88%, and 16.33%).

5.5 Results for pre-specified factors

In this section, we attempt to answer the second question in this paper: given the prominent FF five factors, are there other factors that provide incremental information for explaining the cross section of stock returns?

Table 10 considers the case when the pre-specified factors are the FF five factors and the rest are the PCA, PLS, and RRA factors, respectively. In Panel A where the testing assets are the target assets (i.e., 48 industry portfolios), the total adj-R^2 with the FF five factors is 57.77%. When the FF five factors are augmented with an additional PCA, PLS, and RRA factor, the total adj-R^2s only slightly increase to 59.43%, 60.81%, and 61.54%, respectively. The reductions in the root-mean-squared pricing error are also negligible. With the FF five factors, it is 4.10%; with one more factor, it is 4.03% with the PCA, 3.96% with the PLS, and 3.93% with the RRA. Even when we augment the FF five-factor model with five more factors, the increase in the total adj-R^2 is less than 10% and the reduction in the root-mean-squared pricing error is less than 0.5%.

Although the testing assets are different from the target assets, Panels B, C, and D show similar patterns as Panel A. That is, including more factors to the FF five-factor model cannot substantially increase the pricing power, regardless how the additional factors are extracted. This is especially true at the stock level, suggesting that existing factor proxies do not provide much information relative to the FF five factors. The pattern continues to hold when the target assets are the 202 characteristic portfolios (Table 11). Therefore, in contrast to Harvey, Liu, and Zhu (2016) that it seems there are too many factors in the asset pricing literature, our results suggest that there are too few factors that are useful, and the future research is to identify new factor proxies that contain incremental information relative to the FF five factors.

5.6 Results for TAB and RRA factors

In this section, we ask whether extracting factors from both the target assets and the factor proxies can improve the pricing performance, relative to those extracted from the factor proxies alone. Following (32), we construct a composite K-factor model by extracting J TAB factors from the target assets and $K-J$ RRA factors from the 70 factor proxies. In this way, we can compare the results with the RRA K-factor model in Table 2.

Table 12 shows the results when the 48 industry portfolios are the target assets, from which we extract up to 10 TAB factors. In the first row of Panel A, all the factors are the TAB factors. As the testing assets are the target assets, by default these factors should generate the highest total adj-R^2s and lowest root-mean-squared pricing errors. As expected, the total adj-R^2 is now 67.11% with a three-factor model, and 73.76% with a five-factor model, which are much larger than those with the RRA three- and five-factor models (61.04% and 64.63%), as well as with the FF and PLS models. However, when we augment the J TAB factors with $K-J$ RRA factors, the pricing power improves and is close to the performance with K TAB factors alone.

For example, when $J = 3$ and $K = 5$, a model with three TAB factors and two RRA factors generates a total adj-R^2 of 72.19%, close to the value of 73.76% with the five TAB factors alone.

In Panel B when the testing assets are the 202 characteristic portfolios and different from the target portfolios, the pricing performance with the TAB factors alone is worse than that with the FF and RRA factors. For example, when $K = 5$, the total adj-R^2 with the TAB factors is 80.93%, which are smaller than that using the FF and RRA five factors (86.94% and 84.45%), respectively. When we use a combination of three TAB factors and two RRA factors, the total adj-R^2 is still less than that with five RRA factors, suggesting that the pricing performance of the TAB factors is sensitive to the choice of the testing assets. When the testing assets are not the target assets, the performance is worse than that with the RRA factors.

The results in Panels C and D confirm Panel B that the TAB factors can only improve the pricing performance when the testing assets are the target assets. When the testing assets are different, the performance is similar and even weaker than that with the RRA factors alone, which are extracted from the factor proxies and incorporate the information in the target assets. Table 13 confirms this finding when the target assets are 202 characteristic portfolios.

6 Conclusion

In this paper, we propose a simple reduced-rank approach (RRA) for shrinking factor dimension, a solution to deal with the large number of factors discovered by the empirical literature. In contrast to other dimension reduction tools like the PCA and the PLS, the RRA is designed to explain the cross section of stock returns and is implemented analytically.

We apply the RRA to 70 potential factor proxies, including the FF five factors, momentum factor, Pastor and Stambaugh (2003) liquidity factor, Hou, Xue, and Zhang (2015) ROE factor, and 62 anomalies from Green, Hand, and Zhang (2017) and Hou, Xue, and Zhang (2018). We find that the 70 factor proxies do not provide much new information at the stock level beyond the FF five factors. In addition, we apply an extended RRA to the factor proxies with the FF five factors as pre-specified factors, and find that linear combinations of the remaining factor proxies improve little the performances.

Future research is to identify new factors that can provide independent information beyond the FF five factors. Kelly, Pruitt, and Su (2018) identify such factors using their IPCA in conjunction with firm characteristics. Since the RRA improves the PCA in many contexts, it will be of interest to extend it further along the direction of Kelly, Pruitt, and Su (2018). In addition, it will be of interest to apply the RRA to both international equity markets and to other asset markets such as bonds and currencies, and to extend the RRA framework to conditional asset pricing models.

References

Abarbanell, J. S., Bushee, B. J., 1998. Abnormal returns to a fundamental analysis strategy. The Accounting Review 73, 19–45.

Anderson, C. W., Garcia-Feijoo, L., 2006. Empirical evidence on capital investment, growth options, and security returns. Journal of Finance 61, 171–194.

Anderson, T., 1951. Estimating linear restrictions on regression coefficients for multivariate normal distributions. Annals of Mathematical Statistics 22, 327–351.

Asness, C. S., Porter, R. B., Stevens, R. L., 2000. Predicting stock returns using industry-relative firm characteristics. Working Paper.

Bai, J., 2003. Inferential theory for factor models of large dimensions. Econometrica 71, 135–171.

Balakrishnan, K., Bartov, E., Faurel, L., 2010. Post loss/profit announcement drift. Journal of Accounting and Economics 50, 20 – 41.

Ball, R., Gerakos, J., Linnainmaa, J. T., Nikolaev, V. V., 2015. Deflating profitability. Journal of Financial Economics 117, 225 – 248.

Balvers, R. J., Stivers, A., 2018. Efficient factor selection: Explaining risk and mean returns jointly. Working paper.

Barbee Jr., W. C., Mukherji, S., Raines, G. A., 1996. Do sales-price and debt-equity explain stock returns better than book-market and firm size? Financial Analysts Journal 52, 56–60.

Basu, S., 1977. Investment performance of common stocks in relation to their price-earnings ratios: A test of the efficient market hypothesis. Journal of Finance 32, 663–682.

Belo, F., Lin, X., Vitorino, M. A., 2014. Brand capital and firm value. Review of Economic Dynamics 17, 150 – 169.

Boudoukh, J., Michaely, R., Richardson, M. P., Roberts, M. R., 2007. On the importance of measuring payout yield: Implications for empirical asset pricing. Journal of Finance 62, 877–915.

Bradshaw, M. T., Richardson, S. A., Sloan, R. G., 2006. The relation between corporate financing activities, analysts forecasts and stock returns. Journal of Accounting and Economics 42, 53 – 85.

Brown, D. P., Rowe, B., 2007. The productivity premium in equity returns. Working Paper.

Chan, L. K. C., Lakonishok, J., Sougiannis, T., 2001. The stock market valuation of research and development expenditures. Journal of Finance 56, 2431–2456.

Clarke, C., 2016. The level, slope and curve factor model for stocks. Working paper.

Cochrane, J. H., 2011. Presidential address: Discount rates. Journal of Finance 66, 1047–1108.

Cooper, M. J., Gulen, H., Schill, M. J., 2008. Asset growth and the cross-section of stock returns. Journal of Finance 63, 1609–1651.

Daniel, K., Titman, S., 2006. Market reactions to tangible and intangible information. Journal of Finance 61, 1605–1643.

de Jong, S., 1993. Simpls: An alternative approach to partial least squares regression. Chemometrics and Intelligent Laboratory Systems 18, 251–263.

Dechow, P. M., Sloan, R. G., Soliman, M. T., 2004. Implied equity duration: A new measure of equity risk. Review of Accounting Studies 9, 197–228.

Desai, H., Rajgopal, S., Venkatachalam, M., 2004. Value-glamour and accruals mispricing: One anomaly or two? The Accounting Review 79, 355–385.

Dichev, I. D., 1998. Is the risk of bankruptcy a systematic risk? The Journal of Finance 53, 1131–1147.

Eisfeldt, A. L., Papanikolaou, D., 2013. Organization capital and the cross-section of expected returns. Journal of Finance 68, 1365–1406.

Fairfield, P. M., Whisenant, J. S., Yohn, T. L., 2003. Accrued earnings and growth: Implications for future profitability and market mispricing. The Accounting Review 78, 353–371.

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

Fama, E. F., French, K. R., 1997. Industry costs of equity. Journal of Financial Economics 43, 153–193.

Fama, E. F., French, K. R., 2015. A five-factor asset pricing model. Journal of Financial Economics 116, 1–22.

Fama, E. F., French, K. R., 2018. Choosing factors. Journal of Financial Economics 128, 234–252.

Feng, G., Giglio, S., Xiu, D., 2017. Taming the factor zoo. Working paper.

Freyberger, J., Neuhierl, A., Weber, M., 2018. Dissecting characteristics nonparametrically. Working paper.

Giglio, S., Xiu, D., 2018. Asset pricing with omitted factors. Working paper.

Green, J., Hand, J. R. M., Zhang, X. F., 2017. The characteristics that provide independent information about average U.S. monthly stock returns. Review of Financial Studies 30, 4389–4436.

Gu, S., Kelly, B. T., Xiu, D., 2018. Empirical asset pricing via machine learning. Working paper.

Han, Y., He, A., Rapach, D. E., Zhou, G., 2018. How many firm characteristics drive US stock returns? Working paper.

Hansen, L. P., 1982. Large sample properties of generalized method of moments estimators. Econometrica 50, 1029–1054.

Harvey, C. R., Liu, Y., 2018. Lucky factors. Working paper.

Harvey, C. R., Liu, Y., Zhu, H., 2016. ... and the cross-section of expected returns. Review of Financial Studies 29, 5–68.

Haugen, R. A., Baker, N. L., 1996. Commonality in the determinants of expected stock returns. Journal of Financial Economics 41, 401 – 439.

He, A., Huang, D., Zhou, G., 2018. Failure of existing factor models in explaining individual stock returns: Evidence from a predictability test. Working paper.

Hirshleifer, D., Hou, K., Teoh, S. H., Zhang, Y., 2004. Do investors overvalue firms with bloated balance sheets? Journal of Accounting and Economics 38, 297 – 331.

Hou, K., Xue, C., Zhang, L., 2015. Digesting anomalies: An investment approach. Review of Financial Studies 28, 650–705.

Hou, K., Xue, C., Zhang, L., 2018. Replicating anomalies. Working paper.

Kelly, B. T., Pruitt, S., Su, Y., 2018. Characteristics are covariances: A unified model of risk and return. Journal of Financial Economics, forthcoming.

Kishore, R., Brandt, M. W., Santa-Clara, P., Venkatachalam, M., 2008. Earnings announcements are full of surprises. Working Paper.

Kozak, S., Nagel, S., Santosh, S., 2018. Shrinking the cross section. Journal of Financial Economics, forthcoming.

Lamont, O., Polk, C., Saaa-Requejo, J., 2001. Financial constraints and stock returns. Review of Financial Studies 14, 529–554.

Lev, B., Nissim, D., 2004. Taxable income, future earnings, and equity values. The Accounting Review 79, 1039–1074.

Litzenberger, R. H., Ramaswamy, K., 1979. The effect of personal taxes and dividends on capital asset prices: Theory and empirical evidence. Journal of Financial Economics 7, 163–195.

Lockwood, L., Prombutr, W., 2010. Sustainable growth and stock returns. Journal of Financial Research 33, 519–538.

Novy-Marx, R., 2011. Operating leverage. Review of Finance 15, 103–134.

Novy-Marx, R., 2013. The other side of value: The gross profitability premium. Journal of Financial Economics 108, 1 – 28.

Ou, J. A., Penman, S. H., 1989. Financial statement analysis and the prediction of stock returns. Journal of Accounting and Economics 11, 295 – 329.

Pastor, L., Stambaugh, R., 2003. Liquidity risk and expected stock returns. Journal of Political Economy 111, 642–685.

Pearson, K., 1901. Liii. on lines and planes of closest fit to systems of points in space. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science 2, 559–572.

Penman, S. H., Richardson, S. A., Tuna, I., 2007. The book-to-price effect in stock returns: Accounting for leverage. Journal of Accounting Research 45, 427–467.

Piotroski, J. D., 2000. Value investing: The use of historical financial statement information to separate winners from losers. Journal of Accounting Research 38, 1–41.

Pontiff, J., Woodgate, A., 2008. Share issuance and cross-sectional returns. Journal of Finance 63, 921–945.

Reinsel, G., Velu, R., 1998. Multivariate reduced rank regression. Lecture Notes in Statistics. New York: Springer.

Richardson, S. A., Sloan, R. G., Soliman, M. T., Tuna, I., 2005. Accrual reliability, earnings persistence and stock prices. Journal of Accounting and Economics 39, 437 – 485.

Shanken, J., Zhou, G., 2007. Estimating and testing beta pricing models: Alternative methods and their performance in simulations. Journal of Financial Economics 84, 40–86.

Shumway, T., 1997. The delisting bias in CRSP data. Journal of Finance 52, 327–340.

Sloan, R. G., 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? The Accounting Review 71, 289–315.

Soliman, M. T., 2008. The use of DuPont analysis by market participants. The Accounting Review 83, 823–853.

Thomas, J. K., Zhang, F. X., 2011. Tax expense momentum. Journal of Accounting Research 49, 791–821.

Thomas, J. K., Zhang, H., 2002. Inventory changes and future returns. Review of Accounting Studies 7, 163–187.

Titman, S., Wei, K. C. J., Xie, F., 2004. Capital investments and stock returns. Journal of Financial and Quantitative Analysis 39, 677–700.

Velu, R., Zhou, G., 1999. Testing multi-beta asset pricing models. Journal of Empirical Finance 6, 219–241.

Wold, H., 1966. Estimation of principal component and related models by iterative least squares. In P. R. Krishnaiah (Ed.), Multivariate analysis, (pp. 391–420). New York: Academic Press.

Zhou, G., 1994. Analytical GMM tests: Asset pricing with time-varying risk premiums. Review of Financial Studies 7, 687–709.

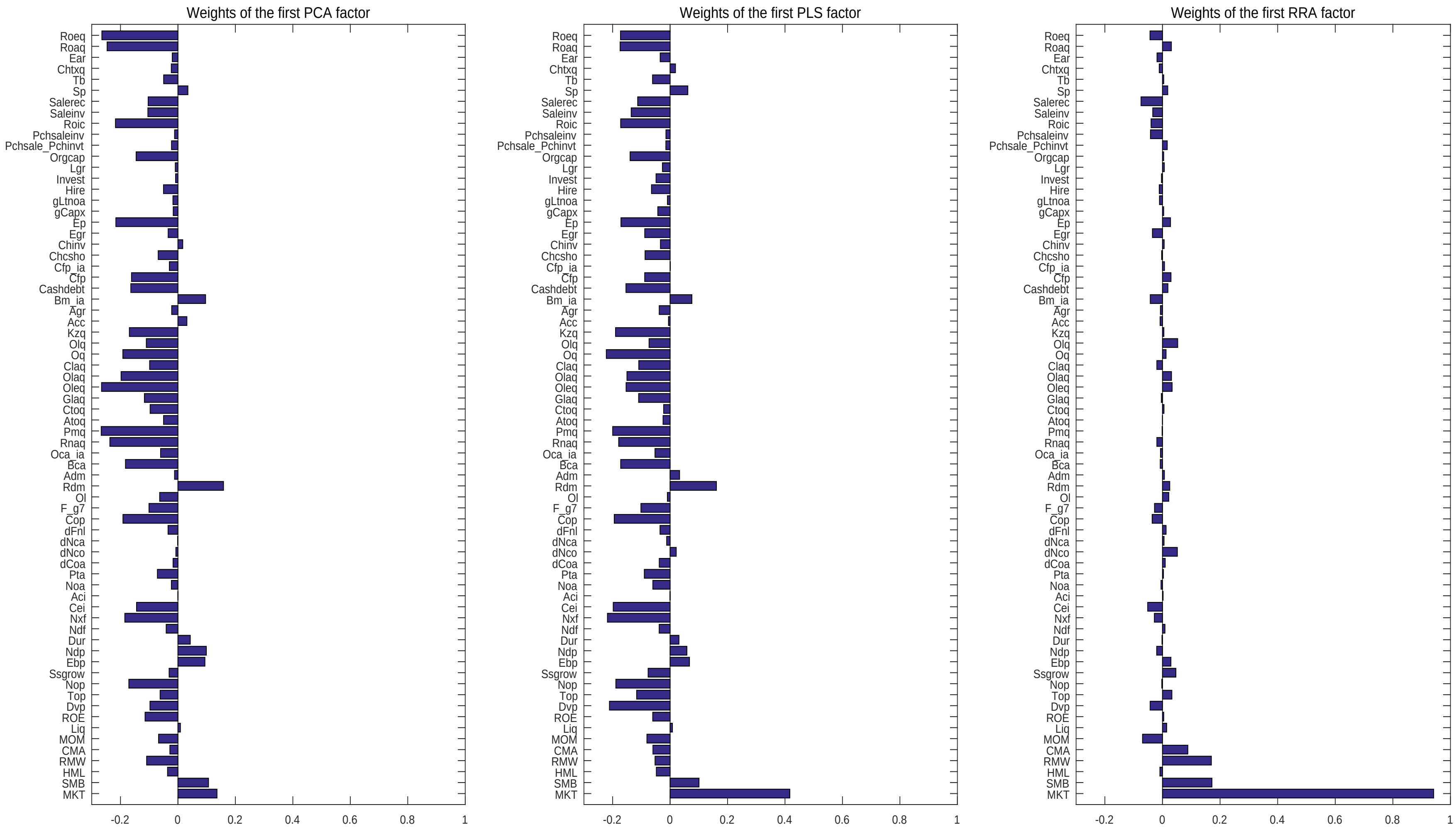

Figure 1 Weights of the first PCA, PLS, and RRA factors

This figure plots the weights of the first PCA, PLS, and RRA factors on the 70 factor proxies, respectively. For comparison, the weights are normalized to have an L_2-norm of 1 for each factor. The target assets in extracting the PLS and RRA factors are FF 48 industry portfolios. The sample period is 1974:01–2016:12.

Table 1 Summary statistics of factor proxies from which factors are extracted

This table reports the average returns, CAPM alphas, and t-values of 70 factor proxies, including the FF (2015) five factors, momentum factor, Pastor and Stambaugh (2003) liquidity factor, Hou, Xue, and Zhang (2015) ROE factor, and value-weighted decile spread portfolios of 62 anomalies that have significant CAPM alpha. The sample period is 1974:01–2016:12.

Factor proxy	Mean	α_{CAPM}	t_{CAPM}	Factor proxy	Mean	α_{CAPM}	t_{CAPM}
MKT	0.59			Ctoq	0.48	0.54	3.31
SMB	0.28	0.20	1.52	Glaq	0.45	0.58	3.28
HML	0.36	0.46	3.68	Oleq	0.66	0.87	3.81
RMW	0.30	0.38	3.79	Olaq	0.71	0.88	4.42
CMA	0.34	0.44	5.48	Claq	0.78	0.91	5.57
MOM	0.60	0.67	3.46	Oq	0.29	0.52	2.75
Liq	0.45	0.45	2.90	Olq	0.62	0.75	4.42
ROE	0.56	0.62	5.49	Kzq	0.22	0.42	2.30
Dvp	0.30	0.60	2.95	Acc	0.48	0.48	3.42
Top	0.44	0.64	3.67	Agr	0.48	0.57	3.56
Nop	0.52	0.79	4.25	Bm_ia	0.52	0.46	2.47
Ssgrow	0.33	0.46	2.94	Cashdebt	0.15	0.33	1.96
Ebp	0.43	0.42	2.17	Cfp	0.57	0.74	3.36
Ndp	0.63	0.58	2.51	Cfp_ia	0.34	0.36	2.40
Dur	0.65	0.69	3.52	Chcsho	0.55	0.70	5.02
Ndf	0.29	0.36	2.86	Chinv	0.44	0.48	3.53
Nxf	0.30	0.58	3.59	Egr	0.43	0.57	3.69
Cei	0.53	0.81	5.03	Ep	0.47	0.75	3.18
Aci	0.34	0.33	2.38	gCapx	0.36	0.45	3.03
Noa	0.53	0.57	4.03	gLtnoa	0.45	0.46	3.17
Pta	0.25	0.38	2.49	Hire	0.27	0.39	2.54
dCoa	0.22	0.30	2.05	Invest	0.51	0.58	4.21
dNco	0.25	0.24	2.34	Lgr	0.21	0.28	2.16
dNca	0.48	0.51	3.73	Orgcap	0.41	0.58	2.60
dFnl	0.33	0.39	3.35	Pchsale_Pchinvt	0.33	0.35	2.53
Cop	0.54	0.76	4.65	Pchsaleinv	0.30	0.30	2.12
F_g7	0.24	0.38	2.68	Roic	0.18	0.38	2.11
Ol	0.35	0.38	2.44	Saleinv	0.23	0.39	2.96
Rdm	0.69	0.52	2.20	Salerec	0.41	0.57	3.71
Adm	0.62	0.66	2.72	Sp	0.58	0.59	2.88
Bca	0.24	0.47	2.21	Tb	0.20	0.28	1.98
Oca_ia	0.60	0.69	5.24	Chtxq	0.53	0.46	2.42
Rnaq	0.48	0.69	3.45	Ear	0.77	0.79	5.39
Pmq	0.47	0.71	3.23	Roaq	0.57	0.78	3.78
Atoq	0.61	0.64	4.07	Roeq	0.60	0.81	3.52

Table 2 Performance of factor models targeted at explaining 48 industry portfolios

This table reports the total adj-R^2s and root-mean-squared pricing errors of different factor models in explaining four sets of testing assets: 48 industry portfolios, 202 characteristic portfolios (Giglio and Xiu, 2018), all stocks, and all-but-micro stocks, respectively. FF refers to the Fama-French model, where 1, 3, 5, and 6 factor(s) are the market factor, FF (1993) three factors, FF (2015) five factors, and FF five factors plus the momentum factor, respectively. PCA, PLS, and RRA refer to the factors based on the principal component analysis, partial least squares, and reduced-rank approach, respectively. The target assets that represent the cross section of stock returns are FF 48 industry portfolios, and the factor proxies are those listed in Table 1. The sample period is 1974:01–2016:12.

	Total adj-R^2 (%)					Root-mean-squared pricing error (%)				
Model	1 factor	3 factors	5 factors	6 factors	10 factors	1 factor	3 factors	5 factors	6 factors	10 factors
Panel A: 48 industry portfolios (i.e., target assets)										
FF	51.39	55.57	57.77	58.34	–	4.46	4.23	4.10	4.07	–
PCA	16.74	20.49	29.92	33.14	40.78	6.01	5.85	5.49	5.34	4.99
PLS	23.42	47.19	58.97	61.11	64.28	5.75	4.70	4.08	3.95	3.78
RRA	54.28	61.04	64.63	65.27	67.38	4.31	3.95	3.75	3.70	3.57
Panel B: 202 characteristic portfolios										
FF	73.31	85.60	86.94	88.30	–	2.97	2.23	2.13	2.04	–
PCA	35.66	39.26	50.34	52.41	59.39	4.68	4.54	4.11	4.02	3.72
PLS	45.50	68.87	79.93	84.41	85.98	4.32	3.28	2.65	2.35	2.24
RRA	77.05	80.56	84.45	85.69	86.84	2.78	2.58	2.33	2.25	2.17
Panel C: All stocks										
FF	9.37	13.64	14.70	15.55	–	17.48	16.79	16.44	16.23	–
PCA	8.88	10.65	12.39	12.82	14.31	17.58	17.11	16.70	16.52	15.84
PLS	10.02	12.97	14.35	15.10	16.31	17.44	16.88	16.47	16.27	15.63
RRA	10.05	12.76	14.39	15.12	16.13	17.40	16.88	16.47	16.27	15.63
Panel D: All-but-micro stocks										
FF	21.07	26.98	28.39	29.25	–	11.80	11.15	10.88	10.73	–
PCA	13.63	17.53	20.81	21.72	24.97	12.27	11.75	11.37	11.23	10.64
PLS	16.35	24.23	27.96	29.47	31.40	12.06	11.36	10.93	10.72	10.23
RRA	22.04	27.46	29.71	30.57	32.14	11.73	11.18	10.83	10.67	10.18

Table 3 Performance of factor models targeted at explaining 202 characteristic portfolios

This table reports the total adj-R^2s and root-mean-squared pricing errors of different factor models in explaining four sets of testing assets: 48 industry portfolios, 202 characteristic portfolios, all stocks, and all-but-micro stocks, respectively. FF refers to the Fama-French model, where 1, 3, 5, and 6 factor(s) are the market factor, FF (1993) three factors, FF (2015) five factors, and FF five factors plus the momentum factor, respectively. PCA, PLS, and RRA refer to the factors based on the principal component analysis, partial least squares, and reduced-rank approach, respectively. The target assets that represent the cross section of stock returns are 202 characteristic portfolios in Giglio and Xiu (2018), including 25 size-B/M portfolios, 17 industry portfolios, 25 operating profitability-investment portfolios, 25 size-variance portfolios, 35 size-net issuance portfolios, 25 size-accruals portfolios, 25 size-beta portfolios, and 25 size-momentum portfolios. The factor proxies are those listed in Table 1. The sample period is 1974:01–2016:12.

	Total adj-R^2 (%)					Root-mean-squared pricing error (%)				
Model	1 factor	3 factors	5 factors	6 factors	10 factors	1 factor	3 factors	5 factors	6 factors	10 factors
Panel A: 48 industry portfolios										
FF	51.39	55.57	57.77	58.34	–	4.46	4.23	4.10	4.07	–
PCA	16.74	20.49	29.92	33.14	40.78	6.01	5.85	5.49	5.34	4.99
PLS	21.60	47.32	57.99	59.77	63.24	5.82	4.69	4.11	4.01	3.80
RRA	51.07	57.97	60.46	62.32	64.22	4.50	4.10	3.97	3.87	3.75
Panel B: 202 characteristic portfolios (i.e., target assets)										
FF	73.31	85.60	86.94	88.30	–	2.97	2.23	2.13	2.04	–
PCA	35.66	39.26	50.34	52.41	59.39	4.68	4.54	4.11	4.02	3.72
PLS	43.42	71.33	83.68	84.49	87.98	4.39	3.14	2.40	2.35	2.09
RRA	79.74	86.92	88.97	89.46	90.35	2.64	2.14	2.00	1.97	1.89
Panel C: All stocks										
FF	9.37	13.64	14.70	15.55	–	17.48	16.79	16.44	16.23	–
PCA	8.88	10.65	12.39	12.82	14.31	17.58	17.11	16.70	16.52	15.84
PLS	9.89	13.04	14.74	15.18	16.66	17.46	16.88	16.45	16.27	15.60
RRA	11.62	14.35	15.91	16.27	17.29	17.25	16.74	16.32	16.16	15.53
Panel D: All-but-micro stocks										
FF	21.07	26.98	28.39	29.25	–	11.80	11.15	10.88	10.73	–
PCA	13.63	17.53	20.81	21.72	24.97	12.27	11.75	11.37	11.23	10.64
PLS	15.83	24.34	28.44	29.22	31.62	12.09	11.36	10.89	10.75	10.21
RRA	23.57	28.01	29.87	30.62	32.50	11.58	11.10	10.78	10.64	10.15

Table 4 Out-of-sample performance of factor models targeted at explaining 48 industry portfolios

This table reports the out-of-sample total adj-R^2_{OS}s and root-mean-squared pricing errors of different factor models in explaining four sets of testing assets: 48 industry portfolios, 202 characteristic portfolios (Giglio and Xiu, 2018), all stocks, and all-but-micro stocks, respectively. FF refers to the Fama-French model, where 1, 3, 5, and 6 factor(s) are the market factor, FF (1993) three factors, FF (2015) five factors, and FF five factors plus the momentum factor, respectively. PCA, PLS, and RRA refer to the factors based on the principal component analysis, partial least squares, and reduced-rank approach, respectively. The target assets that represent the cross section of stock returns are FF 48 industry portfolios, and the factor proxies are those listed in Table 1. In calculating the out-of-sample performance, we use the first 30-year returns to estimate the weights of the PCA, PLS, and RRA factors, apply the weights to the rest 13-year returns, and calculate the out-of-sample total adj-R^2_{OS}s and root-mean-squared pricing errors for each individual asset, with a requirement of at least 24 observations. As such, the out-of-sample evaluation period is 2004:01–2016:12.

	Total adj-R^2_{OS} (%)					Root-mean-squared pricing error (%)				
Model	1 factor	3 factors	5 factors	6 factors	10 factors	1 factor	3 factors	5 factors	6 factors	10 factors
Panel A: 48 industry portfolios (i.e., target assets)										
FF	55.28	57.88	58.58	59.65	–	3.99	3.82	3.75	3.68	–
PCA	31.77	35.48	41.29	41.38	43.49	5.21	5.05	4.83	4.80	4.65
PLS	35.46	47.69	60.79	62.27	64.82	5.06	4.49	3.80	3.73	3.53
RRA	56.61	60.61	63.72	64.42	65.40	4.00	3.79	3.62	3.58	3.48
Panel B: 202 characteristic portfolios										
FF	79.31	86.62	87.32	88.95	–	2.39	1.85	1.78	1.67	–
PCA	50.63	53.07	57.97	57.90	59.74	3.88	3.78	3.60	3.59	3.48
PLS	54.72	68.92	80.27	84.27	85.63	3.70	3.07	2.42	2.22	2.12
RRA	81.55	83.35	85.98	86.88	87.51	2.39	2.27	2.10	2.03	1.98
Panel C: All stocks										
FF	12.66	14.88	15.67	17.00	–	14.14	13.72	13.42	13.22	–
PCA	10.92	12.36	13.67	13.77	14.11	14.35	14.00	13.72	13.61	13.12
PLS	11.18	13.89	15.61	16.28	17.39	14.32	13.85	13.45	13.29	12.82
RRA	13.37	14.88	16.33	16.66	18.13	14.09	13.74	13.40	13.26	12.72
Panel D: All-but-micro stocks										
FF	24.90	28.58	29.61	31.19	–	9.64	9.24	9.02	8.87	–
PCA	17.96	20.98	23.96	24.09	25.49	10.09	9.76	9.48	9.39	9.01
PLS	19.15	25.02	29.49	30.95	32.85	10.01	9.52	9.08	8.93	8.56
RRA	26.10	29.05	31.44	32.10	33.64	9.60	9.27	8.97	8.85	8.45

Table 5 Out-of-sample performance of factor models targeted at explaining 202 characteristic portfolios

This table reports the out-of-sample total adj-R^2_{OS}s and root-mean-squared pricing errors of different factor models in explaining four sets of testing assets: 48 industry portfolios, 202 characteristic portfolios, all stocks, and all-but-micro stocks, respectively. FF refers to the Fama-French model, where 1, 3, 5, and 6 factor(s) are the market factor, FF (1993) three factors, FF (2015) five factors, and FF five factors plus the momentum factor, respectively. PCA, PLS, and RRA refer to the factors based on the principal component analysis, partial least squares, and reduced-rank approach, respectively. The target assets that represent the cross section of stock returns are 202 characteristic portfolios in Giglio and Xiu (2018). In calculating the out-of-sample performance, we use the first 30-year returns to estimate the weights of the PCA, PLS, and RRA factors, apply the weights to the rest 13-year returns, and calculate the out-of-sample total adj-R^2_{OS}s and root-mean-squared pricing errors for each individual asset, with a requirement of at least 36 observations. As such, the out-of-sample evaluation period is 2004:01–2016:12.

	Total adj-R^2_{OS} (%)					Root-mean-squared pricing error (%)				
Model	1 factor	3 factors	5 factors	6 factors	10 factors	1 factor	3 factors	5 factors	6 factors	10 factors
Panel A: 48 industry portfolios										
FF	55.28	57.88	58.58	59.65	–	3.99	3.82	3.75	3.68	–
PCA	31.77	35.48	41.29	41.38	43.49	5.21	5.05	4.83	4.80	4.65
PLS	34.74	47.07	60.99	62.35	64.92	5.08	4.54	3.82	3.71	3.50
RRA	55.69	58.18	60.55	64.51	65.92	4.08	3.88	3.74	3.59	3.46
Panel B: 202 characteristic portfolios (i.e., target assets)										
FF	79.31	86.62	87.32	88.95	–	2.39	1.85	1.78	1.67	–
PCA	50.63	53.07	57.97	57.90	59.74	3.88	3.78	3.60	3.59	3.48
PLS	54.57	69.31	83.72	84.45	87.44	3.71	3.07	2.27	2.23	1.98
RRA	83.71	86.82	88.99	89.70	90.25	2.27	2.02	1.89	1.84	1.77
Panel C: All stocks										
FF	12.66	14.88	15.67	17.00	–	14.14	13.72	13.42	13.22	–
PCA	10.92	12.36	13.67	13.77	14.11	14.35	14.00	13.72	13.61	13.12
PLS	11.30	13.86	16.08	16.43	17.46	14.31	13.86	13.46	13.34	12.82
RRA	13.93	15.33	17.16	17.81	18.50	14.05	13.70	13.36	13.21	12.72
Panel D: All-but-micro stocks										
FF	24.90	28.58	29.61	31.19	–	9.64	9.24	9.02	8.87	–
PCA	17.96	20.98	23.96	24.09	25.49	10.09	9.76	9.48	9.39	9.01
PLS	19.12	25.04	30.23	31.02	33.22	10.01	9.53	9.08	8.97	8.53
RRA	26.78	28.93	31.57	33.13	34.26	9.56	9.25	8.97	8.81	8.45

Table 6 Performance of RRA factors given mispricing constraints and targeted at explaining 48 industry portfolios

This table reports the total adj-R^2s and root-mean-squared pricing errors of RRA factors in explaining four sets of testing assets: 48 industry portfolios, 202 characteristic portfolios (Giglio and Xiu, 2018), all stocks, and all-but-micro stocks, respectively. The RRA factors are extracted by using the reduced-rank approach and are assumed to have mispricing as $\alpha_i = \eta\sigma_i$, where σ_i is asset i's volatility. The target assets that represent the cross section of stock returns are FF 48 industry portfolios, and the factor proxies are those listed in Table 1. The sample period is 1974:01–2016:12.

	Total adj-R^2 (%)					Root-mean-squared pricing error (%)				
$\alpha_i = \eta\sigma_i$	1 factor	3 factors	5 factors	6 factors	10 factors	1 factor	3 factors	5 factors	6 factors	10 factors
Panel A: 48 industry portfolios (i.e., target assets)										
$\eta = 0$	54.60	60.98	64.73	65.33	67.35	4.24	3.88	3.67	3.63	3.49
$\eta = 0.5/12$	54.33	60.69	64.44	65.02	67.02	4.26	3.90	3.69	3.65	3.52
$\eta = 1/12$	53.61	59.98	63.69	64.27	66.26	4.31	3.95	3.75	3.71	3.58
$\eta = 1.5/12$	52.44	58.86	62.50	63.09	65.12	4.40	4.05	3.85	3.81	3.67
Panel B: 202 characteristic portfolios										
$\eta = 0$	76.70	80.92	84.59	85.83	86.86	2.68	2.43	2.19	2.11	2.02
$\eta = 0.5/12$	76.29	80.48	84.04	85.23	86.30	2.72	2.48	2.25	2.17	2.09
$\eta = 1/12$	75.23	79.40	82.89	84.06	85.14	2.81	2.58	2.36	2.29	2.21
$\eta = 1.5/12$	73.70	77.69	81.14	82.31	83.41	2.94	2.73	2.52	2.46	2.38
Panel C: All stocks										
$\eta = 0$	9.83	12.87	14.43	15.15	16.11	17.40	16.85	16.44	16.25	15.61
$\eta = 0.5/12$	9.83	12.82	14.34	15.05	16.06	17.40	16.86	16.45	16.26	15.62
$\eta = 1/12$	9.76	12.70	14.19	14.90	15.95	17.41	16.88	16.47	16.29	15.64
$\eta = 1.5/12$	9.65	12.50	13.99	14.70	15.81	17.43	16.91	16.51	16.32	15.67
Panel D: All-but-micro stocks										
$\eta = 0$	21.68	27.38	29.77	30.61	32.09	11.73	11.16	10.79	10.63	10.15
$\eta = 0.5/12$	21.56	27.23	29.58	30.42	31.94	11.73	11.17	10.80	10.65	10.16
$\eta = 1/12$	21.25	26.93	29.23	30.06	31.62	11.76	11.19	10.84	10.68	10.19
$\eta = 1.5/12$	20.86	26.48	28.72	29.55	31.16	11.80	11.24	10.89	10.73	10.24

Table 7 Performance of RRA factors given mispricing constraints and targeted at explaining 202 characteristic portfolios

This table reports the total adj-R^2s and root-mean-squared pricing errors of RRA factors in explaining four sets of testing assets: 48 industry portfolios, 202 characteristic portfolios (Giglio and Xiu, 2018), all stocks, and all-but-micro stocks, respectively. The RRA factors are extracted by using the reduced-rank approach and are assume to have mispricing as $\alpha_i = \eta\sigma_i$, where σ_i is asset i's volatility. The target assets that represent the cross section of stock returns are 202 characteristic portfolios, and the factor proxies are those listed in Table 1. The sample period is 1974:01–2016:12.

	Total adj-R^2 (%)					Root-mean-squared pricing error (%)				
$\alpha_i = \eta\sigma_i$	1 factor	3 factors	5 factors	6 factors	10 factors	1 factor	3 factors	5 factors	6 factors	10 factors
Panel A: 48 industry portfolios										
$\eta = 0$	51.73	58.02	60.19	62.23	64.21	4.41	4.04	3.92	3.82	3.68
$\eta = 0.5/12$	51.58	57.96	60.13	62.13	64.13	4.42	4.05	3.92	3.82	3.69
$\eta = 1/12$	51.11	57.56	59.85	61.76	63.81	4.45	4.07	3.94	3.85	3.71
$\eta = 1.5/12$	50.47	56.82	59.26	61.07	63.19	4.49	4.12	3.99	3.90	3.76
Panel B: 202 characteristic portfolios (i.e., target assets)										
$\eta = 0$	79.80	86.91	88.95	89.43	90.32	2.50	1.96	1.82	1.78	1.70
$\eta = 0.5/12$	79.65	86.73	88.75	89.24	90.14	2.52	1.99	1.84	1.80	1.72
$\eta = 1/12$	79.01	86.06	88.07	88.57	89.52	2.59	2.06	1.92	1.88	1.80
$\eta = 1.5/12$	77.89	84.90	86.89	87.40	88.45	2.69	2.19	2.06	2.02	1.92
Panel C: All stocks										
$\eta = 0$	11.40	14.40	15.90	16.27	17.30	17.25	16.73	16.32	16.15	15.55
$\eta = 0.5/12$	11.46	14.41	15.90	16.26	17.29	17.24	16.73	16.33	16.16	15.56
$\eta = 1/12$	11.46	14.35	15.84	16.19	17.25	17.24	16.74	16.34	16.17	15.57
$\eta = 1.5/12$	11.34	14.24	15.72	16.08	17.15	17.26	16.76	16.36	16.19	15.59
Panel D: All-but-micro stocks										
$\eta = 0$	23.47	27.96	29.82	30.58	32.50	11.57	11.09	10.77	10.63	10.15
$\eta = 0.5/12$	23.45	27.89	29.74	30.49	32.45	11.57	11.10	10.78	10.64	10.15
$\eta = 1/12$	23.30	27.67	29.56	30.30	32.29	11.58	11.11	10.79	10.66	10.17
$\eta = 1.5/12$	22.95	27.33	29.26	29.99	32.03	11.61	11.14	10.82	10.68	10.19

Table 8 Out-of-sample performance of RRA factors given mispricing constraints and targeted at explaining 48 industry portfolios

This table reports the out-of-sample total adj-R^2s and root-mean-squared pricing errors of RRA factors in explaining four sets of testing assets: 48 industry portfolios, 202 characteristic portfolios (Giglio and Xiu, 2018), all stocks, and all-but-micro stocks, respectively. The RRA factors are extracted by using the reduced-rank approach and are assumed to have mispricing as $\alpha_i = \eta\sigma_i$, where σ_i is asset i's volatility. The target assets that represent the cross section of stock returns are FF 48 industry portfolios, and the factor proxies are those listed in Table 1. In calculating the out-of-sample performance, we use the first 30-year returns to estimate the weights of the RRA factors, apply the weights to the rest 13-year returns, and calculate the out-of-sample total adj-R^2_{OS}s and root-mean-squared pricing errors for each individual asset, with a requirement of at least 24 observations. As such, the out-of-sample evaluation period is 2004:01–2016:12.

	Total adj-R^2_{OS} (%)					Root-mean-squared pricing error (%)				
$\alpha_i = \eta\sigma_i$	1 factor	3 factors	5 factors	6 factors	10 factors	1 factor	3 factors	5 factors	6 factors	10 factors
Panel A: 48 industry portfolios (i.e., target assets)										
$\eta = 0$	56.60	60.43	63.84	64.24	65.50	4.01	3.81	3.62	3.60	3.48
$\eta = 0.5/12$	56.24	60.24	63.58	64.17	65.50	4.01	3.80	3.61	3.58	3.46
$\eta = 1/12$	54.30	58.62	62.26	63.20	64.82	4.11	3.88	3.67	3.62	3.49
$\eta = 1.5/12$	50.98	55.75	60.04	61.48	63.60	4.28	4.04	3.80	3.71	3.55
Panel B: 202 characteristic portfolios										
$\eta = 0$	81.34	83.10	85.74	86.91	87.58	2.41	2.29	2.11	2.06	1.99
$\eta = 0.5/12$	79.90	81.93	84.43	86.34	87.05	2.45	2.31	2.15	2.05	1.99
$\eta = 1/12$	76.19	78.93	81.66	84.55	85.60	2.64	2.48	2.30	2.14	2.06
$\eta = 1.5/12$	70.74	74.46	77.75	81.81	83.48	2.93	2.72	2.53	2.30	2.19
Panel C: All stocks										
$\eta = 0$	13.26	14.65	16.19	16.62	18.16	14.10	13.76	13.40	13.26	12.72
$\eta = 0.5/12$	13.16	14.55	16.13	16.68	18.19	14.11	13.76	13.40	13.25	12.71
$\eta = 1/12$	12.76	14.21	15.89	16.58	18.08	14.15	13.79	13.42	13.25	12.72
$\eta = 1.5/12$	12.10	13.67	15.50	16.35	17.84	14.22	13.85	13.46	13.28	12.75
Panel D: All-but-micro stocks										
$\eta = 0$	26.03	28.62	31.33	32.06	33.71	9.61	9.29	8.98	8.86	8.46
$\eta = 0.5/12$	25.71	28.31	31.06	31.98	33.61	9.62	9.30	8.98	8.85	8.46
$\eta = 1/12$	24.75	27.46	30.40	31.56	33.24	9.69	9.36	9.02	8.87	8.48
$\eta = 1.5/12$	23.24	26.15	29.41	30.88	32.69	9.78	9.44	9.09	8.91	8.51

Table 9 Out-of-sample performance of RRA factors given mispricing constraints and targeted at explaining 202 characteristic portfolios

This table reports the out-of-sample total adj-R^2s and root-mean-squared pricing errors of RRA factors in explaining four sets of testing assets: 48 industry portfolios, 202 characteristic portfolios (Giglio and Xiu, 2018), all stocks, and all-but-micro stocks, respectively. The RRA factors are extracted by using the reduced-rank approach and are assume to have mispricing as $\alpha_i = \eta\sigma_i$, where σ_i is asset i's volatility. The target assets that represent the cross section of stock returns are 202 characteristic portfolios, and the factor proxies are those listed in Table 1. In calculating the out-of-sample performance, we use the first 30-year returns to estimate the weights of the RRA factors, apply the weights to the rest 13-year returns, and calculate the out-of-sample total adj-R^2_{OS}s and root-mean-squared pricing errors for each individual asset, with a requirement of at least 24 observations. As such, the out-of-sample evaluation period is 2004:01–2016:12.

	Total adj-R^2_{OS} (%)					Root-mean-squared pricing error (%)				
$\alpha_i = \eta\sigma_i$	1 factor	3 factors	5 factors	6 factors	10 factors	1 factor	3 factors	5 factors	6 factors	10 factors
Panel A: 48 industry portfolios										
$\eta = 0$	55.67	58.02	59.54	63.98	65.62	4.09	3.89	3.78	3.60	3.46
$\eta = 0.5/12$	55.83	57.86	59.29	63.47	65.47	4.06	3.89	3.79	3.62	3.47
$\eta = 1/12$	54.83	56.73	58.26	62.34	64.84	4.09	3.95	3.85	3.68	3.51
$\eta = 1.5/12$	52.83	54.73	56.53	60.62	63.79	4.19	4.05	3.95	3.77	3.58
Panel B: 202 characteristic portfolios (i.e., target assets)										
$\eta = 0$	83.54	86.80	88.77	89.57	90.23	2.30	2.03	1.90	1.85	1.78
$\eta = 0.5/12$	83.10	86.16	88.24	89.06	89.84	2.27	2.04	1.92	1.88	1.80
$\eta = 1/12$	80.91	84.18	86.47	87.64	88.72	2.38	2.15	2.03	1.98	1.89
$\eta = 1.5/12$	77.22	81.06	83.55	85.36	87.02	2.59	2.35	2.23	2.13	2.03
Panel C: All stocks										
$\eta = 0$	13.81	15.28	17.02	17.71	18.34	14.06	13.71	13.38	13.23	12.73
$\eta = 0.5/12$	13.85	15.30	16.99	17.65	18.27	14.05	13.70	13.38	13.23	12.73
$\eta = 1/12$	13.66	15.14	16.80	17.49	18.13	14.06	13.72	13.40	13.25	12.75
$\eta = 1.5/12$	13.23	14.81	16.47	17.22	17.92	14.11	13.75	13.43	13.27	12.77
Panel D: All-but-micro stocks										
$\eta = 0$	26.67	28.83	31.07	32.99	34.21	9.57	9.26	9.01	8.83	8.46
$\eta = 0.5/12$	26.63	28.75	30.95	32.80	34.08	9.56	9.26	9.01	8.84	8.47
$\eta = 1/12$	26.09	28.29	30.52	32.37	33.76	9.59	9.29	9.04	8.86	8.49
$\eta = 1.5/12$	25.10	27.51	29.82	31.71	33.30	9.65	9.34	9.08	8.91	8.52

Table 10 Performance of factor models with pre-specified FF five factors and targeted at explaining 48 industry portfolios

This table reports the total adj-R^2s and root-mean-squared pricing errors of different factor models in explaining four sets of testing assets: 48 industry portfolios, 202 characteristic portfolios (Giglio and Xiu, 2018), all stocks, and all-but-micro stocks, respectively. FF5 refers to the FF five-factor model, and FF5+PCA, FF5+PLS, and FF5+RRA refer to K-factor models that include FF5 and $K-5$ PCA, PLS, and RRA factors, respectively. The target assets that represent the cross section of stock returns are FF 48 industry portfolios, and the factor proxies are those listed in Table 1. The sample period is 1974:01–2016:12.

	Total adj-R^2 (%)					Root-mean-squared pricing error (%)				
Model	5 factors	6 factors	7 factors	8 factors	10 factors	5 factors	6 factors	7 factors	8 factors	10 factors
Panel A: 48 industry portfolios (i.e., target assets)										
FF5	57.77	–	–	–	–	4.10	–	–	–	–
FF5+PCA	–	59.43	61.09	61.48	62.45	–	4.03	3.94	3.91	3.84
FF5+PLS	–	60.81	61.92	63.01	64.50	–	3.96	3.90	3.82	3.75
FF5+RRA	–	61.54	63.33	64.88	66.16	–	3.93	3.82	3.73	3.66
Panel B: 202 characteristic portfolios										
FF5	86.94	–	–	–	–	2.13	–	–	–	–
FF5+PCA	–	87.62	88.12	88.22	88.58	–	2.10	2.07	2.06	2.03
FF5+PLS	–	87.54	88.34	88.62	89.09	–	2.10	2.05	2.03	1.99
FF5+RRA	–	87.60	88.39	88.86	89.14	–	2.10	2.05	2.01	1.99
Panel C: All stocks										
FF5	14.70	–	–	–	–	16.44	–	–	–	–
FF5+PCA	–	15.41	15.80	16.01	16.41	–	16.24	16.07	15.92	15.61
FF5+PLS	–	15.29	16.00	16.26	16.63	–	16.25	16.05	15.89	15.60
FF5+RRA	–	15.18	15.78	16.03	16.49	–	16.26	16.07	15.91	15.60
Panel D: All-but-micro stocks										
FF5	28.39	–	–	–	–	10.88	–	–	–	–
FF5+PCA	–	29.40	30.18	30.48	31.34	–	10.72	10.57	10.46	10.21
FF5+PLS	–	29.66	30.54	31.18	31.83	–	10.70	10.55	10.42	10.19
FF5+RRA	–	29.76	30.86	31.46	32.05	–	10.70	10.53	10.41	10.18

Table 11 Performance of factor models with pre-specified FF five factors and targeted at explaining 202 characteristic portfolios

This table reports the total adj-R^2s and root-mean-squared pricing errors of different factor models in explaining four sets of testing assets: 48 industry portfolios, 202 characteristic portfolios, all stocks, and all-but-micro stocks, respectively. FF5 refers to the FF five-factor model, and FF5+PCA, FF5+PLS, and FF5+RRA refer to K-factor models that include FF5 and $K-5$ PCA, PLS, and RRA factors, respectively. The target assets that represent the cross section of stock returns are 202 characteristic portfolios in Giglio and Xiu (2018), including 25 size-B/M portfolios, 17 industry portfolios, 25 operating profitability-investment portfolios, 25 size-variance portfolios, 35 size-net issuance portfolios, 25 size-accruals portfolios, 25 size-beta portfolios, and 25 size-momentum portfolios. The factor proxies are those listed in Table 1. The sample period is 1974:01–2016:12.

	Total adj-R^2 (%)					Root-mean-squared pricing error (%)				
Model	5 factors	6 factors	7 factors	8 factors	10 factors	5 factors	6 factors	7 factors	8 factors	10 factors
Panel A: 48 industry portfolios										
FF5	57.77	–	–	–	–	4.10	–	–	–	–
FF5+PCA	57.77	59.43	61.09	61.48	62.45	4.10	4.03	3.94	3.91	3.84
FF5+PLS	57.77	59.05	61.53	62.55	63.58	4.10	4.04	3.92	3.85	3.78
FF5+RRA	57.77	59.44	62.27	63.41	64.38	4.10	4.02	3.88	3.81	3.74
Panel B: 202 characteristic portfolios (i.e., target assets)										
FF5	86.94	–	–	–	–	2.13	–	–	–	–
FF5+PCA	86.94	87.62	88.12	88.22	88.58	2.13	2.10	2.07	2.06	2.03
FF5+PLS	86.94	87.86	88.46	88.87	89.61	2.13	2.08	2.04	2.01	1.95
FF5+RRA	86.94	88.33	89.07	89.63	90.19	2.13	2.05	2.00	1.96	1.91
Panel C: All stocks										
FF5	14.70	–	–	–	–	16.44	–	–	–	–
FF5+PCA	14.70	15.41	15.80	16.01	16.41	16.44	16.24	16.07	15.92	15.61
FF5+PLS	14.70	15.52	15.99	16.30	16.74	16.44	16.23	16.05	15.90	15.58
FF5+RRA	14.70	15.54	16.09	16.50	17.10	16.44	16.23	16.04	15.87	15.55
Panel D: All-but-micro stocks										
FF5	28.39	–	–	–	–	10.88	–	–	–	–
FF5+PCA	28.39	29.40	30.18	30.48	31.34	10.88	10.72	10.57	10.46	10.21
FF5+PLS	28.39	29.37	30.41	30.96	31.90	10.88	10.72	10.56	10.43	10.18
FF5+RRA	28.39	29.61	30.81	31.32	32.27	10.88	10.71	10.54	10.41	10.16

Table 12 Performance of TAB and RRA factors targeted at explaining 48 industry portfolios

This table reports the total adj-R^2s and root-mean-squared pricing errors of different factor models in explaining four sets of testing assets: 48 industry portfolios (targeted assets), 202 characteristic portfolios (Giglio and Xiu, 2018), all stocks, and all-but-micro stocks, respectively. TAB refers to a model with target asset-based factors alone, and TABJ+RRA refer to a K-factor model with J TAB factors and $K-J$ RRA factors. The sample period is 1974:01–2016:12.

	Total adj-R^2 (%)					Root-mean-squared pricing error (%)				
Model	1 factor	3 factors	5 factors	6 factors	10 factors	1 factor	3 factors	5 factors	6 factors	10 factors
Panel A: 48 industry portfolios (i.e., target assets)										
TAB	55.74	67.11	73.76	75.69	81.37	4.25	3.80	3.40	3.28	2.89
TAB1+RRA	–	62.53	66.51	67.28	69.50	–	3.88	3.68	3.63	3.49
TAB2+RRA	–	66.14	69.63	70.69	73.12	–	3.82	3.60	3.53	3.37
TAB3+RRA	–	–	72.19	73.16	75.84	–	–	3.49	3.42	3.23
Panel B: 202 characteristic portfolios										
TAB	75.60	77.39	80.93	81.09	84.07	2.85	2.76	2.55	2.54	2.35
TAB1+RRA	–	80.71	83.10	84.82	86.58	–	2.57	2.42	2.32	2.19
TAB2+RRA	–	80.31	82.96	83.90	86.76	–	2.61	2.43	2.37	2.18
TAB3+RRA	–	–	82.59	83.29	86.36	–	–	2.46	2.41	2.21
Panel C: All stocks										
TAB	9.51	11.08	13.51	13.86	15.35	17.45	17.04	16.54	16.37	15.69
TAB1+RRA	–	12.93	14.13	14.77	16.01	–	16.87	16.49	16.30	15.64
TAB2+RRA	–	12.67	13.97	14.46	15.93	–	16.91	16.50	16.32	15.64
TAB3+RRA	–	–	14.05	14.23	15.79	–	–	16.50	16.34	15.65
Panel D: All-but-micro stocks										
TAB	21.07	24.68	28.77	29.31	31.88	11.80	11.35	10.91	10.77	10.20
TAB1+RRA	–	27.31	29.49	30.29	32.05	–	11.19	10.85	10.69	10.20
TAB2+RRA	–	27.14	29.18	30.10	32.14	–	11.21	10.87	10.71	10.19
TAB3+RRA	–	–	29.23	29.76	32.21	–	–	10.87	10.74	10.18

Table 13 Performance of TAB and RRA factors targeted at explaining 202 characteristic portfolios

This table reports the total adj-R^2s and root-mean-squared pricing errors of different factor models in explaining four sets of testing assets: 48 industry portfolios, 202 characteristic portfolios, all stocks, and all-but-micro stocks, respectively. TAB refers to a model with target asset-based factors alone, and TABJ+RRA refer to a K-factor model with J TAB factors and $K-J$ RRA factors. The target assets that represent the cross section of stock returns are 202 characteristic portfolios in Giglio and Xiu (2018), including 25 size-B/M portfolios, 17 industry portfolios, 25 operating profitability-investment portfolios, 25 size-variance portfolios, 35 size-net issuance portfolios, 25 size-accruals portfolios, 25 size-beta portfolios, and 25 size-momentum portfolios. The sample period is 1974:01–2016:12.

	Total adj-R^2 (%)					Root-mean-squared pricing error (%)				
Model	1 factor	3 factors	5 factors	6 factors	10 factors	1 factor	3 factors	5 factors	6 factors	10 factors
Panel A: 48 industry portfolios										
TAB	52.47	58.46	62.94	65.26	69.20	4.43	4.08	3.88	3.78	3.53
TAB1+RRA	–	58.49	62.11	63.38	65.12	–	4.08	3.89	3.82	3.70
TAB2+RRA	–	58.34	62.07	63.30	65.16	–	4.08	3.89	3.83	3.70
TAB3+RRA	–	–	61.44	63.31	65.25	–	–	3.92	3.83	3.70
Panel B: 202 characteristic portfolios (i.e., target assets)										
TAB	80.21	87.63	90.06	90.89	92.48	2.61	2.08	1.91	1.85	1.72
TAB1+RRA	–	87.12	89.23	89.70	90.68	–	2.12	1.99	1.95	1.87
TAB2+RRA	–	87.48	89.51	90.15	91.25	–	2.10	1.96	1.91	1.81
TAB3+RRA	–	–	89.62	90.38	91.48	–	–	1.95	1.89	1.79
Panel C: All stocks										
TAB	11.44	14.87	16.67	17.32	18.46	17.26	16.69	16.25	16.06	15.43
TAB1+RRA	–	14.43	15.85	16.34	17.31	–	16.73	16.33	16.15	15.54
TAB2+RRA	–	14.85	16.29	16.90	18.16	–	16.69	16.29	16.10	15.47
TAB3+RRA	–	–	16.15	16.95	18.16	–	–	16.30	16.10	15.48
Panel D: All-but-micro stocks										
TAB	23.58	28.29	30.26	31.16	33.40	11.59	11.08	10.77	10.61	10.09
TAB1+RRA	–	28.05	30.23	30.95	32.62	–	11.10	10.76	10.62	10.14
TAB2+RRA	–	28.14	30.29	30.84	32.77	–	11.09	10.76	10.63	10.14
TAB3+RRA	–	–	30.21	30.94	32.84	–	–	10.76	10.63	10.14

Appendix

A Derivation of the GMM estimators

Following Zhou (1994), we can write the objective function as

$$Q = \mathrm{tr}[W_1(R - X\alpha' - \Theta'G)'P_0(R - X\alpha' - \Theta'G)]/T^2. \tag{A.1}$$

Now we solve α conditional on other parameters. Let U be a $T \times N$ matrix of the residuals or the matrix obtained from stacking the U_t's. Then it can be verified that

$$U'P_0U = (R - X\hat{\alpha}' - \Theta'G)'(R - X\hat{\alpha}' - \Theta'G) + (\alpha' - \hat{\alpha}')'X'P_0X(\alpha' - \hat{\alpha}'). \tag{A.2}$$

Hence, regardless of W_1, Q is minimized at $\alpha = \hat{\alpha}$. Consequently, we only need to minimize

$$Q^* = \mathrm{tr}[W_1(R - \Theta'G)'P(R - \Theta'G)]/T^2. \tag{A.3}$$

The rest of the proof follows similarly from Zhou (1994). Q.E.D.

B D_T representation

For the reader's application convenience, this appendix provides an explicit expression for D_T for both computing the GMM test and for checking the optimality of the estimators with $D_T'W_T\boldsymbol{h}_T$ being zero.

Recall that the regression system is:

$$\begin{aligned} R_t &= \alpha + \beta' f_t + U_t, \quad \beta : K \times N \\ &= \alpha + \Theta' g_t + U_t, \ \Theta : L \times N \\ f_t &= \Phi' g_t, \qquad\qquad \Phi : L \times K, \ \Theta = \Phi\beta. \end{aligned}$$

To make the solution unique (Zhou, 1994), Φ and β can be represented as

$$\Phi = \begin{pmatrix} \phi_{11} & \phi_{12} & \cdots & \phi_{1K} \\ \phi_{21} & \phi_{22} & \cdots & \phi_{2K} \\ \cdots & \cdots & \cdots & \cdots \\ \phi_{L1} & \phi_{L2} & \cdots & \phi_{LK} \end{pmatrix}, \quad \beta = \begin{pmatrix} 1 & 0 & \cdots & 0 & \beta_{1,K+1} & \cdots & \beta_{1N} \\ 0 & 1 & \cdots & 0 & \beta_{2,K+1} & \cdots & \beta_{2N} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 & \beta_{K,K+1} & \cdots & \beta_{KN} \end{pmatrix} = (I_K, \beta_2).$$

The parameters to be estimated are

$$\begin{aligned}\theta &= (\alpha', vec(\Phi)', vec(\beta_2)')' \\ &= (\alpha_1, \cdots, \alpha_N, \phi_{11}, \cdots, \phi_{L1}, \cdots, \phi_{1K}, \cdots, \phi_{LK}, \beta_{1,K+1}, \cdots, \beta_{K,K+1}, \cdots, \beta_{1N}, \cdots, \beta_{KN}).\end{aligned}$$

Let $h_t = U_t(\alpha, \Phi, \beta_2) \otimes Z_t$ with $U_t = R_t - \alpha - (\Phi\beta)' g_t = R_t - \alpha - \begin{pmatrix} \Phi' g_t \\ \beta_2' \Phi' g_t \end{pmatrix}$ and $Z_t = (1, g_t')'$. Then

$$\Phi' g_t = \begin{pmatrix} \sum_{i=1}^{L} \phi_{i1} g_{it} \\ \sum_{i=1}^{L} \phi_{i2} g_{it} \\ \cdots \\ \sum_{i=1}^{L} \phi_{iK} g_{it} \end{pmatrix} \quad \text{and} \quad \beta_2' \Phi' g_t = \begin{pmatrix} \sum_{j=1}^{K} [\beta_{j,K+1} \sum_{i=1}^{L} \phi_{ij} g_{it}] \\ \sum_{j=1}^{K} [\beta_{j,K+2} \sum_{i=1}^{L} \phi_{ij} g_{it}] \\ \cdots \\ \sum_{j=1}^{K} [\beta_{j,N} \sum_{i=1}^{L} \phi_{ij} g_{it}] \end{pmatrix}.$$

Since

$$\frac{\partial U_t}{\partial \alpha} = -I_N,$$

$$\frac{\partial \Phi' g_t}{\partial vec(\Phi)} = \begin{pmatrix} g_{1t} & \cdots & g_{Lt} & \cdot & \cdots & \cdot & 0 & \cdots & 0 \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ 0 & \cdots & 0 & \cdot & \cdots & \cdot & g_{1t} & \cdots & g_{Lt} \end{pmatrix} = I_K \otimes g_t',$$

$$\frac{\partial \beta_2' \Phi' g_t}{\partial vec(\Phi)} = \begin{pmatrix} \beta_{1,K+1} g_{1t} & \cdots & \beta_{1,K+1} g_{Lt} & \cdot & \cdots & \cdot & \beta_{K,K+1} g_{1t} & \cdots & \beta_{K,K+1} g_{Lt} \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ \beta_{1,N} g_{1t} & \cdots & \beta_{1,N} g_{Lt} & \cdot & \cdots & \cdot & \beta_{K,N} g_{1t} & \cdots & \beta_{K,N} g_{Lt} \end{pmatrix} = \beta_2' \otimes g_t',$$

and

$$\frac{\partial \beta_2' \Phi' g_t}{\partial vec(\beta_2)} = \begin{pmatrix} \sum_{i=1}^{L} \phi_{i1} g_{it} & \cdots & \sum_{i=1}^{L} \phi_{iK} g_{it} & \cdot & \cdots & \cdot & 0 & \cdots & 0 \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ 0 & \cdots & 0 & \cdot & \cdots & \cdot & \sum_{i=1}^{L} \phi_{i1} g_{it} & \cdots & \sum_{i=1}^{L} \phi_{iK} g_{it} \end{pmatrix} = I_{N-K} \otimes g_t' \Phi,$$

the first-order derivative of h_t with respect to θ is

$$D_t = \frac{\partial h_t}{\partial \theta} = \frac{\partial U_t}{\partial \theta} \otimes Z_t = -\begin{pmatrix} I_K & 0 & I_K \otimes g_t' & 0 \\ 0 & I_{N-K} & \beta_2' \otimes g_t' & I_{N-K} \otimes g_t' \Phi \end{pmatrix} \otimes Z_t.$$

Thus,

$$D_T = \frac{1}{T} \sum_{t=1}^{T} D_t. \qquad \text{Q.E.D.}$$

C GMM extracted factors

By taking expectation in (29), it is clear that $e_t = (I - BQ')(R_t - \mu)$, and so $\boldsymbol{h}_T^e$ is well defined. Then, following Zhou (1994), we have

$$Q_2 = \mathrm{tr}[W_1(I - BQ')\hat{\Sigma}_1[(I - BQ')], \tag{C.1}$$

after replacing μ by its GMM estimator $\hat{\mu}$ with $\hat{\Sigma}_1 = (R - 1_T\hat{\mu}')'P_0(R - 1_T\hat{\mu}')/T^2$.

Noting the solution of Balvers and Stivers (2018) to their (A.2) and the fact that W_1 is unrelated to B, we need only to minimize

$$Q_2^* = \mathrm{tr}[\hat{\Sigma}_q - \hat{\Sigma}_q C(C'\hat{\Sigma}_q C)^{-1}C'\hat{\Sigma}_q]. \tag{C.2}$$

Then the solution follows. Q.E.D.

D Definitions of 62 anomalies

Dvp Dividend to price ratio (Litzenberger and Ramaswamy, 1979): Annual total dividends payouts divided by the market value at the end of June

Top Total payouts (Boudoukh, Michaely, Richardson, and Roberts, 2007): Dividends (dvc) on common stock plus repurchases

Nop Net payout yields (Boudoukh, Michaely, Richardson, and Roberts, 2007): Total payouts minus equity issuances

Ssgrow Sustainable growth (Lockwood and Prombutr, 2010): Annual growth in book value of equity

Ebp Enterprise component of book to price (Penman, Richardson, and Tuna, 2007): Book value of net operating asset (net debt plus book equity) to the net operating assets (net debt plus market equity)

Ndp Net debt to price (Penman, Richardson, and Tuna, 2007): Net debt to the market equity

Dur Equity duration (Dechow, Sloan, and Soliman, 2004): Weighted average of the time to each of the respective net cash distributions divided by market equity

Ndf Net debt financing (Bradshaw, Richardson, and Sloan, 2006): Cash proceeds from the issuance of long-term debt (dltis) minus cash payments for long-term debt reductions (dltr), plus the net changes in current debt (dlcch)

Nxf Net external financing (Bradshaw, Richardson, and Sloan, 2006): Sale of common and preferred stocks (sstk) minus cash payments for the repurchases of preferred stocks (prstkc), minus cash payments for dividends (dv)

Cei Composite equity issuance (Daniel and Titman, 2006): Log growth rate in the market equity not attributable to stock returns

Aci Abnormal capital investment (Titman, Wei, and Xie, 2004): Capital expenditure (capx) for the fiscal year divided by the average of last three years of capital expenditure minus one

Noa Net operating asset (Hirshleifer, Hou, Teoh, and Zhang, 2004): Operating assets (at-che) minus operating liabilities (at−dlc−dltt−mib−pstk−ceq)

Pta Percentage total accruals (Richardson, Sloan, Soliman, and Tuna, 2005): Total accruals scaled by the absolute value of net income (ni)

dCoa Change in current operating assets (Richardson, Sloan, Soliman, and Tuna, 2005): Change in current assets (act) minus change in cash and short term investment (che)

dNco Change in net non-current operating assets (Richardson, Sloan, Soliman, and Tuna, 2005): Change in non-current operating assets minus change in non-current liabilities

dNca Change in non-current operating assets (Richardson, Sloan, Soliman, and Tuna, 2005): Change in total asset (at) minus change in current assets (act), minus change in long-term investments (ivao)

dFnl Change in financial liabilities (Richardson, Sloan, Soliman, and Tuna, 2005): Change in short-term investments plus change in long-term investments

Cop Cash-based operating profitability (Ball, Gerakos, Linnainmaa, and Nikolaev, 2015): Total revenue (revt) minus cost of goods sold (cogs), minus selling, general, and administrative expenses (xsga), plus research and development expenditures (xrd), minus change in accounts receivable (rect), minus change in inventory (invt), minus change in prepaid expenses (xpp), plus change in deferred revenue (drc+drlt), plus change in trade accounts payable (ap), and plus change in accrued expenses (xacc), all scaled by book assets (at)

F_g7 F-score (Piotroski, 2000): The sum of nine firm's fundamental signals as either good or bad depending on the signals' implications for future stock prices and profitability

Ol Operating leverage (Novy-Marx, 2011): Operate costs to total assets

Rdm R&D to market (Chan, Lakonishok, and Sougiannis, 2001): R&D expenses (xrd) divided by the market value at the end of December

Adm Advertising expenses-to-market equity (Chan, Lakonishok, and Sougiannis, 2001): Advertising expenses (xad) to market value at the end of December

Bca Brand capital to assets (Belo, Lin, and Vitorino, 2014): Accumulating advertising expensed with the perpetual inventory method

Oca_ia Industry-adjusted Organizational capital to assets (Eisfeldt and Papanikolaou, 2013): Organizational capital to assets with 17 industry adjusted with the perpetual inventory method

Rnaq Quarterly return on net operating assets (Soliman, 2008): Quarterly operating income after depreciation (oiadpq) divided by one-quarter-lagged net operating assets

Pmq Quarterly profit margin (Soliman, 2008): Quarterly operating income after depreciation (oiadpq) divided by quarterly slaes (saleq)

Atoq Quarterly asset turnover (Soliman, 2008): Quarterly sales (saleq) divided by one-quarter-lagged net operating assets

Ctoq Quarterly capital turnover (Haugen and Baker, 1996): Quarterly sales (saleq) divided by one-quarter-lagged total assets (atq)

Glaq Quarterly gross profits to lagged assets (Novy-Marx, 2013): Quarterly total revenue (revtq) minus cost of goods sold (cogsq), divided by one-quarter-lagged total assets (atq)

Oleq Quarterly operating profits to lagged equity (Fama and French, 2015): Quarterly total revenue (revtq) minus cost of goods sold (cogsq), minus selling, general, and administraive expenses (xsgaq), minus interest expenses (xintq), all scaled by one-quarter-lagged book equity

Olaq Quarterly operating profits to lagged assets (Ball, Gerakos, Linnainmaa, and Nikolaev, 2015): Quarterly total revenue (revtq) minus cost of goods sold (cogsq), minus selling, general, and administraive expenses (xsgaq), plus research and development expenditures (xrdq), all scaled by one-quarter-lagged book assets (atq)

Claq Quarterly cash-based operating profits to lagged assets (Ball, Gerakos, Linnainmaa, and Nikolaev, 2015): Quarterly total revenue (revtq) minus cost of goods sold (cogsq), minus selling, general, and administraive expenses (xsgaq), plus research and development expenditures (xrdq), minus change in accounts receivable (rectq), minus change in inventory (invtq), plus change in deferred revenue (drcq+drltq), plus change in trade accounts payable (apq), all scaled by one-quarter-lagged book assets (atq)

Oq Quarterly O-score (Dichev, 1998): Replace annual O-score components as quarterly components

Olq Quarterly operating leverage (Novy-Marx, 2011): Quarterly operating costs (cogsq+xsgaq) divided by assets (atq) for the fiscal quarter ending at least four months ago

Kzq Quarterly Kaplan-Zingales index (Lamont, Polk, and Saaa-Requejo, 2001): Replace annual KZ index components as quarterly components

Acc Working capital accruals (Sloan, 1996): Annual income before extraordinary items (ib) minus operating cash flows (oancf) divided by average total assets (at); if oancf is missing then set to change in act−change in che−change in lct+change in dlc+change in txp−dp

Agr Asset growth (Cooper, Gulen, and Schill, 2008): Annual percent change in total assets (at)

Bm_ia Industry-adjusted book to market (Asness, Porter, and Stevens, 2000): Industry adjusted book-to-market ratio

Cashdebt Cash flow to debt (Ou and Penman, 1989): Earnings before depreciation and extraordinary items (ib+dp) divided by average total liabilities (lt)

Cfp Cash flow to price ratio (Desai, Rajgopal, and Venkatachalam, 2004): Operating cash flows divided by fiscal-year-end market capitalization

Cfp_ia Industry-adjusted cash flow to price ratio (Asness, Porter, and Stevens, 2000): Industry adjusted cfp

Chcsho Change in shares outstanding (Pontiff and Woodgate, 2008): Annual percent change in shares outstanding (csho)

Chinv Change in inventory (Thomas and Zhang, 2002): Change in inventory (inv) scaled by average total assets (at)

Egr Growth in common shareholder equity (Richardson, Sloan, Soliman, and Tuna, 2005): Annual percent change in book value of equity (ceq)

Ep Earnings to price (Basu, 1977): Annual income before extraordinary items (ib) divided by end of fiscal year market cap

gCapx Growth in capital expenditures (Anderson and Garcia-Feijoo, 2006): Percent change in capital expenditures from year $t-2$ to year t

gLtnoa Growth in long term net operating assets (Fairfield, Whisenant, and Yohn, 2003): Growth in long term net operating assets

Hire Employee growth rate (Belo, Lin, and Vitorino, 2014): Percent change in number of employees (emp)

Invest Capital expenditures and inventory (Hou, Xue, and Zhang, 2018): Annual change in gross property, plant, and equipment (ppegt) + annual change in inventories (invt) all scaled by lagged total assets (at)

Lgr Growth in long-term debt (Richardson, Sloan, Soliman, and Tuna, 2005): Annual percent change in total liabilities (lt)

Orgcap Organizational capital (Eisfeldt and Papanikolaou, 2013): Capitalized SG&A expenses

Pchsale_Pchinvt % change in sales−% change in inventory (Abarbanell and Bushee, 1998): Annual percent change in sales (sale) minus annual percent change in inventory (invt)

Pchsaleinv % change sales-to-inventory (Ou and Penman, 1989): Percent change in saleinv

Roic Return on invested capital (Brown and Rowe, 2007): Annual earnings before interest and taxes (ebit) minus non-operating income (nopi) divided by non-cash enterprise value (ceq+lt−che)

Saleinv Sales to inventory (Ou and Penman, 1989): Annual sales divided by total inventory

Salerec Sales to receivables (Ou and Penman, 1989): Annual sales divided by accounts receivable

Sp Sales to price (Barbee Jr., Mukherji, and Raines, 1996): Annual revenue (sale) divided by fiscal-year-end market capitalization

Tb Tax income to book income (Lev and Nissim, 2004): Tax income, calculated from current tax expense divided by maximum federal tax rate, divided by income before extraordinary items

Chtxq Quarterly change in tax expense (Thomas and Zhang, 2011): Percent change in total taxes (txtq) from quarter $t-4$ to t

Ear Earnings announcement return (Kishore, Brandt, Santa-Clara, and Venkatachalam, 2008): Sum of daily returns in three days around earnings announcement. Earnings announcement from Compustat quarterly file (rdq)

Roaq Return on assets (Balakrishnan, Bartov, and Faurel, 2010): Income before extraordinary items (ibq) divided by one quarter lagged total assets (atq)

Roeq Return on equity (Hou, Xue, and Zhang, 2015): Earnings before extraordinary items divided by lagged common shareholders' equity