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: June 2018

*We are grateful to seminar participants at Washington University in St. Louis for insightful and helpful comments.

Please 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

In this paper, we propose a simple reduced-rank approach (RRA) to reduce a large number of factors to a few parsimonious ones. In contrast with the popular PCA and PLS, the factors from RRA are designed to explain the cross section of stock returns, and are unrelated to maximizing the variations of the factors or the covariances between the factors and returns. Out of 70 potential factors, a composite of five factors identified by using the RRA outperforms the Fama-French (2015) five-factor model substantially for 48 industry portfolios, but only marginally for all individual stocks. The same conclusion is true with alternative target portfolios. Our results suggest that existing potential factors do not have 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, 2017). 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 (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 potential factors that we use below including the market factor, we are interested in finding a few out of them. If we restrict to only one factor, RRA will find a linear combination of the 70 that best explains the cross section of stock returns. Interestingly, under the one factor restriction, the factor found by RRA is almost identical to the market factor, indicating that the market factor is the most important one among the 70 candidates. In contrast, the popular principal component analysis (PCA) and partial least square (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 5-factor restriction, the five factor composites chosen by RRA outperform the Fama and French (2015) 5-factor model substantially for 48 industry portfolios, which are the target portfolios used to select the factors.[1] However, the RRA factors are only marginally better when applied to all individual stocks. The same conclusion is true with alternative target portfolios. Note that this is unlikely due to the methodology as RRA is statistically designed to pick up the best factors to explain the returns. Hence, we interpret our results as suggesting that the 70 factors do not have enough new information at the stock level beyond the Fama and French (2015) five factors.

In the statistical learning literature, RRA is a dimension reduction tool. It imposes a rank restriction on regression coefficients, so that a lower rank restriction can effectively be used to reduce a large number of regressors/factors into a small number of their composites/linear combinations. Anderson (1951) appears to

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

be the first study in statistics. Reinsel and Velu (1998) provide a book-level analysis on its properties and applications. In finance, Velu and Zhou (1999) apply RRA to test multi-beta asset pricing models, and Zhou (1994) extends it to the generalized method of moment (GMM) framework of Hansen (1982). In the spirit of Zhou (1994), this paper develops a reduced-rank approach to shrinking factor dimension.

This paper also extends the RRA to pre-specify Fama and French (2015) five factors as part of the true factors. Then we ask whether a few composite factors from the remaining 65 candidates can help explaining the cross section of stock returns. Interestingly, we find that, for both the target portfolios and the cross section of all stocks, the additional factors add little extra explanatory power, suggesting again no much can be gained out of the common candidate factors beyond the Fama and French (2015) five factors. In light of Harvey, Liu, and Zhu (2016), it seems there are too many factors in finance. However, our paper in fact 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 candidates by applying the PCA. Based on firm characteristics, Kelly, Pruitt, and Su (2017) find seven factors that are significant using their instrumented principal components analysis (IPCA). Feng, Giglio, and Xiu (2017) find 14 out of 99 with LASSO, and Freyberger, Neuhierl, and Weber (2017) 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 (2017, 2018) find that the best linear combinations of the candidates in explaining target returns 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 with the various LASSO methods, our RRA estimates are analytically done and their asymptotic statistical properties are known.

The rest of the paper is organized as follows.

2 Methodology

In this section, we provide a reduced-rank approach and compare it with the PCA and the PLS. The extension with pre-specified factors is outlined at the end.

2.1 RRA

Following most studies, we assume that asset returns are governed by a multi-factor model:

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

where R_{it} is the return on asset i in period t $(1 \leq i \leq N)$, f_{jt} is the realization of the j-th factor in period t $(1 \leq j \leq K)$, ε_{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 common interest is that 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 $g_{1t}, \cdots, g_{Lt}$ are L variables that can be highly correlated with factors f. 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 many factor candidates. This assumption is also made for the PCA and PLS, two popular dimension reduction methods. However, they estimate the combination coefficients differently from our approach below. A detailed comparison of the three methods will be provided later.

The univariate regressions can be written in a matrix manner,

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

where $R_t = (R_{1t}, \cdots, R_{Nt})'$ is an $N \times 1$ vector of the returns, G_t is an $L \times 1$ vector of the proxies, and Θ is an

$L \times N$ matrix of the parameters. Equation (2) implies

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

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

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

In other words, when the K factors can be expressed as a linear combination of the proxies, the rank of the regression coefficients cannot exceed K. On the other hand, if the regression coefficients have a rank of K, there must have a reduction of L proxies from K factors. Hence, the estimation of Φ and B is from 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 obtaining easily the asymptotic distribution of the parameters. Let Z_t be an $M \times 1$ vector of the instruments available at time t. Then, the moment condition is

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

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:

$$\boldsymbol{h}_T(\alpha, \Theta) = \frac{1}{T}\sum_{t=1}^{T} h_t(\alpha, \Theta). \tag{7}$$

Then the GMM estimator solves the following minimization problem:

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

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

$$Z_t = [1, G_t']'.$$

Since Φ and B 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:

$$W_T \equiv W_1 \otimes W_2, \qquad W_1 : N \times N, \quad W_2 : M \times M, \tag{9}$$

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

$$\hat{\alpha} = (X'P_0X)^{-1}X'P_0(R - G\Theta), \tag{10}$$

where X is a $T \times 1$ 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 proxies. The proof, based on Zhou (1994), is provided in the appendix.

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

$$\hat{\Phi} = (G'PG/T^2)^{-1/2}E, \qquad \hat{B} = (G^{*\prime}PG^*)^{-1}G^{*\prime}PR, \tag{11}$$

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:

$$(G'PG/T^2)^{-1/2\prime}(G'PR/T^2)W_1(G'PR/T^2)'(G'PG/T^2)^{-1/2}. \tag{12}$$

In summary, the estimators are computed easily in practice. By using the identity weighting matrix, we can compute sequentially from (12) to (11) and to (10), obtaining all the parameter estimates. The estimated factors will be given by

$$\hat{f}_t = \hat{\Phi}' G_t,$$

where f_t is a $K \times 1$ vector of the factor realizations, and G_t is an $L \times 1$ vector of the realizations of the g_{lt}'s. We can standardize f_t to make it has identity covariance matrix.

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

$$\hat{S}_T = \left(\frac{1}{T}\sum \hat{U}_t \hat{U}_t'\right) \otimes \left(\frac{1}{T}\sum \hat{Z}_t \hat{Z}_t'\right) \tag{13}$$

as the weighting matrix, where $\hat{U}_t$ is evaluated at the previous estimator with identity weighting matrix. In our applications below, N or L or both are usually large. In this case, one can simply use only the diagonal elements of $\hat{S}_T$. Of course, the best estimator is obtained by using the optimal weighting matrix, but, as pointed out earlier, this is not feasible due to the lack of an analytical solution.

2.2 Comparison with PCA

Principal component analysis (PCA) has a long history. Since its introduction by Pearson (1901), it has been widey used in all sciences. The idea is to transform a set of random variables (factors) into independent ones, so that the first has the largest variance, the second one has the next largest, and so on. Mathematically, we find $\phi_k^{\text{PCA}} = (\phi_{k1}^{\text{PCA}}, \ldots, \phi_{kL}^{\text{PCA}},)'$ to solve successively

$$\max \operatorname{Var}(Z_t' \phi_k^{\text{PCA}}) \tag{14}$$

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, $\hat{\phi}_1^{\text{PCA}}, \ldots, \hat{\phi}_K^{\text{PCA}}$. Then

$$\hat{f}_{kt}^{\text{PCA}} = \hat{\phi}_{k1}^{\text{PCA}} g_{1t} + \cdots + \hat{\phi}_{kL}^{\text{PCA}} g_{Lt}, \quad k = 1, \ldots, K, \tag{15}$$

are the PCA factors.

By design, the first K PCA factors represent the best factors that explain the variations of the factors. There is no guarantee that they are in any way close to the best factors that explain the 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. 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 PCA is designed to explain the factor variations, rather than returns.

2.3 Comparison with PLS

While PCA is popular, it replaces G_t by a few independent factors of maximal variances that makes no information of the target to be explained. Recognizing this, Wold (1966) introduces the partial least square (PLS) method to make the factor selection related 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 the extracted factors should have the maximum covariance. Hence, we choose a linear combination of returns too. Mathematically, the PLS solves

$$\max \mathrm{Cov}(R_t' \psi_k^{\mathrm{PLS}}, Z_t' \phi_k^{\mathrm{PLS}}), \tag{16}$$

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 similar as before.

In summary of PCA, PLS and RRA, PCA ignores the target to get independent factors. In contrast, 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 second moment of covariances. RRA seems particularly suited for finance applications. It extracts

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

the 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 GMM-based so that it is flexible in adding instrumental variables and is capable of drawing inferences and testing hypotheses within the popular GMM framework.

2.4 GMM Test

In our RRA framework, there exists an analytical GMM test for the number of factors. First, we use the weighting matrix given by (9) to obtain the analytical estimates. Since this matrix is not necessarily optimal under general heteroscedasticity, we cannot use the usual Hansen over-identification test.

Nevertheless, based on Zhou (1994), we can compute

$$H_z = T\left(M_T \boldsymbol{h}_T\right)' V_T \left(M_T \boldsymbol{h}_T\right), \tag{17}$$

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

$$\Omega_T = [I - D_T(D_T' W_T D_T)^{-1} D_T' W_T] S_T [I - D_T(D_T' W_T D_T)^{-1} D_T' W_T]', \tag{18}$$

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, and D_T is the first order derivatives of $\boldsymbol{h}_T$ with respect to α and Θ. Under the rank K hypothesis, H_z is asymptotically chi-squared distributed with the degree of freedom $[N(L+1) - q]$, with $q = N + LK + KN - K^2$.

2.5 Extension with 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. Currently, the five factors of Fama and French (2015) are one set of the most studied. It is hence of interest to include this set or some other factors as the true factors, while searching the best linear combination factors from the rest.

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

$$R_{it} = \alpha_i + \beta_0' R_{0t} + \beta_{i1} f_{1t} + \cdots + \beta_{iK} f_{Kt} + \varepsilon_{it}, \quad i = 1, \ldots, N, \ t = 1, \ldots, T, \tag{19}$$

where R_{0t} is a K_0-vector of pre-specified factors and β_0 are the factor loading, while the rest are similar as before. That is, we assume the other factors are related to a number of candidates,

$$f_{kt} = \phi_{k1} g_{1t} + \cdots + \phi_{kL} g_{Lt}, \quad k = 1, \ldots, K, \tag{20}$$

where $g_{1t}, \ldots, g_{Lt}$ are L proxies of the candidate factors excluding R_{0t}.

The estimation can be done as easily as before. The only difference is re-define X 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 the market beta. If there are more than one pre-specified factors, one can expand X and α accordingly. However, the degree of the GMM test has to be adjusted down by K_0 to reflect the additional K_0 parameters.

2.6 Performance Measures

In this paper, we are interested how the extracted factors explain individual stock returns. We use several measures to assess the performance of various models and estimates.

The first measure is the total R^2 as defined by Kelly, Pruitt, and Su (2017, 2018),

$$\text{Total } R^2 = 1 - \frac{\sum_{i,t} (R_{it} - \tilde{\alpha}_i - \tilde{\beta}_{i1} \tilde{f}_{1t} - \cdots - \tilde{\beta}_{iK} \tilde{f}_{Kt})^2}{\sum_{i,t} R_{it}^2}, \tag{21}$$

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. Although PLS and RRA factors are extracted based on the same target, the above measure is straightforward to compute from regressions of all excess returns on the factors.

The second measure is the aggregate mean-squared pricing error,

$$\text{PE} = \frac{1}{NT} \sum_{i,t} (R_{it} - \tilde{\beta}_{i1} \tilde{f}_{1t} - \cdots - \tilde{\beta}_{iK} \tilde{f}_{Kt})^2. \tag{22}$$

PE assesses to what extent there are returns not attributable to the extracted factors.

3 Empirical Results

3.1 Data

3.1.1 Target portfolios

We explore two sets of target portfolios to proxy for the cross section of stock returns. The first set of target portfolios consists of the Fama-French 48 industry portfolios, and the second set consists of 202 characteristic portfolios explored in 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 portfolio is that portfolios can efficiently reduce idiosyncratic noises in individual stock returns.

3.1.2 Factor candidates

We consider 70 factor candidates, which include Fama and French (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 63 anomalies that have significant CAPM alpha. We independently replicate about 120 anomalies that are explored in Green, Hand, and Zhang (2017) and Hou, Xue, and Zhang (2017) and have data since 1974. 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 candidates.

3.1.3 Testing assets

We consider four sets of testing assets to evaluate the pricing performance of the RRA factors. The first two sets of testing assets are the two sets of target portfolios, 48 industry portfolios and 202 characteristic portfolios. The third set is the universe of all common stocks (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 beginning of the month. Both Fama and French (2015) and Hou, Xue, and Zhang (2015) find that it is all-but-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).

3.2 Optimal number of factors

We perform the GMM test in Section 2.4 to explore how many factors we need to explain the cross section of stock returns. Especially, to describe the 48 industries portfolios or 202 characteristic portfolios, what is the minimum number of factors?

3.3 In-sampler performance

Table 2 reports the total R^2s and aggregate mean-squared pricing errors (PEs), defined as (21) and (22), of different factor models in explaining the testing assets. As a benchmark, we use Fama-French models, where 1-factor refers to the market factor, 3-factors to Fama and French (1993), 5-factors to Fama and French (2015), and 6-factors to Fama and French (2018) (i.e., 5-factors plus the momentum factor), respectively. The PCA K-factors refer to the K principal components corresponding to the largest K eigenvalues of the covariance matrix of the 70 factor candidates. The PLS and RRA factors refer to those that are extracted from the 70 factor candidates with the target to explain the cross section of the 48 industry portfolios.

Panel A of Table 2 shows that the MKT factor (1-factor), Fama and French (1993) three factors, Fama and French (2015) five factors, and Fama and French (2018) six factors explain 46.62%, 61.44%, 62.55%, and 63.84% of cross-sectional variation in the 48 industry portfolios, respectively. Among the three dimension shrinkage methods, the PCA performs the worst, whose first 1-, 3-, 5-, and 6-components explain only 32.36%, 35,30%, 46.71%, and 48.06% of variation, which is worse than the Fama-French factors. In contrast, the RRA factors perform the best, which explain 61.54%, 67.53%, 68.87%, and 69.35% of variation, thereby outperforming the Fama-French factors. When we extend the number of factors to 10 from 5, the pricing power increases from 46.71%, 64.39%, and 68.87% to 52.86%, 68.03%, and 70.84%, with the PCA, PLS, and RRA methods, respectively.

In Panel A, the RRA factors outperform the Fama-French factors in explaining the 48 industry portfolios. One may argue that this is because the RRA method uses the information in the 48 industry portfolios when extracting the factors. Panel B of Table 2 presents the total R^2s and PEs when the testing assets are the 202 characteristic portfolios. Since most of the portfolios are finer sorts of the Fama and French (2015) five factors, the pricing power should tilt toward the the Fama-French 5-factor model. As expected, the Fama-French performs better than the PCA, PLS, and RRA factors.

Panel C shows the most important result in this paper, where the testing assets are all individual stocks. That is, in addition to the market factor, the other four factors in Fama and French (2015) are based on only four firm characteristics, but their pricing power is comparable with that of the PCA, PLS, and RRA factors, which are based on 69 firm characteristics. For example, the total R^2s with Fama-French 3-, 5-, and 6-factors are 16.29%, 19.03%, and 20.68%, respectively. The corresponding values of the RRA factors are 17.44%, 20.23%, and 21.46%, which outperform the PCA and PLS factors. Therefore, to explain the cross section of individual stock returns, searching new factors seems meaningless.

Panel D repeats Panel C but excludes micro stocks in evaluating the factor models. The reason for this test is that both Fama and French (2015) and Hou, Xue, and Zhang (2015) argue that 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 the same as Panel C. That is, the total R^2s with the Fama-French factors and the RRA factors are very close each other, and both outperform that with the PCA and PLS factors.

The left-hand side of Table 2 reports PEs and the results are consistent with the total R^2.

Table 3

3.4 Out-of-sample performance

4 Results with Pre-specified Factors

Table 4

Table 5

5 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 candidates, including Fama and French (2015) five factors, momentum factor, Pastor and Stambaugh (2003) liquidity factor, Hou, Xue, and Zhang (2015) ROE factor, and 63 anomalies from Green, Hand, and Zhang (2017) and Hou, Xue, and Zhang (2017). We find that the 70 factors do not provide much new information at the stock level beyond the Fama and French (2015) five factors. In addition, we apply an extended RRA to the factor candidates with Fama and French (2015) five factors as pre-specified factors, and find that linear combinations of the remaining factor candidates improve little the performances.

Future research is to identify new factors that can provide independent information beyond the Fama and French (2015) five factors. Kelly, Pruitt, and Su (2017) 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 (2017). 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.

References

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

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.

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

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., 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., 2017. 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., Zhu, H., 2016. ... and the cross-section of expected returns. Review of Financial Studies 29, 5–68.

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., 2017. Replicating anomalies. Working paper.

Kelly, B. T., Pruitt, S., Su, Y., 2017. Instrumented principal component analysis. Working paper.

Kelly, B. T., Pruitt, S., Su, Y., 2018. Characteristics are covariances: A unified model of risk and return. Working paper.

Kozak, S., Nagel, S., Santosh, S., 2017. Shrinking the cross section. Working paper.

Kozak, S., Nagel, S., Santosh, S., 2018. Interpreting factor models. Journal of Finance, forthcoming.

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.

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

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

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.

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

This table reports the average returns, CAPM alphas, and t-values of 70 factor candidates, including Fama and French (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 63 anomalies that have significant CAPM alpha. The sample period is 1974:01–2016:12.

Factor candidate	Mean	α_{CAPM}	t_{CAPM}	Factor candidate	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 In-sample performance of factor models that are targeted at explaining 48 industry portfolios

This table reports the total R^2s and aggregate 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, in which 1-, 3-, 5-, and 6-factor(s) are the market factor, Fama and French (1993) three factors, Fama and French (2015) five factors, and Fama and French (2015) five factors plus the momentum factor, respectively. PCA, PLS, and RRA refer to the models that extract factors using the principal component analysis, partial least squares, and reduced-rank approach, respectively. The target returns that represent the cross section of stock returns for extracting the PLS and RRA factors are Fama-French 48 industry portfolios, and the factor candidates are those listed in Table 1. The sample period is 1974:01–2016:12.

	Total R^2 (%)					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	46.62	61.44	62.55	63.84	–	0.29	0.21	0.20	0.20	–
PCA	32.36	35.30	46.71	48.06	52.86	0.36	0.35	0.29	0.28	0.25
PLS	37.93	56.29	64.39	65.08	68.03	0.33	0.24	0.19	0.19	0.18
RRA	61.54	67.53	68.87	69.35	70.84	0.21	0.18	0.17	0.17	0.16
Panel B: 202 characteristic portfolios										
FF	73.36	85.68	87.07	88.44	–	0.10	0.05	0.05	0.05	–
PCA	35.79	39.61	50.82	52.96	60.18	0.22	0.21	0.17	0.16	0.14
PLS	42.10	67.87	81.31	82.64	85.45	0.20	0.11	0.07	0.06	0.06
RRA	75.19	78.66	80.17	83.94	86.37	0.09	0.08	0.07	0.06	0.05
Panel C: All stocks										
FF	10.29	16.29	19.03	20.68	–	4.21	3.90	3.72	3.61	–
PCA	9.79	13.33	16.79	18.07	22.94	4.18	3.96	3.77	3.69	3.38
PLS	10.61	15.54	18.98	20.46	24.96	4.15	3.91	3.72	3.63	3.32
RRA	13.40	17.44	20.23	21.46	25.87	4.07	3.84	3.66	3.59	3.29
Panel D: All-but-micro stocks										
FF	21.86	29.13	31.90	33.40	–	1.66	1.46	1.38	1.34	–
PCA	14.45	19.86	24.57	26.20	32.17	1.74	1.59	1.49	1.45	1.30
PLS	15.95	25.80	31.68	33.29	38.02	1.71	1.51	1.39	1.35	1.21
RRA	23.33	29.71	32.83	34.52	39.11	1.60	1.45	1.36	1.32	1.19

Table 3 In-sample performance of factor models that are targeted at explaining 202 characteristic portfolios

This table reports the total R^2s and aggregate 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, in which 1-, 3-, 5-, and 6-factor(s) are the market factor, Fama and French (1993) three factors, Fama and French (2015) five factors, and Fama and French (2015) five factors plus the momentum factor, respectively. PCA, PLS, and RRA refer to the models that extract factors using the principal component analysis, partial least squares, and reduced-rank approach, respectively. The target returns that represent the cross section of stock returns for extracting the PLS and RRA factors 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 candidates are those listed in Table 1. The sample period is 1974:01–2016:12.

	Total R^2 (%)					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	46.62	61.44	62.55	63.84	–	0.29	0.21	0.20	0.20	–
PCA	32.36	35.30	46.71	48.06	52.86	0.36	0.35	0.29	0.28	0.25
PLS	37.76	58.09	62.21	63.31	66.47	0.33	0.23	0.21	0.20	0.18
RRA	59.16	63.63	65.31	66.59	68.21	0.22	0.20	0.19	0.18	0.17
Panel B: 202 characteristic portfolios (i.e., target assets)										
FF	73.36	85.68	87.07	88.44	–	0.10	0.05	0.05	0.05	–
PCA	35.79	39.61	50.82	52.96	60.18	0.22	0.21	0.17	0.16	0.14
PLS	43.53	71.50	83.84	84.67	88.21	0.20	0.10	0.06	0.06	0.05
RRA	79.55	86.93	89.04	89.56	90.52	0.08	0.05	0.04	0.04	0.04
Panel C: All stocks										
FF	10.29	16.29	19.03	20.68	–	4.21	3.90	3.72	3.61	–
PCA	9.79	13.33	16.79	18.07	22.94	4.18	3.96	3.77	3.69	3.38
PLS	10.79	15.69	19.07	20.35	25.12	4.14	3.91	3.71	3.63	3.32
RRA	12.68	17.08	20.24	21.40	25.67	4.11	3.87	3.66	3.58	3.30
Panel D: All-but-micro stocks										
FF	21.86	29.13	31.90	33.40	–	1.66	1.46	1.38	1.34	–
PCA	14.45	19.86	24.57	26.20	32.17	1.74	1.59	1.49	1.45	1.30
PLS	16.63	26.55	31.94	33.39	38.32	1.69	1.51	1.38	1.35	1.21
RRA	24.38	30.15	33.32	34.75	39.13	1.58	1.45	1.36	1.32	1.19

Table 4 In-sample performance of factor models that pre-specify Fama and French (2015) five factors and are targeted at explaining 48 industry portfolios

This table reports the total R^2s and aggregate 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, in which 1-, 3-, 5-, and 6-factor(s) are the market factor, Fama and French (1993) three factors, Fama and French (2015) five factors, and Fama and French (2015) five factors plus the momentum factor, respectively. PCA, PLS, and RRA refer to the models that extract factors using the principal component analysis, partial least squares, and reduced-rank approach, respectively. The target returns that represent the cross section of stock returns for extracting the PLS and RRA factors are Fama-French 48 industry portfolios, and the factor candidates are those listed in Table 1. The sample period is 1974:01–2016:12.

	Total R^2 (%)					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)										
PCA	62.55	64.41	65.69	65.85	66.28	0.20	0.19	0.19	0.19	0.18
PLS	62.55	65.14	66.58	67.39	68.52	0.20	0.19	0.18	0.18	0.17
RRA	62.55	65.48	68.22	68.89	69.80	0.20	0.19	0.17	0.17	0.17
Panel B: 202 characteristic portfolios										
PCA	87.07	87.77	88.28	88.40	88.80	0.05	0.05	0.05	0.05	0.04
PLS	87.07	87.76	88.53	88.76	89.16	0.05	0.05	0.04	0.04	0.04
RRA	87.07	88.09	88.67	88.98	89.34	0.05	0.05	0.04	0.04	0.04
Panel C: All stocks										
PCA	19.03	20.56	21.78	22.82	24.87	3.72	3.62	3.54	3.47	3.32
PLS	19.03	20.58	21.99	23.14	25.27	3.72	3.62	3.53	3.45	3.31
RRA	19.03	21.14	22.51	23.69	25.95	3.72	3.59	3.50	3.42	3.27
Panel D: All-but-micro stocks										
PCA	31.90	33.55	34.97	35.93	38.08	1.38	1.34	1.30	1.27	1.21
PLS	31.90	33.81	35.24	36.33	38.55	1.38	1.34	1.30	1.27	1.20
RRA	31.90	33.56	35.51	36.79	39.02	1.38	1.34	1.30	1.26	1.20

Table 5 In-sample performance of factor models that pre-specify Fama and French (2015) five factors and are targeted at explaining 202 characteristic portfolios

This table reports the total R^2s and aggregate 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, in which 1-, 3-, 5-, and 6-factor(s) are the market factor, Fama and French (1993) three factors, Fama and French (2015) five factors, and Fama and French (2015) five factors plus the momentum factor, respectively. PCA, PLS, and RRA refer to the models that extract factors using the principal component analysis, partial least squares, and reduced-rank approach, respectively. The target returns that represent the cross section of stock returns for extracting the PLS and RRA factors 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 candidates are those listed in Table 1. The sample period is 1974:01–2016:12.

	Total R^2 (%)					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										
PCA	62.55	64.41	65.69	65.85	66.28	0.20	0.19	0.19	0.19	0.18
PLS	62.55	64.38	66.14	66.60	67.23	0.20	0.19	0.19	0.18	0.18
RRA	62.55	64.38	65.43	66.46	67.72	0.20	0.19	0.19	0.18	0.18
Panel B: 202 characteristic portfolios (i.e., target assets)										
PCA	87.07	87.77	88.28	88.40	88.80	0.05	0.05	0.05	0.05	0.04
PLS	87.07	88.00	88.62	89.04	89.81	0.05	0.05	0.04	0.04	0.04
RRA	87.07	88.54	89.14	89.78	90.33	0.05	0.04	0.04	0.04	0.04
Panel C: All stocks										
PCA	19.03	20.56	21.78	22.82	24.87	3.72	3.62	3.54	3.47	3.32
PLS	19.03	20.67	21.95	23.09	25.18	3.72	3.61	3.53	3.45	3.31
RRA	19.03	20.79	22.05	23.25	25.49	3.72	3.60	3.52	3.44	3.30
Panel D: All-but-micro stocks										
PCA	31.90	33.55	34.97	35.93	38.08	1.38	1.34	1.30	1.27	1.21
PLS	31.90	33.52	35.19	36.39	38.60	1.38	1.34	1.30	1.27	1.20
RRA	31.90	33.55	35.12	36.41	38.81	1.38	1.34	1.30	1.27	1.20

Appendix

A Derivatives of the objective function

For the reader's convenience, this appendix provides an explicit expression for D_T for the analytical GMM test by examining whether $D_T' W_T \boldsymbol{h}_T$ is zero.

Suppose the regression system is as follows:

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

Let $h_t = U_t(\alpha, \Theta) \otimes Z_t = U_t(\alpha, \Phi, B) \otimes Z_t$ and $\theta = (\alpha, \Phi, B)$, where $Z_t = (1, G_t')'$. Then

$$D_t = \frac{\partial h_t}{\partial \theta} = \frac{\partial U_t}{\partial \theta} \otimes Z_t = -\left(\begin{array}{cccc} I_K & 0 & U_{1t} & 0 \\ 0 & I_{N-K} & U_{3t} & U_{4t} \end{array} \right) \otimes Z_t,$$

where $U_{1t} = I_K \otimes G_t'$, $U_{3t} = B(1:K, K+1:N)' \otimes G_t'$, and $U_{4t} = I_{N-K} \otimes (\Phi' G(t,:)')'$