Basis Assets

Dong-Hyun Ahn
School of Economics, Seoul National University

Jennifer Conrad
Kenan-Flagler Business School, University of North Carolina

Robert F. Dittmar
Stephen M. Ross School of Business, University of Michigan

This paper proposes a new method of forming basis assets. We use return correlations to sort securities into portfolios and compare the inferences drawn from this set of basis assets with those drawn from other benchmark portfolios. The proposed set of portfolios appears capable of generating measures of risk–return trade-off that are estimated with a lower error. In tests of asset pricing models, we find that the returns of these portfolios are significantly and positively related to both CAPM and Consumption CAPM risk measures, and there are significant components of these returns that are not captured by the three-factor model. (*JEL* C10, G11)

1. Introduction

A fundamental object in asset pricing is the investment opportunity set, the set of assets that investors use in making portfolio decisions. The pioneering work of Markowitz (1952), Cass and Stiglitz (1970), and Ross (1978) shows that this set can be reduced to a group of portfolios that dominate the opportunity set represented by individual assets. The subsequent empirical literature that tests the implications of asset pricing models has used this insight to focus on models' ability to describe the returns on a relatively small set of portfolios rather than a large number of individual assets. The implicit assumption in this literature is that the portfolios analyzed span the *ex ante* opportunity set available to investors. We term this representative set of portfolios a set of basis assets.

We thank Michael Brandt and Wayne Ferson and two anonymous referees for helpful comments, as well as seminar participants at the 2004 American Finance Association Meetings (San Diego, CA), Arizona State University, Baruch College, Case-Western Reserve, Federal Reserve Bank of Chicago, DePaul University, Harvard University, the Joint University of Alberta/University of Calgary Conference, London Business School, London School of Economics, Texas A&M, UCLA, the University of North Carolina, University of Oregon, University of Washington, the 2005 Utah Winter Finance Conference, and Vanderbilt University for helpful comments and suggestions. Portfolio data used in this paper are available at http://webuser.bus.umich.edu/rdittmar by following the "Research" link. The usual disclaimer applies. Send correspondence to Jennifer Conrad, CB#3490, McColl Building, University of North Carolina-Chapel Hill, Chapel Hill, NC 27599-3490; telephone: 919-962-3132; fax: 919-962-2068. E-mail: j_conrad@unc.edu.

doi:10.1093/rfs/hhp065 Advance Access publication October 20, 2009

In recent years, it has become increasingly common to use size- and book-to-market-sorted portfolios as basis assets, and to ask whether asset pricing models can explain the dispersion in the returns on these portfolios. Forming portfolios based on these characteristics has the advantage that it generates a large dispersion in returns, and hence presents a challenge to any asset pricing model. Fama and French (1996), for example, show that their three-factor model can explain more than 90% of the returns of these portfolios and that the unexplained portion of returns is economically small. However, the practice of using characteristic-sorted portfolios as basis assets has sparked a debate about whether they are appropriate to use in order to draw inferences about asset pricing models. Cochrane (2001) advances the opinion that such a procedure is precisely what researchers should do, as this approach generates dispersion in expected returns along dimensions of interest. In contrast, Lo and MacKinlay (1990) suggest that sorting on characteristics that are known to be correlated with returns generates a data-snooping bias.[1] Conrad, Cooper, and Kaul (2003) show that the increasing tendency of researchers to sort on multiple characteristics, and consequently to form larger number of portfolios, exacerbates the data-snooping bias. In this case, the dispersion may weaken or disappear in out-of-sample tests because the relation between returns and the characteristic is not robust over time.

In addition to the concern about data snooping, Daniel and Titman (2005) argue that firm characteristics such as book-to-market equity serve as a "catch-all," and capture differences in the sensitivities of firms' returns to a number of different fundamental factors. Consequently, asset pricing model tests on size- and book-to-market-sorted portfolios will find that factors based on these characteristics are important, but are unable to inform us as to the importance of other, perhaps more fundamental factors, because there is insufficient variation along any other specific factor in these portfolios. Daniel and Titman go on to argue that a new set of portfolios is required if statistical tests of asset pricing models are to have sufficient power to reject the model, or to identify important factors. As Jagannathan and Wang (1996, p. 36) suggest, "... we need to devise methods for evaluating the economic importance of the data sets used in empirical studies of asset-pricing models."

In this paper, we suggest a particular approach for forming basis assets that we argue is well-motivated economically, alleviates some of the problems inherent in the usual approaches to forming these assets, and may have the ability to generate more informative tests of asset pricing models. The method that we propose focuses on the properties of the covariance matrix of returns rather than the *ex post* vector of mean returns. This focus is sensible since

[1] Berk (2000) and Kan (1999) extend this analysis to consider subsorts within groups of characteristic-sorted portfolios. These authors suggest that the characteristics upon which researchers sort, such as size, may be mechanically related to expected returns.

the covariance matrix is the central object in the portfolio theory pioneered by Markowitz (1952). In particular, we suggest that the appropriate sort should attempt to group (separate) firms that are highly (less) correlated, as opposed to grouping firms that have similar realized returns. We utilize a correlation-based distance measure to form portfolios. In focusing on the covariance matrix as an interesting object for economically differentiating sets of basis assets, we also examine an important, but little discussed, characteristic of sorted portfolios: the conditioning of the covariance matrix formed by the basis assets. This feature of the covariance matrix is a critical determinant of the precision with which any inferences related to the basis assets can be drawn.

We use cluster analysis and a distance measure proposed by Ormerod and Mounfield (2000) to sort firms into portfolios. A small number of these portfolios generates significant dispersion in subsequent returns: the spread in average returns when ten cluster portfolios are formed is 47 basis points per month. This dispersion in mean returns is comparable to, or better than, that observed in the ten size-, book-to-market-, or beta-sorted portfolios of 51, 54, and 18 basis points per month, respectively. When 25 cluster portfolios are formed, the spread in mean returns increases substantially to 79 basis points per month; this large dispersion is accompanied by substantially higher volatility of returns. In general, the maximum Sharpe ratio of the opportunity set formed from cluster portfolios is somewhat smaller than the maximum Sharpe ratios of the opportunity set formed using characteristic-sorted portfolios. Nonetheless, these results suggest that forming benchmark assets using the past comovement of individual securities' returns can generate significant differences in future returns, without the potential data-snooping bias associated with the use of characteristics that have already been shown to be related to dispersion in returns.

We present the results of a battery of tests to gauge the statistical and economic advantages and weaknesses of this alternative method of estimating the investor's opportunity set. We show that the clustering procedure generates significantly smaller correlations across portfolios than characteristic-based sorting methods. As a consequence, the cluster portfolios' covariance matrix is better conditioned than the alternative portfolios. We examine the implications of the relative conditioning of the different sets of basis assets for asset pricing inferences. The relatively poor conditioning of the characteristic-sorted basis assets leads to efficient frontiers that are more sensitive to small changes in the data. In contrast, the cluster portfolios, with their better conditioning, generate inferences that are relatively insensitive to small perturbations in the data. In addition, measurement error in the data is associated with a positive bias in the Sharpe ratio, and we find that this bias is larger for basis assets (such as size-sorted portfolios) whose covariance matrix is relatively ill-conditioned.

Finally, we compare the inferences drawn from the cluster portfolios with those drawn from size- and book-to-market-sorted portfolios using two standard approaches in the empirical literature. First, we perform Gibbons, Ross, and Shanken (1989, GRS hereafter) tests to gauge the performance of the CAPM and the Fama and French (1993) three-factor model in describing the returns on the two sets of basis assets. The GRS specification test suggests that the models perform similarly using both the cluster and characteristic-sorted portfolios. However, we observe larger pricing errors, and a relatively poor model fit, in the individual cluster portfolios compared with the characteristic-sorted portfolios. That is, there appear to be some features of the cluster portfolios returns that cannot be explained by the CAPM or the three-factor model.

Second, we use cross-sectional regression tests to test the CAPM, the consumption CAPM (CCAPM), the three-factor model, and the conditional CAPM of Jagannathan and Wang (1996) on the cluster portfolios and characteristic-sorted portfolios. Generally, we find strong evidence that traditional risk measures, such as CAPM and CCAPM betas, are significantly and positively related to cluster portfolio returns; this result is robust to the inclusion of firm characteristics such as size and book-to-market. Furthermore, the conditional specification in Jagannathan and Wang (1996) describes a large portion of the cross-sectional variation in average returns on the cluster portfolios. However, the three Fama and French (1993) factors explain virtually none of this variation. In contrast, when size- and book-to-market-sorted portfolios are used as basis assets, only the characteristics and their factor-mimicking portfolio risk exposures are significantly related to returns; traditional beta risk measures are significantly negatively related to the returns of these portfolios when characteristics are included.

Despite the characteristics' lack of explanatory power in specification tests using the cluster portfolios, we show that the cluster portfolios exhibit significant cross-sectional differences in size and book-to-market, as well as CAPM beta. This evidence suggests two conclusions. First, it indicates that there is sufficient dispersion in the characteristics to enable a powerful test of the relation between cluster portfolio returns and these characteristics. Second, the evidence provides insight into the question of whether the characteristics are related to the comovement that determines the cluster grouping. The results suggest that the comovement that determines membership in a particular cluster is related to fundamental economic characteristics of the firm.

The remainder of the paper is outlined as follows. In Section 2, we discuss the theoretical reasoning underpinning the formation of basis assets for asset pricing tests. In addition, the specific clustering procedure following Ormerod and Mounfield (2000) is discussed. Section 3 describes the data that we use and examines the effects of differences in the conditioning of covariance matrices on test results. Section 4 examines the economic differences associated with different sets of basis assets when testing asset pricing models. Section 5 describes differences in the average characteristics of the cluster portfolios. Section 6 concludes.

2. Forming Basis Assets

2.1 Definition of the basis

The challenge that we consider is to find a set of portfolios that best characterizes an investor's opportunity set, consisting of all marketed securities as the basis assets. Practical constraints, including econometric problems, data availability, and computational resources, prevent researchers from considering the entire set of marketed claims. Instead, the opportunity set is reduced to a "representative set" of portfolios that seeks to approximate the investor's opportunity set as well as possible. In typical applications, researchers analyze between 10 and 100 portfolios of stocks in order to judge the opportunity set available to investors or to analyze pricing models.

Various rules for the division of assets in portfolios have been proposed in the literature. These rules are based on firm-specific characteristics that are hypothesized to be related to dispersion in expected returns, or on characteristics that have been shown to be related to dispersion in subsequent realized returns. Using characteristics, rather than variables that a particular theory implies are related to expected return, is appealing insofar as the procedure generates dispersion in returns (and therefore empirical power in tests of asset pricing models). The principal difficulty with using characteristics that are known to be related *ex post* to mean returns is that this procedure induces a data-snooping bias. In Lo and MacKinlay (1990), for example, the authors argue that one will always be able to find *ex post* deviations from a "true" asset pricing model and, moreover, that such biases will appear to be significant when they are considered in a group. Consequently, finding that firm characteristics are related *ex post* to average returns and then grouping firms into portfolios based on these characteristics may constitute a grouping of *ex post* deviations from a pricing model. Unfortunately, the magnitude of this bias is difficult to quantify in practice. MacKinlay (1995) presents evidence that suggests that this bias may be quite severe in the context of size and book-to-market portfolios.

Consequently, the decision about the choice of basis assets in an asset pricing test poses a significant conundrum for the researcher. On the one hand, the researcher could choose a set of variables that are *ex ante* related to expected returns on the basis of a theoretical model of asset prices. However, if the model is not correctly specified or returns are sufficiently noisy, these variables may have no significant relation to estimates of expected returns and, consequently, generate insufficient dispersion in returns for the empirical tests to have power. In contrast, sorting on the basis of characteristics known to be related to *ex post* returns generates dispersion in average returns, lending apparent power to the test of the asset pricing model. However, data-snooping issues limit the conclusions that one can draw from such tests. And, as a practical matter, the ability of some characteristics to produce dispersion in returns can vary substantially through time. For example, much of the well-known relation between size and returns appears to evaporate after 1985. A similar effect has occurred over the 1990s to the returns on book-to-market portfolios.

What, then, is the correct approach for identifying a basis for tests of asset pricing models, given that we wish to minimize the bias induced by searching over *ex post* average returns, while generating sufficient dispersion over these returns in order to generate empirical power? In this paper, we use a statistical method, cluster analysis, to generate a set of basis assets in which stocks should be highly correlated within groups, but have minimal correlation across groups. King (1966) argues that this criterion defines a set of basis assets well, and suggests that industry-sorted portfolios generate such a basis. Daniel and Titman (1997) use a similar argument to suggest that size and book-to-market do not represent risk exposures because the within-group covariation of these firms is not high.

2.2 Cluster analysis

The goal of cluster analysis is to reduce the dimensionality of a set of data by sorting individual observations into groups that are either similar (within the group) or different (across groups). "Similarity" and "difference" are calculated using some measure taken between the data points; for example, a Euclidean norm might be used as a distance measure. In our setting, we are particularly concerned with the covariance or correlation matrix of returns. Consequently, we specify a distance measure that is based on the correlation between the returns on two firms.

The intuition behind this measure is straightforward. Consider the problem of how a new security affects the menu of opportunities facing a hypothetical investor. Barring the trivial case in which this new security is perfectly correlated with an existing asset (or a combination of these assets), the new security will contribute something to the investor's set of choices. However, if the new asset is highly correlated with another security (or portfolio of securities), then grouping it with the highly correlated assets costs relatively little, and maintains a small number of portfolios. In contrast, an asset that has a low degree of correlation with other assets would add relatively more to the opportunity set and may warrant being placed in a separate portfolio.[2]

The distance measure d_{ij} suggested by Ormerod and Mounfield (2000) captures the intuition behind the use of the correlation coefficient well:

$$d_{ij} = \sqrt{2 * (1 - \rho_{ij})}, \tag{1}$$

[2] As an additional justification behind this criterion, we will analyze the contribution of the correlation structure of the assets to the stability of the covariance matrix. The stability of the covariance matrix is important since, as mentioned above, this matrix is a central object in most asset pricing analyses.

where ρ_{ij} denotes the sample correlation between the return on firms i and j.[3] Firms with perfectly correlated returns will be assigned a distance of 0 to each other, whereas perfectly negatively correlated firms are assigned the maximum distance of 2. Ormerod and Mounfield (2000) show that this function satisfies the conditions that are required to be used as a distance metric in a clustering algorithm.[4]

Once the initial distance measures are calculated, we must specify the method by which these distance measures are used to identify groups. We use Ward's minimum variance method (Ward 1963). In this method, one seeks to minimize the increase in the sum of squared errors generated when combining any two smaller clusters. The sum of squared errors for any cluster is the sum of the squared distances between each cluster member and the cluster centroid; it can also be calculated as the average (across cluster members) squared distance between all members of the cluster.

The algorithm for applying this distance measure is intuitive. Firms are initially each placed into their own individual clusters; thus, if there are N firms, the algorithm starts with N clusters. By definition, the sum of squared errors at this point is zero; each firm is its own centroid. The algorithm proceeds sequentially by optimally joining the individual firms, and later, groups of firms. That is, for every possible combination of smaller clusters i and j, the algorithm seeks to minimize the following:

$$D_{ij} = ESS(C_{ij}) - [ESS(C_i) + ESS(C_j)], \tag{2}$$

where $ESS(C_{ij})$ is the error sum of squares obtained in the new aggregate cluster, and $ESS(C_i)$, $ESS(C_j)$ are the error sum of squares for clusters i and j, respectively. Thus, Ward's method seeks to minimize the information loss, or the deterioration in fit, that occurs as clusters are combined. This procedure can be repeated until only two clusters remain; the researcher may stop the clustering process at any desired number of portfolios. In practice, we also analyze differences in results when the final number of clusters changes.

The clustering algorithm we use throughout the paper is designed to maximize within-group correlation, minimize across-group correlations, and thus reduce the off-diagonal terms in the correlation matrix of returns. Note the contrast between the cluster algorithm's focus on the correlation matrix, as opposed to the typical sorting method's focus on *ex post* mean returns. While

---

[3] The conditions required for a measure to be a proper or admissible distance metric rule out the use of covariance, although standardized measures of comovement, as Ormerod and Mounfield (2000) show, meet these conditions. Consequently, the distance metric we use in our analysis is based on correlation, rather than covariance.

[4] There are other distance measures and clustering algorithms that can be specified. For example, Brown, Goetzmann, and Grinblatt (1997) also form portfolios using a clustering method; they use the resulting portfolios as factors and find that these factors have relatively high explanatory power for industry returns, both in-sample and out-of-sample. However, the distance measure in their algorithm is related to the difference in observed mean returns, rather than comovement in returns. In our analysis of conditioning, we explore further the benefits of focusing on comovement in grouping firms into portfolios.

the focus on mean returns seems sensible, given the desirability of generating dispersion in returns, the covariance matrix is at the heart of much of the estimation performed in the asset pricing literature. Because of the sensitivity of inference to the properties of the covariance matrix, we suggest that the assets constructed by our algorithm may possess certain advantages. In particular, we suggest that the covariance matrix resulting from our clustering procedure is better conditioned, i.e., has a lower condition number. Appendix A describes the general implications of better conditioning for inferences related to asset pricing models, and we investigate the specific consequences of the differences in conditioning across cluster and characteristic-sorted portfolios in Section 3.2.

3. Data

Our starting point for analysis is all CRSP-covered firms with common shares outstanding over the period 1955–2003. We are particularly interested in comparing the clustering methodology, and the portfolios generated, to characteristic-based sorts. Consequently, we reduce the set of firms according to the availability of data for the characteristics. In particular, we follow the procedures outlined by Fama and French (1993) and Fama, Davis, and French (2000) for defining the set of firms to be covered. More specifically, we analyze the intersection of CRSP and Compustat data where firms' book values as of June of the portfolio formation year are available. To avoid Compustat bias issues, firms are included in the sample only if they have been covered by Compustat for at least two years.

At each time t we start with a set of individual firms' return data covering the months $t - 60$ through $t - 1$. For the calculation of betas (for the characteristic-sorted portfolios) and correlations (for the clustering algorithm), we require that a firm has a minimum of 36 months of returns data available in this period. The correlation matrix of the returns on the firms over this time period is computed, and the distance measures from Equation (1) are calculated from these correlations. In each subperiod, we trim the extreme 2.5% of distance measures (and the corresponding firms) because the clustering algorithm tends to bias toward maintaining these firms in their own clusters.[5] The clusters are then determined by using these distance measures and the algorithm described above. Using this cluster assignment, we form a value-weighted portfolio return of the securities in that cluster for months t through $t + 11$. Thus, all our analysis will be conducted on returns that are "out of sample" relative to the period from which the clusters are formed. At the end of the month $t + 11$, we roll the entire analysis forward by one year and continue throughout the entire sample period.

[5] For one randomly chosen subperiod, we examined the proportion of firms removed from the sample after trimming 2.5% of the pairwise distance measures. The trimming based on distance removed approximately 2.5% of firms.

We form portfolios of 10 and 25 clusters using the algorithm described in the previous section.

One issue that arises in this procedure is that there need be no time consistency in the cluster numbers we use as identifiers. That is, if 25 clusters are formed from 1965 to 2000, there is no requirement that "Cluster 1" in 1965 be related to "Cluster 1" in 1966. Although the clustering procedure minimizes within-group distances and maximizes across-group distances, the cluster number itself has no intrinsic meaning. In order to add some structure, we impose an auxiliary criterion. In each year τ, we compute the similarity in member firms of cluster j to all clusters i formed in year $\tau - 1$, where similarity is defined as the number of firms common to the cluster in each year. We assign to cluster j in year τ the index variable associated with the most similar cluster i at $\tau - 1$. As a result, across adjacent years each cluster i will have the most consistent firm membership possible through our sample period. Clearly, this criterion is not the only possible method for ranking the clusters; moreover, the index number associated with a cluster does not affect its composition in any way.[6] However, the procedure assures some time consistency in the cluster definitions without losing the *ex ante* spirit of the portfolio formation exercise.

3.1 Descriptive statistics

Summary statistics for 10 and 25 cluster portfolios are presented in Tables 1 and 2, respectively. In Table 1, Panel A, we see that the clustering method results in significant dispersion in the means and the standard deviations of the resulting portfolios. The means vary from 0.87% per month to 1.34% per month, generating a dispersion in mean returns of 47 basis points per month; standard deviations range from 5.03% per month to 8.11% per month, with an average of 6.01%. The average number of securities in each of the ten portfolios is reasonably large; the smallest number of securities in any portfolio during any subperiod is 25, and the average number of securities in each portfolio is 294. There is a tendency for the number of securities in each portfolio to increase through the sample period, corresponding to an increase in the overall sample; the number of firms increases from 906 to 5242 through the sample period.

In Table 1, Panel B, we present the correlation matrix for the returns of the ten cluster portfolios. The algorithm generates portfolios that have relatively low cross-correlations. Although the clustering algorithm is designed to group highly correlated securities, it is important to realize that the result in Panel B is not guaranteed, since the clustering algorithm uses historical returns, while the correlations presented in Panel B are for returns subsequent to those used by the clustering algorithm. Given the relatively large number of securities in the

[6] As a robustness check on this auxiliary criterion, we have also used Sharpe ratio, volatility, and the cluster number initially assigned by the clustering algorithm as identifying variables for clusters as we move through the sample period; our results are qualitatively similar. That is, for each method we generate roughly similar dispersion in returns, and lower cross-correlations in returns, compared with characteristic-sorted portfolios.

Table 1
Descriptive statistics: 10 portfolios

Panel A: Means and standard deviations

	Means				Standard deviations			
	Cluster	MV	BM	Beta	Cluster	MV	BM	Beta
1	1.184	1.399	0.819	0.944	8.111	6.604	5.061	3.647
2	1.060	1.304	0.995	0.997	5.025	6.144	4.756	3.762
3	1.189	1.243	0.906	0.979	5.394	5.734	4.685	4.012
4	0.980	1.227	0.985	0.971	5.185	5.593	4.654	4.365
5	1.006	1.211	1.101	1.051	6.665	5.266	4.431	4.592
6	0.872	1.083	1.070	1.010	6.177	5.028	4.308	4.997
7	1.323	1.112	1.151	0.946	6.546	4.978	4.292	5.248
8	1.342	1.100	1.170	0.867	5.750	4.828	4.361	5.850
9	1.242	0.977	1.308	0.906	5.871	4.465	4.595	6.700
10	1.262	0.892	1.361	0.960	5.413	4.285	5.090	8.231

Panel B: Correlation matrix

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	1.00									
C2	0.29	1.00								
C3	0.39	0.31	1.00							
C4	0.47	0.55	0.62	1.00						
C5	0.59	0.36	0.45	0.49	1.00					
C6	0.48	0.44	0.46	0.53	0.50	1.00				
C7	0.47	0.45	0.51	0.59	0.49	0.44	1.00			
C8	0.40	0.49	0.52	0.63	0.38	0.48	0.58	1.00		
C9	0.65	0.44	0.52	0.62	0.68	0.56	0.57	0.54	1.00	
C10	0.51	0.51	0.58	0.66	0.52	0.52	0.60	0.63	0.64	1.00

Panel C: Higher moments and comoments

	Betas				Skewness			
	Cluster	MV	BM	Beta	Cluster	MV	BM	Beta
1	1.117	1.132	1.064	0.319	2.085	0.157	−0.236	−0.475
2	0.806	1.169	1.053	0.558	0.457	−0.193	−0.425	−0.092
3	0.900	1.163	1.005	0.722	0.524	−0.314	−0.510	−0.065
4	1.014	1.155	1.015	0.864	0.095	−0.441	−0.385	−0.231
5	1.092	1.126	0.927	0.996	1.140	−0.416	−0.415	−0.363
6	1.030	1.102	0.933	1.127	1.051	−0.461	−0.209	−0.348
7	1.002	1.106	0.923	1.264	1.225	−0.347	0.070	−0.258
8	0.995	1.058	0.909	1.429	0.412	−0.365	−0.108	−0.310
9	1.086	1.016	0.947	1.652	0.254	−0.330	−0.041	−0.303
10	1.032	0.938	0.995	2.098	0.250	−0.327	−0.047	−0.047

	Kurtosis				Coskewness			
	Cluster	MV	BM	Beta	Cluster	MV	BM	Beta
1	17.611	3.616	1.425	1.889	−1.614	−1.466	0.093	−0.531
2	3.767	3.336	1.630	1.240	0.765	−1.145	−0.341	−0.014
3	6.191	2.810	2.298	1.778	−0.575	−0.879	−0.500	0.294
4	1.761	2.744	2.448	1.474	−0.546	−0.948	−0.412	0.221
5	7.838	3.011	2.937	1.959	−1.001	−0.733	−0.536	−0.073
6	6.556	2.592	1.803	1.931	−0.064	−0.546	0.090	−0.155
7	8.661	2.586	2.256	1.715	−2.003	−0.238	0.217	0.068
8	3.479	1.836	2.337	1.583	−1.072	−0.199	0.078	−0.277
9	3.044	1.569	2.841	1.478	−0.265	0.030	0.044	−0.502
10	2.956	1.543	4.479	1.426	−0.975	−0.038	−0.383	−0.178

(*continued overleaf*)

Table 1
(Continued)

Panel D: Sharpe ratios

Cluster	MV	BM	Beta
0.196	0.180	0.239	0.184

This table presents descriptive statistics for 10 portfolios formed on various dimensions. The column labeled "Cluster" represents portfolios formed on the basis of clustering of the correlation matrix over the prior 60 months, "MV" represents portfolios formed on market value, "BM" represents portfolios formed on book-to-market ratios, and "Beta" represents portfolios formed on the basis of CAPM betas. The cluster, market value, book-to-market ratio, and beta are measured as of June of each year. Data are formed from CRSP and Compustat data, and require that each firm have a valid market value, book-to-market ratio, and CAPM beta at the time of portfolio formation. Panel A presents means and standard deviations for these four sets of portfolios. Panel B presents the correlation matrix of the cluster portfolios. Panel C reports additional summary statistics for the returns; displayed are the skewness, kurtosis, and beta and coskewness as measured with respect to the value-weighted CRSP index. Coskewness is measured as the slope coefficient from regressing the portfolio return on the return on the market portfolio squared. Panel D presents the maximum Sharpe ratios of efficient portfolios formed from the assets. Data are sampled monthly from July 1959 through December 2003.

portfolios, it is surprising how low some of these correlations are; for example, the correlation between portfolios 1 and 10 is only 0.51, and no correlations exceed 0.68. The average pairwise correlation is 0.51.

As a point of comparison, we also present summary statistics for portfolios formed on three widely used firm characteristics: the book-to-market ratio, the market capitalization of equity, and the market beta. Firms are sorted into decile portfolios on the basis of either size or book-to-market ratio at the end of June of each year; firms are sorted into decile portfolios based on beta estimates calculated over the immediately preceding 60-month period. Summary statistics for these portfolios are also presented in Table 1, Panel A. The dispersion in mean returns generated by these portfolios is similar to that generated by the clustering algorithm. The difference in mean returns on high and low book-to-market portfolios is 0.54% per month, whereas the difference in mean returns on small and large capitalization stocks is 0.51% per month. The dispersion in beta-sorted portfolio returns is substantially smaller, at 18 basis points per month. Furthermore, these portfolios exhibit appreciably higher interportfolio correlations than that of the cluster portfolios; the average correlations for the book-to-market, size portfolios, and beta portfolios are 0.84, 0.89, and 0.80, respectively.

We also report skewness, beta, kurtosis, and coskewness measures for the cluster portfolios, as well as the characteristic-sorted portfolios, in Table 1, Panel C. The cluster portfolios tend to have positive skewness, in contrast to all three sets of characteristic-sorted portfolios. In addition, the magnitude of the skewness is somewhat larger. In terms of coskewness, there are no substantive differences between any of the four sets of portfolios.

In Table 2, we present the results after forming 25 cluster portfolios. When a larger number of clusters are formed, there is a marked increase in the dispersion in mean returns. These results are displayed in Panel A, and indicate a minimum mean return of 0.80% and a maximum of 1.60%, for a dispersion in

Table 2
Summary statistics: 25 portfolios

Panel A: Means and standard deviations

Mean				Standard deviation			
C1	1.002	S1B1	0.869	C1	5.861	S1B1	8.103
C2	1.253	S1B2	1.235	C2	6.770	S1B2	6.861
C3	1.186	S1B3	1.335	C3	6.939	S1B3	6.044
C4	1.077	S1B4	1.530	C4	5.893	S1B4	5.664
C5	1.229	S1B5	1.560	C5	6.333	S1B5	6.037
C6	1.412	S2B1	0.989	C6	6.922	S2B1	6.955
C7	1.275	S2B2	1.135	C7	7.157	S2B2	5.999
C8	1.193	S2B3	1.273	C8	6.002	S2B3	5.293
C9	0.893	S2B4	1.420	C9	6.304	S2B4	5.122
C10	1.348	S2B5	1.404	C10	6.874	S2B5	5.549
C11	1.176	S3B1	0.997	C11	5.946	S3B1	6.314
C12	1.176	S3B2	1.069	C12	6.478	S3B2	5.281
C13	0.804	S3B3	1.136	C13	6.343	S3B3	4.800
C14	1.257	S3B4	1.212	C14	5.975	S3B4	4.771
C15	1.028	S3B5	1.388	C15	4.907	S3B5	5.265
C16	1.207	S4B1	0.983	C16	7.117	S4B1	5.775
C17	1.176	S4B2	0.962	C17	6.362	S4B2	5.087
C18	1.417	S4B3	1.179	C18	8.125	S4B3	4.773
C19	1.429	S4B4	1.283	C19	6.527	S4B4	4.591
C20	1.228	S4B5	1.362	C20	7.384	S4B5	5.148
C21	1.595	S5B1	0.882	C21	8.555	S5B1	4.764
C22	1.114	S5B2	0.893	C22	6.357	S5B2	4.518
C23	1.444	S5B3	1.034	C23	7.659	S5B3	4.323
C24	0.900	S5B4	1.027	C24	7.026	S5B4	4.250
C25	1.236	S5B5	1.159	C25	6.677	S5B5	4.613

Panel B: Sharpe ratios

Cluster	SZBM
0.241	0.407

This table presents summary statistics for 25 portfolios formed on two criteria. The portfolios labeled “C1” through “C25” represent portfolios based on the clustering of the correlation matrix measured over the past 60 months. Rows labeled “S1B1” through “S5B5” represent portfolios formed on the intersection of size and book-to-market ratio quintiles. Firms are assigned to cluster and size/book-to-market portfolios as of the end of June of each year. Data are formed using CRSP and Compustat, and require that each firm have a valid market value, book-to-market ratio, and beta at the time of portfolio formation. Panel A presents the means and the standard deviations of the portfolios, and Panel B presents maximum Sharpe ratios from the efficient frontier constructed using these portfolios. Data are sampled at the monthly frequency from July 1959 through December 2003.

mean returns of 79 basis points per month (the difference is due to rounding). Thus, increasing the number of clusters to 25 yields more than a 30 basis point increase in the spread in average monthly returns. This increased dispersion in mean returns comes at the cost of somewhat higher standard deviation, however. The standard deviation of returns in the cluster portfolios ranges from 4.91% to 8.56%. The pairwise correlations across returns in the set of portfolios (not shown) decrease; the average correlation falls to 0.42. For comparison, we also compute summary statistics on a set of 25 size- and book-to-market-sorted portfolios constructed as in Fama and French (1993). The 25 portfolios yield a spread of 0.69% per month, ranging from a low of 0.87% per month to a high of 1.56% per month. The average pairwise correlation across these portfolios, 0.81, is comparable to the sets of 10 portfolios sorted on individual characteristics and substantially higher than the correlations among the 25 cluster portfolios. In other (unreported) comparisons, we find that, as in Table 1, the 25 cluster portfolio returns are positively skewed, whereas the size- and book-to-market-sorted portfolios are negatively skewed. In addition, there are no significant differences in coskewness between the two sets of portfolios.

In Panel D of Table 1 and Panel B of Table 2, we present the maximum Sharpe ratios for the mean-variance frontier generated by each of these sets of basis assets. For both sets of cluster portfolios, 10 and 25, we see that the higher standard deviation in the individual portfolios, particularly relative to the book-to-market-sorted portfolios, results in a lower Sharpe ratio for the mean-variance frontier, compared with the characteristic-sorted portfolios. For the ten cluster portfolios, the Sharpe ratio of 0.196 is lower than the book-to-market portfolios' Sharpe ratio of 0.239, although it is slightly higher than the size- and beta-sorted portfolios' Sharpe ratios of 0.180 and 0.184, respectively. The Sharpe ratio of the 25 cluster portfolios, at 0.241, is also substantially lower than the ratio observed in the Fama–French portfolios of 0.407.[7]

Overall, the clustering algorithm is able to generate significant dispersion in expected returns. When ten portfolios are formed, the dispersion is roughly comparable with that observed in portfolios sorted along size or book-to-market characteristics, and higher than that generated by sorting on beta. When 25 cluster portfolios are formed, the dispersion is higher than that observed in size- and book-to-market-sorted portfolios, although this higher dispersion is accompanied by a substantially higher standard deviation of returns.

Finally, we calculate the condition numbers of the covariance matrices of the ten cluster portfolios; the ten book-to-market, size-, and beta-sorted portfolios; the 25 cluster portfolios; and the 25 size- and book-to-market-sorted portfolios. The set of ten cluster portfolios has a condition number of 26; the condition numbers of the (ten) size-sorted, book-to-market-sorted, and beta-sorted portfolios are all larger, at 548, 138, and 159, respectively, indicating that they are less well-conditioned. Comparing the two sets of 25 basis assets, the condition number of the 25 portfolios sorted on the basis of size and book-to-market is 671, whereas the condition number of the 25 cluster portfolios is 50—again, the cluster algorithm generates a better conditioned covariance matrix.

To interpret these differences in condition numbers, we examine both asymptotic p-values and bootstrapping results. Edelman and Sutton (2005) derive asymptotic results for the tails of the distribution of (the square root of the) condition numbers of Wishart matrices that we have calculated above. Using their results, none of the condition numbers reported above are significantly different at conventional levels. However, when we compare these condition

[7] When short sales are precluded, the performance of the cluster portfolios improves further relative to the characteristic-sorted portfolios. Specifically, the maximum constrained Sharpe ratios of portfolios of the ten size-, book-to-market-, and beta-sorted portfolios are 0.143, 0.178, and 0.142, respectively, compared with 0.171 for the ten cluster-sorted portfolios. The constrained Sharpe ratio of the 25 size and book-to-market portfolios is 0.187, compared with a constrained Sharpe ratio of 0.193 for the 25 cluster portfolios.

Table 3
Simulated Sharpe ratios

Panel A: 10 Portfolios

		$\sigma_\mu = 1e-4$			$\sigma_\mu = 5e-4$			$\sigma_\mu = 1e-3$		
Portfolio	Actual	Mean	Med.	Std.	Mean	Med.	Std.	Mean	Med.	Std.
Cluster	0.1960	0.1962	0.1961	0.0019	0.1999	0.1994	0.0095	0.2110	0.2095	0.0190
MV	0.1801	0.1831	0.1826	0.0084	0.2426	0.2383	0.0420	0.37108	0.3624	0.0856
BM	0.2386	0.2395	0.2395	0.0044	0.2597	0.2581	0.0227	0.3136	0.3092	0.04875
Beta	0.1844	0.1854	0.1854	0.0043	0.2080	0.2067	0.0222	0.2655	0.2620	0.0461

Panel B: 25 Portfolios

		$\sigma_\mu = 1e-4$			$\sigma_\mu = 5e-4$			$\sigma_\mu = 1e-3$		
Portfolio	Actual	Mean	Med.	Std.	Mean	Med.	Std.	Mean	Med.	Std.
Cluster	0.2408	0.2411	0.2410	0.0016	0.2466	0.2463	0.0082	0.2629	0.2622	0.0164
SZBM	0.4074	0.4088	0.4088	0.0061	0.4441	0.4431	0.0304	0.5400	0.5379	0.0605

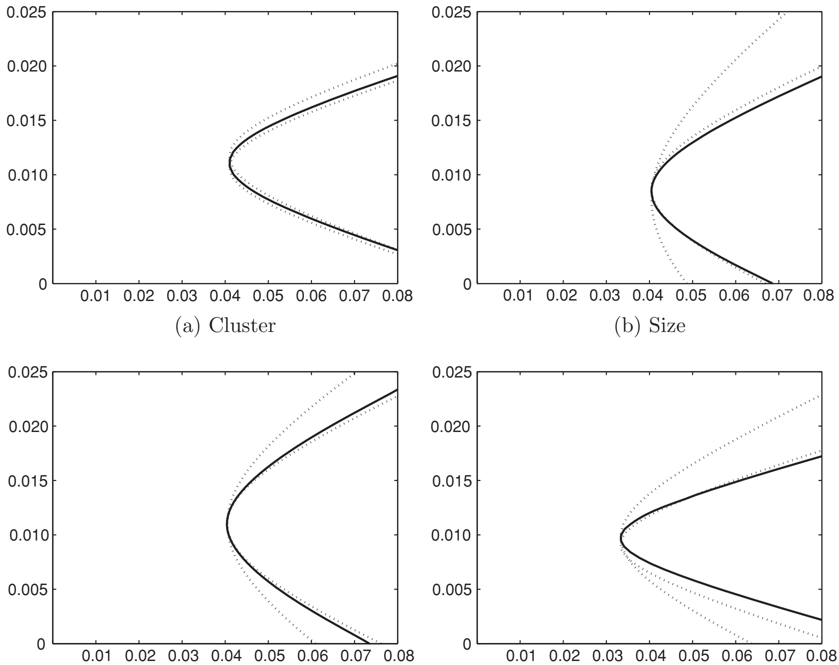
This table presents results of simulations of the efficient frontier represented by various sets of portfolios. It presents the mean, the median, and the standard deviation of maximum Sharpe ratios obtained by perturbing the mean return vector of cluster, market value, size, book-to-market, and beta-sorted portfolios. The perturbation is distributed normally with mean zero and standard deviation $\sigma_\mu = \{1e-4, 5e-4, 1e-3\}$. Panel A presents results for sets of 10 portfolios formed by clustering and on market value, book-to-market, and beta. Panel B presents analogous results for sets of 25 portfolios formed on clusters and size and book-to-market. Results are derived from 5000 simulated sets of portfolios.

numbers with those obtained from bootstrapping (without replacement) the individual securities into 10 (25) randomly chosen value-weighted portfolios, we find that the characteristic-sorted portfolios have strikingly large condition numbers, with empirically observed p-values less than 0.05, while the cluster portfolios have empirically observed p-values of greater than 0.95.[8] Consequently, in the next section, we explore the specific effect that these differences in conditioning have on inferences with respect to Sharpe ratios (or efficient frontiers) formed by these different sets of basis assets.

3.2 Consequences of conditioning: Sharpe ratios and efficient frontiers

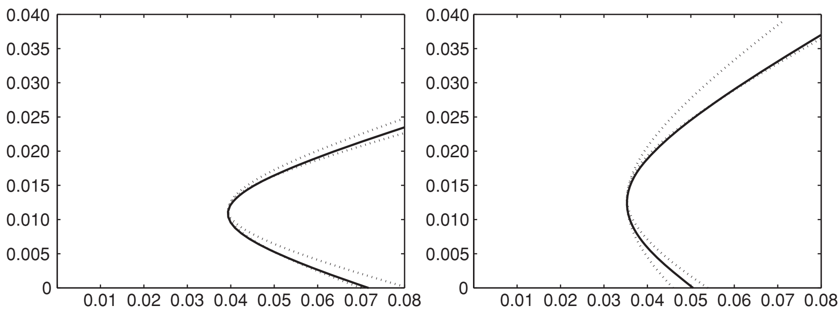
We analyze the impact of the differences in the conditioning of the covariance matrix of characteristic-sorted and cluster portfolios on applications in asset pricing by performing a simulation experiment similar to that analyzed by MacKinlay (1995). The simulation experiment adds random pricing errors of varying magnitude to the mean returns on the two different sets of basis assets and examines the resulting changes in the investor's opportunity set; the details of this simulation are described in Appendix B. In Table 3, we present means, medians, and standard deviations of maximum Sharpe ratios for three values of the standard deviation of the measurement error. Panel A presents results for sets of 10 portfolios, and Panel B presents results for sets of 25 portfolios.

[8] These results are available on request from the authors. An alternative approach is to use the asymptotic results of Edelman and Sutton (2005) to determine these p-values; differences between calculated p-values and those of our bootstrap exercise suggest that the asymptotic assumption does not hold well in this case.



(a) Cluster (b) Size (c) Book-to-Market (d) Beta

Figure 1
Distribution of efficient frontiers: 10 portfolios



(a) Cluster (b) Size and book-to-market

Figure 2
Distribution of efficient frontiers: 25 portfolios

Additionally, we present the efficient frontiers corresponding to the data, the 5th and 95th percentiles of simulated Sharpe ratios in Figures 1 and 2.

It is apparent from Panel A of Table 3 that an increase in the standard deviation of the measurement error (σ_μ) increases the variation in the Sharpe ratios of all sets of basis assets. However, the range is always much larger for the characteristic-sorted portfolios. Furthermore, the volatility in Sharpe ratios is always largest for the size-ranked portfolios, which have a much larger condition number than the other characteristic-sorted portfolios, and smallest for the cluster portfolios, which have the smallest condition number. That is, the condition number acts as an amplifier to the noise with which the mean return is measured.

The effect of the measurement error on the simulated efficient frontiers is also presented in Figure 1. The 5th and 95th percentiles for the cluster portfolios from the simulations (the dashed frontiers) plot almost on top of the empirically observed (solid) efficient frontier. However, the same bounds for the size and book-to-market portfolios are always larger, corresponding to much greater variation in the (estimated) opportunity set that investors face. This variation is associated with the larger condition number of the covariance matrix of characteristic-sorted portfolio returns. For example, note that the range of efficient frontiers varies more for the single-sorted size portfolios; recall that these portfolios are associated with a relatively high condition number of 548. The book-to-market portfolios, with their condition number of 138, generate less variable frontiers. Clearly, conditioning matters when making inferences about the investor's opportunity set.

Panel B of Table 3 and Figure 2 present similar results for sets of 25 portfolios. The standard deviation of Sharpe ratios for the 25 size- and book-to-market-sorted portfolios is generally more than three times the standard deviation of the Sharpe ratios for the cluster portfolios for all three values of the standard deviation of measurement errors. Figure 2 also shows the importance of conditioning in the sensitivity of the efficient frontier to measurement errors in mean returns. Intuitively, measurement error in the mean returns gets translated, through the condition number of the covariance matrix, to higher variability in the investor's opportunity set.

Importantly, it is not just the volatility of the efficient frontier that is affected by measurement error in the data. The higher condition numbers of the characteristic-sorted portfolios are associated with an increase in the average Sharpe ratio in the simulations as the measurement error increases. For example, in Table 3, Panel A, note that as σ_μ, or the standard deviation in the measurement error, increases from 1 to 5 basis points, the average Sharpe ratio increases by 37 basis points for the cluster portfolios. In contrast, the average Sharpe ratio increases by 202 basis points (from 0.2395 to 0.2597) for book-to-market-sorted portfolios, and by 595 basis points (from 0.1831 to 0.2426) for the ten size-ranked portfolios. Intuitively, since the Sharpe ratio is the maximum price of risk in the sample, the higher volatility in the opportunity set results in a higher Sharpe ratio for covariance matrices which are more ill-conditioned. The upward bias in the average Sharpe ratio across the 5000 simulations is seen clearly in Figure 3. In this figure, we present the relation between measurement error and the average Sharpe ratio across 5000 simulations for the four

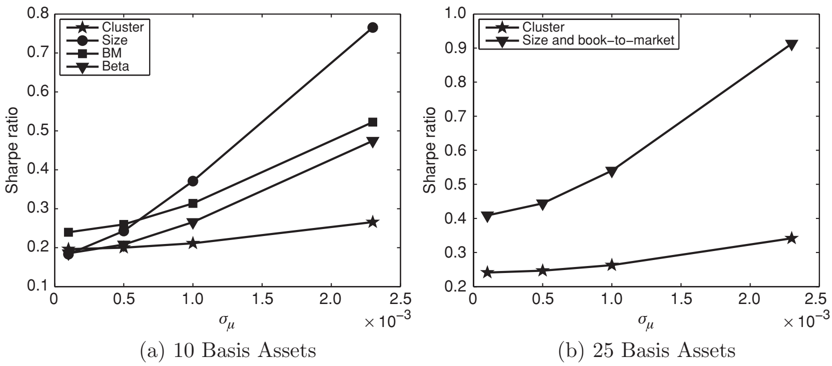


Figure 3
Sharpe ratio bias

sets of ten basis assets. Note that the sensitivity of the average Sharpe ratio to measurement error (or the slope of the line) is the largest for the size-sorted portfolios, which are the most ill-conditioned set of basis assets, and lowest for the cluster portfolios, which have the lowest condition number of all four sets of portfolios.

These results are related to those of MacKinlay (1995). In his paper, he examines the effects of two types of model errors: risk-based factors (i.e., missing factors) and non-risk-based factors, which he describes as data snooping, market frictions, or market irrationalities. Measurement error in the data would fall in the second, non-risk-based category. MacKinlay (1995) notes that the distribution of the test statistic for squared Sharpe ratios follows a noncentral F-distribution for both risk-based and non-risk-based factors. However, there is an upper bound on the noncentrality parameter for risk-based factors, which is due to a link between the magnitude of the excess returns and their volatility. For non-risk-based factors, there is no such link, and hence no upper bound on the noncentrality parameter. In his sample, he shows that the differences between these two types of distributions are large; non-risk-based model errors can lead to very large Sharpe ratios.

Our results indicate that the differences in the distributions, and hence the expected difference in the Sharpe ratios, are particularly large when the covariance matrix of the test assets is ill-conditioned. To demonstrate this, we compare the squared Sharpe ratio observed for each set of basis assets with a null hypothesis in which the CAPM holds, and an alternative hypothesis where deviations are caused by non-risk-based factors. Following MacKinlay (1995), we assume that the test statistic for the squared Sharpe ratio under the alternative hypothesis is drawn from a noncentral F-distribution with a noncentrality parameter given by

$$\lambda = T\,(\hat{\mu} - \mu)'\,\Sigma^{-1}\,(\hat{\mu} - \mu)\,, \tag{3}$$

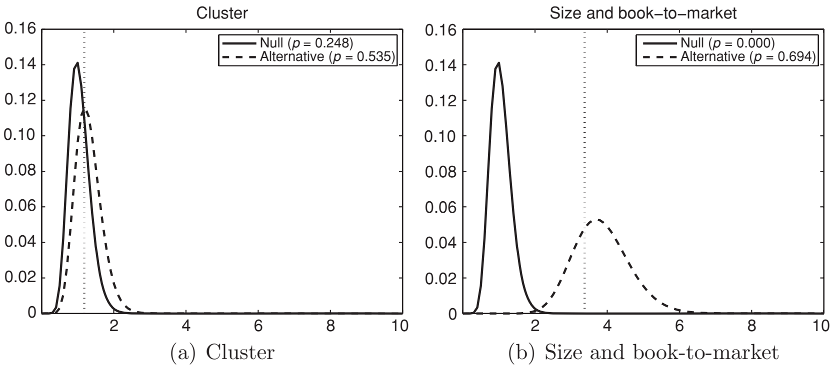


Figure 4
Distribution of squared Sharpe ratios: 25 portfolios

where $(\hat{\mu} - \mu)$ is equivalent to the simulated measurement error in our experiment (and so would have standard deviation equal to σ_μ). The simulated distributions under the null and alternative hypotheses are plotted in Figure 4 and correspond to MacKinlay's (1995) Figure 1, in which he shows the rightward shift in the F-distribution compared with the null for deviations due to risk-based and non-risk-based factors.

For brevity, we present results only for the sets of 25 portfolios. As shown in the figure, the Sharpe ratio of the tangency portfolio generated by the cluster portfolios is not significantly different from the null hypothesis, while the Sharpe ratio of the size and book-to-market portfolios is quite high compared with the distribution under the null. However, when evaluated under the non-risk-based alternative (associated with our simulated measurement error), the difference in the inferences regarding the portfolios' squared Sharpe ratios varies dramatically across the two types of basis assets. Measurement errors, market frictions, irrationalities, and data-snooping biases have only a small effect on the distribution of Sharpe ratios, and hence on inferences, when the covariance matrix of returns is relatively well-conditioned. In contrast, the distribution of the squared Sharpe ratio for the size and book-to-market portfolios shifts far to the right. Moreover, the observed squared Sharpe ratio of the size- and book-to-market-sorted portfolios is not statistically different than zero under the alternative distribution, with a p-value of 0.69. As a consequence, a Sharpe ratio that was considered very unlikely under the null is not startling at all in the presence of small measurement errors and a poorly conditioned covariance matrix. That is, an ill-conditioned covariance matrix combined with even economically small errors in the returns data could easily generate a spuriously large Sharpe ratio.

In the simulation results discussed above, we have assumed that the measurement error is uncorrelated across portfolios. In Appendix B, we discuss other experiments where the measurement error is assumed to be correlated across portfolios. The results suggest that positively correlated measurement errors reduce the advantages of better conditioning, although, except in extreme cases, conditioning continues to affect the precision of inferences and potential bias in the Sharpe ratio. Negatively correlated measurement errors (which might obtain if, as Lo and MacKinlay (1990) suggest, characteristic-sorted portfolios also contain a sorting on measurement errors) increase the advantages of better conditioning. These results demonstrate that the sources of, and correlation in, measurement errors are an important determinant of the consequences of ill-conditioning.

Overall, these results indicate that the outside observer's inferences regarding the ability of characteristic-sorted portfolios to capture the opportunity set of investors are more sensitive to measurement errors in the vector of mean returns. Ill-conditioning in the basis assets serves to magnify the effect of measurement error in returns in the generation of the efficient frontier. Importantly, measurement error can affect both the precision with which the opportunity set is measured and the maximum Sharpe ratio associated with the opportunity set.

4. Tests of Asset Pricing Models

In the previous section, we discussed the statistical properties of the cluster portfolios and showed that conditioning affects the precision with which the opportunity set and, consequently, properties such as the Sharpe ratio are measured. In this section, we investigate the economic differences in the cluster portfolios compared with characteristic-sorted portfolios when testing asset pricing models.

4.1 Multivariate tests

We first perform multivariate specification tests of two asset pricing models using the characteristic and cluster portfolios: the Sharpe–Lintner CAPM and the Fama and French (1993) three-factor model. Results of these tests are presented in Tables 4–7. Results for the CAPM are standard. The GRS test indicates that the CAPM is rejected using the ten cluster, size-sorted, and book-to-market-sorted portfolios with low p-values, but fails to reject the CAPM with a p-value of 0.273 using the beta-sorted portfolios. In contrast, the GRS tests fail to reject the three-factor model for the cluster portfolios, beta-, size-, and book-to-market-sorted portfolios. Thus, multivariate model specification inferences using the cluster- and the characteristic-sorted portfolios are similar.

Although the GRS test results suggest that all four sets of assets have similar implications for the validity of these pricing models, it is also instructive to examine the results for individual portfolios. In Table 5, the estimates of three-factor regression intercepts for the size-sorted and book-to-market-sorted portfolios are very small. These intercepts range from −5 to 8 basis points for the size-sorted and −6 to 14 basis points for the book-to-market

Table 4
CAPM specification tests: 10 assets

Port.	α	β	R^2	Port.	α	β	R^2	Port.	α	β	R^2	Port.	α	β	R^2
C1	0.19	1.17	40.84	M1	0.43	1.12	56.46	B1	−0.13	1.05	85.68	BT1	0.21	0.59	51.55
	(0.27)	(0.06)			(0.19)	(0.04)			(0.08)	(0.02)			(0.11)	(0.03)	
C2	0.28	0.69	37.41	M2	0.31	1.16	70.39	B2	0.07	1.02	91.04	BT2	0.23	0.68	64.32
	(0.17)	(0.04)			(0.15)	(0.03)			(0.06)	(0.01)			(0.10)	(0.02)	
C3	0.33	0.87	51.13	M3	0.27	1.12	75.21	B3	−0.02	1.00	90.42	BT3	0.16	0.79	76.03
	(0.17)	(0.04)			(0.13)	(0.03)			(0.06)	(0.01)			(0.09)	(0.02)	
C4	0.09	0.93	63.69	M4	0.25	1.12	78.72	B4	0.08	0.96	84.60	BT4	0.10	0.89	82.04
	(0.14)	(0.03)			(0.11)	(0.03)			(0.08)	(0.02)			(0.08)	(0.02)	
C5	0.04	1.10	53.58	M5	0.26	1.07	81.85	B5	0.23	0.90	80.62	BT5	0.15	0.96	85.51
	(0.20)	(0.04)			(0.10)	(0.02)			(0.09)	(0.02)			(0.08)	(0.02)	
C6	−0.04	0.98	50.25	M6	0.14	1.06	87.60	B6	0.20	0.88	83.12	BT6	0.06	1.06	88.29
	(0.19)	(0.04)			(0.08)	(0.02)			(0.08)	(0.02)			(0.08)	(0.02)	
C7	0.40	1.02	47.71	M7	0.16	1.06	90.25	B7	0.30	0.85	77.51	BT7	−0.03	1.12	90.25
	(0.21)	(0.05)			(0.07)	(0.02)			(0.09)	(0.02)			(0.07)	(0.02)	
C8	0.46	0.92	51.06	M8	0.16	1.04	92.45	B8	0.32	0.85	75.57	BT8	−0.16	1.24	88.27
	(0.18)	(0.04)			(0.06)	(0.01)			(0.09)	(0.02)			(0.09)	(0.02)	
C9	0.29	1.08	66.60	M9	0.07	0.97	93.42	B9	0.44	0.88	72.32	BT9	−0.19	1.39	85.02
	(0.15)	(0.03)			(0.05)	(0.01)			(0.11)	(0.02)			(0.11)	(0.03)	
C10	0.36	0.96	61.56	M10	0.00	0.93	93.94	B10	0.45	0.97	72.25	BT10	−0.26	1.66	80.55
	(0.15)	(0.03)			(0.05)	(0.01)			(0.12)	(0.03)			(0.16)	(0.04)	
	GRS: 2.834 (0.002)				GRS: 3.384 (0.000)				GRS: 3.025 (0.001)				GRS: 1.223 (0.273)		

This table presents specification tests for four sets of 10 portfolios. We conduct specification tests based on the regression specification:

$$r_{it} - r_f = \alpha_i + \beta_i (r_{mt} - r_f) + \epsilon_{it},$$

where r_{it} is the return on portfolio i; r_f is the risk-free rate, measured as the return on the T-bill closest to one month to maturity; and r_{mt} is the return on the value-weighted CRSP index. The parameters α_i and β_i are estimated via least squares, and the restriction $\alpha = 0$ is tested using the GRS specification test. The table presents the least squares coefficient estimates of α and β and accompanying standard errors, along with regression R^2. The test statistics and p-values for the specification tests are presented below the coefficient estimates. Tests are conducted on four separate sets of assets: cluster, size, book-to-market, and beta-sorted portfolios. Data are sampled at the monthly frequency from July 1959 through December 2003.

Table 5
Three-factor specification tests: 10 assets

	α	β	γ	δ	R^2		α	β	γ	δ	R^2
C1	0.292	0.967	0.698	−0.229	48.995	M1	0.070	1.042	1.081	0.528	81.747
	(0.260)	(0.063)	(0.084)	(0.097)			(0.127)	(0.031)	(0.041)	(0.047)	
C2	0.013	0.770	0.189	0.446	43.214	M2	0.011	1.096	0.915	0.445	91.343
	(0.169)	(0.041)	(0.055)	(0.063)			(0.081)	(0.020)	(0.026)	(0.030)	
C3	0.255	0.909	−0.036	0.131	51.625	M3	−0.016	1.079	0.771	0.430	92.832
	(0.168)	(0.041)	(0.054)	(0.063)			(0.069)	(0.017)	(0.022)	(0.026)	
C4	−0.136	0.991	0.184	0.378	67.806	M4	0.009	1.080	0.674	0.369	92.834
	(0.132)	(0.032)	(0.043)	(0.049)			(0.067)	(0.016)	(0.022)	(0.025)	
C5	0.160	1.034	0.045	−0.207	54.356	M5	0.018	1.060	0.551	0.370	93.207
	(0.202)	(0.049)	(0.065)	(0.076)			(0.062)	(0.015)	(0.020)	(0.023)	
C6	−0.113	0.994	0.102	0.116	50.650	M6	−0.051	1.060	0.382	0.291	93.942
	(0.194)	(0.047)	(0.063)	(0.073)			(0.056)	(0.013)	(0.018)	(0.021)	
C7	0.124	1.021	0.537	0.424	55.285	M7	−0.007	1.087	0.244	0.272	93.746
	(0.196)	(0.048)	(0.064)	(0.073)			(0.056)	(0.014)	(0.018)	(0.021)	
C8	0.189	0.976	0.313	0.434	56.532	M8	0.011	1.078	0.151	0.243	94.643
	(0.170)	(0.041)	(0.055)	(0.064)			(0.050)	(0.012)	(0.016)	(0.019)	
C9	0.365	0.985	0.262	−0.153	69.274	M9	−0.045	1.015	0.037	0.196	94.713
	(0.146)	(0.035)	(0.047)	(0.055)			(0.046)	(0.011)	(0.015)	(0.017)	
C10	0.124	1.002	0.280	0.387	66.499	M10	0.078	0.954	−0.248	−0.111	97.054
	(0.141)	(0.034)	(0.046)	(0.053)			(0.033)	(0.008)	(0.011)	(0.012)	
			GRS: 1.692 (0.079)						GRS: 1.538 (0.122)		

(*continued overleaf*)

Table 5
(Continued)

	α	β	γ	δ	R^2		α	β	γ	δ	R^2
B1	0.142	0.967	−0.157	−0.445	91.140	BT1	−0.007	0.680	0.030	0.373	58.620
	(0.068)	(0.016)	(0.022)	(0.025)			(0.105)	(0.026)	(0.034)	(0.039)	
B2	0.109	1.015	−0.063	−0.070	91.291	BT2	0.106	0.757	−0.126	0.212	68.142
	(0.063)	(0.015)	(0.020)	(0.024)			(0.095)	(0.023)	(0.031)	(0.036)	
B3	−0.054	1.023	−0.010	0.069	90.585	BT3	0.089	0.844	−0.119	0.125	77.803
	(0.065)	(0.016)	(0.021)	(0.024)			(0.085)	(0.021)	(0.027)	(0.032)	
B4	−0.063	1.027	0.001	0.249	86.576	BT4	−0.004	0.945	−0.038	0.184	83.458
	(0.076)	(0.019)	(0.025)	(0.029)			(0.080)	(0.019)	(0.026)	(0.030)	
B5	0.068	0.974	−0.022	0.277	83.438	BT5	0.078	0.996	−0.034	0.127	86.154
	(0.081)	(0.020)	(0.026)	(0.030)			(0.077)	(0.019)	(0.025)	(0.029)	
B6	0.004	0.968	0.022	0.339	87.329	BT6	0.005	1.079	0.018	0.101	88.562
	(0.069)	(0.017)	(0.022)	(0.026)			(0.076)	(0.018)	(0.025)	(0.028)	
B7	−0.015	0.969	0.104	0.530	87.740	BT7	−0.074	1.135	0.031	0.074	90.391
	(0.067)	(0.016)	(0.022)	(0.025)			(0.073)	(0.018)	(0.024)	(0.027)	
B8	−0.060	0.990	0.148	0.633	89.798	BT8	−0.194	1.203	0.220	0.042	89.537
	(0.063)	(0.015)	(0.020)	(0.023)			(0.085)	(0.021)	(0.027)	(0.032)	
B9	0.022	1.014	0.252	0.703	88.875	BT9	−0.226	1.336	0.314	0.036	87.006
	(0.069)	(0.017)	(0.022)	(0.026)			(0.108)	(0.026)	(0.035)	(0.041)	
B10	0.014	1.089	0.374	0.723	88.141	BT10	−0.204	1.523	0.508	−0.137	84.587
	(0.079)	(0.019)	(0.025)	(0.029)			(0.145)	(0.035)	(0.047)	(0.054)	
			GRS: 1.345 (0.203)						GRS: 1.270 (0.244)		

This table presents specification tests for four sets of 10 portfolios. We conduct specification tests based on the regression specification:

$$r_{it} - r_f = \alpha_i + \beta_i (r_{mt} - r_f) + \gamma_i r_{SMB,t} + \delta_i r_{HML,t} + \epsilon_{it},$$

where r_{it} is the return on portfolio i; r_f is the risk-free rate, measured as the return on the T-bill closest to one month to maturity; r_{mt} is the return on the value-weighted CRSP index; $r_{SMB,t}$ is the return on a small size portfolio in excess of the return on a large size portfolio; and $r_{HML,t}$ is the return on a high book-to-market portfolio in excess of the return on a low book-to-market portfolio. The parameters are estimated via least squares, and the restriction $\alpha = 0$ is tested using the GRS specification test. The table presents the coefficient estimates and accompanying standard errors, along with regression R^2. The test statistics and p-values for the specification tests are presented below the coefficient estimates. Tests are conducted on four separate sets of assets: cluster, size, book-to-market, and beta-sorted portfolios. Data are sampled at the monthly frequency from July 1959 through December 2003.

Table 6
CAPM specification tests: 25 assets

Cluster portfolios

Port.	α	β	R^2	Port.	α	β	R^2	Port.	α	β	R^2	Port.	α	β	R^2	Port.	α	β	R^2
C1	0.09	0.99	56.11	C6	0.59	0.78	25.25	C11	0.27	0.96	52.08	C16	0.26	1.07	44.78	C21	0.67	1.01	27.52
	(0.17)	(0.04)			(0.26)	(0.06)			(0.18)	(0.04)			(0.23)	(0.05)			(0.32)	(0.07)	
C2	0.33	1.02	44.56	C7	0.38	0.94	34.15	C12	0.29	0.92	39.91	C17	0.28	0.94	43.45	C22	0.27	0.84	34.31
	(0.22)	(0.05)			(0.25)	(0.06)			(0.22)	(0.05)			(0.21)	(0.05)			(0.22)	(0.05)	
C3	0.27	0.98	39.68	C8	0.29	0.97	51.03	C13	−0.09	0.95	44.23	C18	0.41	1.19	42.48	C23	0.55	0.95	30.15
	(0.23)	(0.05)			(0.18)	(0.04)			(0.21)	(0.05)			(0.27)	(0.06)			(0.28)	(0.06)	
C4	0.20	0.91	46.75	C9	0.03	0.87	38.06	C14	0.37	0.92	46.88	C19	0.53	0.96	42.96	C24	0.02	0.90	32.89
	(0.19)	(0.04)			(0.22)	(0.05)			(0.19)	(0.04)			(0.22)	(0.05)			(0.25)	(0.06)	
C5	0.29	1.04	53.73	C10	0.37	1.12	52.38	C15	0.27	0.65	34.70	C20	0.26	1.11	45.09	C25	0.37	0.88	34.76
	(0.19)	(0.04)			(0.21)	(0.05)			(0.17)	(0.04)			(0.24)	(0.05)			(0.24)	(0.05)	

GRS: 1.930 (0.047)

Size and book-to-market portfolios

Port.	α	β	R^2	Port.	α	β	R^2	Port.	α	β	R^2	Port.	α	β	R^2	Port.	α	β	R^2
S1B1	−0.23	1.40	58.77	S2B1	−0.09	1.35	73.99	S3B1	−0.05	1.28	81.47	S4B1	−0.04	1.22	88.19	S5B1	−0.04	1.00	87.05
	(0.23)	(0.05)			(0.15)	(0.03)			(0.12)	(0.03)			(0.09)	(0.02)			(0.07)	(0.02)	
S1B2	0.22	1.22	62.11	S2B2	0.14	1.15	73.17	S3B2	0.11	1.08	82.21	S4B2	0.01	1.06	85.73	S5B2	0.00	0.95	86.92
	(0.18)	(0.04)			(0.14)	(0.03)			(0.10)	(0.02)			(0.08)	(0.02)			(0.07)	(0.02)	
S1B3	0.38	1.08	62.66	S2B3	0.34	1.02	73.08	S3B3	0.24	0.95	77.12	S4B3	0.27	0.97	81.03	S5B3	0.18	0.85	76.34
	(0.16)	(0.04)			(0.12)	(0.03)			(0.10)	(0.02)			(0.09)	(0.02)			(0.09)	(0.02)	
S1B4	0.61	1.01	62.10	S2B4	0.52	0.96	69.82	S3B4	0.33	0.91	72.61	S4B4	0.41	0.91	76.94	S5B4	0.20	0.80	70.18
	(0.15)	(0.03)			(0.12)	(0.03)			(0.11)	(0.02)			(0.10)	(0.02)			(0.10)	(0.02)	
S1B5	0.62	1.05	59.46	S2B5	0.47	1.04	69.40	S3B5	0.47	1.00	71.54	S4B5	0.45	0.98	71.32	S5B5	0.32	0.82	61.92
	(0.17)	(0.04)			(0.13)	(0.03)			(0.12)	(0.03)			(0.12)	(0.03)			(0.12)	(0.03)	

GRS: 3.353 (0.000)

This table presents specification tests for two sets of 25 portfolios. We conduct specification tests based on the regression specification:

$$r_{it} - r_f = \alpha_i + \beta_i (r_{mt} - r_f) + \epsilon_{it},$$

where r_{it} is the return on portfolio i; r_f is the risk-free rate, measured as the return on the T-bill closest to one month to maturity; and r_{mt} is the return on the value-weighted CRSP index. The parameters α_i and β_i are estimated via least squares, and the restriction $\alpha = 0$ is tested using the GRS specification test. The table presents the least squares coefficient estimates of α and β and accompanying standard errors, along with regression R^2. The test statistics and p-values for the specification tests are presented below the coefficient estimates. Tests are conducted on four separate sets of assets: cluster, size, book-to-market, and beta-sorted portfolios. Data are sampled at the monthly frequency from July 1959 through December 2003.

Table 7
Three-factor specification tests: 25 assets

Panel A: Cluster portfolios

	α	β	γ	δ	R^2
C1	−0.133	1.004	0.379	0.349	61.296
	(0.164)	(0.040)	(0.053)	(0.061)	
C2	0.351	0.990	0.068	−0.047	44.714
	(0.226)	(0.055)	(0.073)	(0.084)	
C3	0.014	1.017	0.374	0.418	43.937
	(0.233)	(0.056)	(0.075)	(0.087)	
C4	0.020	0.998	−0.050	0.314	48.925
	(0.189)	(0.046)	(0.061)	(0.071)	
C5	0.181	1.057	0.153	0.175	54.610
	(0.191)	(0.046)	(0.062)	(0.071)	

	α	β	γ	δ	R^2
C6	0.405	0.793	0.333	0.294	28.037
	(0.263)	(0.064)	(0.085)	(0.098)	
C7	0.263	0.890	0.477	0.168	38.191
	(0.252)	(0.061)	(0.082)	(0.094)	
C8	0.228	0.922	0.324	0.081	53.630
	(0.183)	(0.045)	(0.059)	(0.069)	
C9	−0.190	0.906	0.310	0.355	41.676
	(0.216)	(0.052)	(0.070)	(0.081)	
C10	0.425	1.077	0.083	-0.094	52.698
	(0.212)	(0.051)	(0.069)	(0.079)	

	α	β	γ	δ	R^2
C11	0.108	0.990	0.213	0.267	54.206
	(0.180)	(0.044)	(0.058)	(0.067)	
C12	0.220	0.912	0.186	0.111	40.729
	(0.223)	(0.054)	(0.072)	(0.084)	
C13	−0.099	0.863	0.408	−0.019	48.124
	(0.205)	(0.050)	(0.066)	(0.077)	
C14	0.163	0.955	0.276	0.340	50.318
	(0.189)	(0.046)	(0.061)	(0.071)	
C15	0.019	0.742	0.089	0.421	39.681
	(0.171)	(0.041)	(0.055)	(0.064)	

	α	β	γ	δ	R^2
C16	0.130	1.013	0.516	0.175	49.550
	(0.226)	(0.055)	(0.073)	(0.085)	
C17	0.147	0.923	0.366	0.205	46.689
	(0.208)	(0.050)	(0.067)	(0.078)	
C18	0.463	1.061	0.487	−0.127	46.268
	(0.267)	(0.065)	(0.086)	(0.100)	
C19	0.359	0.963	0.346	0.259	46.018
	(0.215)	(0.052)	(0.070)	(0.081)	
C20	0.391	0.967	0.397	−0.261	49.216
	(0.236)	(0.057)	(0.076)	(0.088)	

	α	β	γ	δ	R^2
C21	0.496	1.072	0.065	0.295	28.327
	(0.324)	(0.079)	(0.105)	(0.121)	
C22	0.178	0.815	0.285	0.135	36.201
	(0.227)	(0.055)	(0.074)	(0.085)	
C23	0.545	0.843	0.480	−0.029	33.875
	(0.279)	(0.068)	(0.090)	(0.105)	
C24	−0.084	0.884	0.315	0.161	34.819
	(0.254)	(0.062)	(0.082)	(0.095)	
C25	0.194	0.884	0.364	0.272	38.009
	(0.236)	(0.057)	(0.076)	(0.088)	

GRS: 1.550 (0.045)

Panel B: Size and book-to-market portfolios

	α	β	γ	δ	R^2
S1B1	−0.301	1.188	1.102	0.036	75.768
	(0.179)	(0.043)	(0.058)	(0.067)	
S1B2	0.000	1.094	1.006	0.295	81.434
	(0.133)	(0.032)	(0.043)	(0.050)	
S1B3	0.060	1.014	0.949	0.477	86.133
	(0.101)	(0.024)	(0.033)	(0.038)	
S1B4	0.226	0.977	0.916	0.584	88.682
	(0.086)	(0.021)	(0.028)	(0.032)	
S1B5	0.112	1.063	0.978	0.791	89.304
	(0.089)	(0.022)	(0.029)	(0.033)	

	α	β	γ	δ	R^2
S2B1	−0.065	1.177	0.722	-0.094	84.532
	(0.123)	(0.030)	(0.040)	(0.046)	
S2B2	−0.045	1.074	0.757	0.269	87.666
	(0.094)	(0.023)	(0.031)	(0.035)	
S2B3	0.028	1.013	0.675	0.491	90.601
	(0.073)	(0.018)	(0.024)	(0.027)	
S2B4	0.112	0.998	0.675	0.641	91.871
	(0.066)	(0.016)	(0.021)	(0.025)	
S2B5	−0.022	1.099	0.736	0.779	93.839
	(0.062)	(0.015)	(0.020)	(0.023)	

	α	β	γ	δ	R^2
S3B1	0.049	1.134	0.470	−0.207	88.081
	(0.098)	(0.024)	(0.032)	(0.037)	
S3B2	−0.083	1.071	0.436	0.304	89.407
	(0.077)	(0.019)	(0.025)	(0.029)	
S3B3	−0.083	1.006	0.402	0.523	89.121
	(0.071)	(0.017)	(0.023)	(0.027)	
S3B4	−0.065	1.003	0.415	0.646	88.810
	(0.072)	(0.017)	(0.023)	(0.027)	
S3B5	0.004	1.095	0.525	0.754	90.728
	(0.072)	(0.017)	(0.023)	(0.027)	

	α	β	γ	δ	R^2
S4B1	0.087	1.133	0.138	−0.222	90.021
	(0.082)	(0.020)	(0.026)	(0.031)	
S4B2	−0.139	1.084	0.207	0.247	88.303
	(0.078)	(0.019)	(0.025)	(0.029)	
S4B3	−0.003	1.050	0.188	0.461	87.793
	(0.075)	(0.018)	(0.024)	(0.028)	
S4B4	0.058	1.021	0.194	0.580	88.056
	(0.071)	(0.017)	(0.023)	(0.027)	
S4B5	0.024	1.101	0.318	0.710	85.542
	(0.088)	(0.021)	(0.028)	(0.033)	

	α	β	γ	δ	R^2
S5B1	0.172	0.953	−0.220	−0.341	91.615
	(0.062)	(0.015)	(0.020)	(0.023)	
S5B2	−0.056	1.005	−0.159	0.103	88.671
	(0.068)	(0.017)	(0.022)	(0.025)	
S5B3	0.066	0.943	−0.186	0.212	80.353
	(0.086)	(0.021)	(0.028)	(0.032)	
S5B4	−0.111	0.959	−0.091	0.536	82.467
	(0.080)	(0.019)	(0.026)	(0.030)	
S5B5	−0.051	0.985	−0.004	0.639	75.312
	(0.103)	(0.025)	(0.033)	(0.038)	

GRS: 2.286 (0.000)

This table presents specification tests for two sets of 25 portfolios. We conduct specification tests based on the regression specification:

$$r_{it} - r_f = \alpha_i + \beta_i(r_{mt} - r_f) + \gamma_i r_{SMB,t} + \delta_i r_{HML,t} + \epsilon_{it},$$

where r_{it} is the return on portfolio i; r_f is the risk-free rate, measured as the return on the T-bill closest to one month to maturity; r_{mt} is the return on the value-weighted CRSP index; $r_{SMB,t}$ is the return on a small size portfolio in excess of the return on a large size portfolio; and $r_{HML,t}$ is the return on a high book-to-market portfolio in excess of the return on a low book-to-market portfolio. The parameters are estimated via least squares, and the restriction $\alpha = 0$ is tested using the GRS specification test. The table presents the coefficient estimates and accompanying standard errors, along with regression R^2. The test statistics and p-values for the specification tests are presented below the coefficient estimates. The test assets are sorted on clusters and size and book-to-market. Data are sampled at the monthly frequency from July 1959 through December 2003.

portfolios. These small intercepts are consistent with the results of Fama and French (1996), and lead to a small GRS test statistic and a failure to reject the three-factor model. In contrast, although the GRS test fails to reject for the cluster portfolios, the estimates of the intercepts are much larger, ranging from −14 to 37 basis points per month. The reason for the failure to reject in this case is not small pricing errors, but the relatively large covariance matrix associated with the intercepts. That is, the R^2 presented in Tables 4 and 5 are substantially lower for the cluster portfolios than for any of the characteristic-sorted portfolios. These low R^2 measures translate into larger residual variances, larger intercept standard errors, and thus a lower GRS test statistic. Using the cluster portfolios, we see the mirror image of the interpretation in Fama and French (1996); while they emphasize that the three-factor model, with its superior fit, can make even small intercepts distinguishable from zero, when we use the cluster portfolios to test the model, the low explanatory power of the model makes even large intercepts indistinguishable from zero.[9]

We present results of tests of the CAPM and the three-factor model using 25 test portfolios in Tables 6 and 7. As shown in the tables, when 25 cluster or size and book-to-market portfolios are used to test model specification, both the CAPM and the three-factor model are strongly rejected. As in tests using ten portfolios, although the inferences from the GRS test statistic are similar across the sets of basis assets, the results for individual portfolios are quite different. The regression R^2 values again suggest a systematically poorer fit of the time series of the cluster portfolios than the characteristic portfolios. Additionally, the regression intercepts of the cluster portfolios are again substantially larger in absolute value than those of the characteristic portfolios.

The results from these specification tests indicate that while multivariate tests of the CAPM and the three-factor model lead to similar inferences, individual cluster portfolios have substantially larger pricing errors, and the CAPM and Fama–French factors fit the cluster portfolios returns substantially less well than the traditional characteristic-sorted portfolios. Since the three-factor model was generated, at least in part, by fitting the returns process in the characteristic-sorted portfolios, its superior fit on these portfolios is not

[9] We explored the possibility that the low R^2 for the cluster portfolios is driven by instability in the cluster portfolios' composition. Specifically, we examined the time-series standard deviation of factor loadings for cluster and characteristic-sorted portfolios, as well as autocorrelations in these loadings. While the standard deviation of the factor loadings in the overall sample period for cluster portfolios is higher for the market, SMB and HML factors (by 50%, 60%, and 18%, respectively), the autocorrelations of the loadings across the two sets of basis assets are quite similar (and above 0.9 in all cases). We also break the sample period into four equal-length subperiods and compare the R^2's for the cluster and characteristic-sorted portfolios, the CAPM, and the three-factor model. If instability in the cluster portfolios' composition, and hence a lack of comparability through time in the factor loadings, is driving the low R^2 in the overall sample period results, the difference in R^2 should diminish when subsamples are analyzed, since the cluster portfolios would exhibit greater stability over shorter sample periods. In fact, with one exception (the first subperiod of 1959–1970, for the CAPM model, where the R^2 for the clusters is 0.153 and for the characteristic-sorted portfolios is 0.135), the difference in explanatory power across the cluster and characteristic-sorted portfolios is even larger than in the overall sample. Consequently, it does not appear that instability in the cluster portfolios, and consequently in their loadings, is the source of the difference in R^2's in Tables 4–7.

surprising.[10] What may be surprising is the model's relatively poor fit for the cluster portfolios. Substantial portions of cluster portfolio returns are unexplained by the three-factor model. As a consequence, and related to the arguments of Daniel and Titman (2005), the cluster portfolios may have the ability to provide more powerful tests of some asset pricing models.

4.2 Cross-sectional regression tests

In addition to the multivariate tests of return-based factor models above, we perform cross-sectional tests of the CAPM, the consumption CAPM, the three-factor model, and the conditional CAPM of Jagannathan and Wang (1996) (which includes factors related to labor income and time-variation in risk premia, captured in the default spread). To gain a better understanding of where the pricing models may fail using different sets of basis assets, we include other right-hand-side variables, such as firm size and book-to-market equity. Because our (real) consumption data are measured quarterly, all tests are run on real quarterly returns. Results are reported in Table 8.

There are striking differences in inferences across the two sets of portfolios. The cluster portfolios' returns are positively and significantly related to both CAPM betas and consumption betas (regressions 1 and 3, respectively). The relation between returns and (both) betas is unaffected by the inclusion of firm size and book-to-market equity characteristics, which are not significantly related to returns (regressions 2 and 4). In contrast, when the Fama–French portfolio returns are the dependent variable (in Panel B), the coefficient on CAPM beta is negative and insignificant (regression 1); when size and book-to-market variables are included (in regression 2), the coefficient on beta remains negative and becomes significant. The coefficient on the consumption CAPM beta is positive and significant for the Fama–French portfolios (regression 3), but does not survive the inclusion of size and book-to-market characteristics (regression 4).

Not surprisingly, the size- and book-to-market-sorted portfolio returns are much more strongly related to SMB and HML, as well as size and book-to-market characteristics, than the cluster portfolios. As a result, the $\bar{R}^2$ of cross-sectional regressions using the Fama–French portfolios is higher when these variables are included in the specification. In fact, the cluster portfolio returns have no significant relation to size and book-to-market characteristics or the exposure to the SMB and HML factors, in any of the specifications. One

[10] For example, the size sorting in the 25 Fama–French portfolios can easily be seen in the magnitude of the coefficients on SMB in Table 8, Panel B. The coefficient, γ, declines steadily across the S1–S5 sort dimension and is negative for the S5 (largest) firms; the range in γ is large, from a maximum of 1.10 to -0.22. In contrast, while there is significant variability in the γ coefficient across the cluster portfolios in Panel A, the dispersion is much smaller—only one coefficient is (insignificantly) negative, and the coefficient ranges from -0.05 to 0.52. In Section 5, we examine whether there is statistically significant dispersion in such characteristics as market value across the cluster portfolios; however, it is worth pointing out that the average γ across the cluster portfolios is consistent with the γ of the mid-cap firms (or size categories 3 and 4) in the Fama–French sorting. Thus, the cluster portfolios do not appear to be dominated by small (or large) capitalization firms.

Table 8
Cross-sectional regressions

Panel A: Cluster portfolios

Specification		γ_0	γ_m	$\gamma_{\Delta c}$	$\gamma_{\Delta l}$	γ_{DS}	γ_{SMB}	γ_{HML}	$\ln MV$	BM	$\bar{R}^2$
(1)	Coeff.	1.225	1.539								0.112
	t-stat	1.574	2.008								
(2)	Coeff.	−0.685	1.684						0.125	−0.227	0.096
	t-stat	−0.128	2.152						0.381	−0.998	
(3)	Coeff.	2.321		0.144							0.146
	t-stat	10.295		2.260							
(4)	Coeff.	1.160		0.167					0.087	−0.295	0.159
	t-stat	0.226		2.558					0.274	−1.319	
(5)	Coeff.	1.567	1.883		0.039	−21.777					0.422
	t-stat	2.465	2.969		0.339	−3.333					
(6)	Coeff.	−3.140	1.927		0.080	−20.709			0.293	−0.071	0.415
	t-stat	−0.716	3.015		0.664	−2.968			1.093	−0.366	
(7)	Coeff.	1.781	1.038				−0.443	−0.678			0.072
	t-stat	1.569	0.591				−0.521	−0.826			
(8)	Coeff.	1.413	1.250				−0.226	−0.449	0.018	−0.203	0.012
	t-stat	0.189	0.622				−0.246	−0.415	0.042	−0.823	

Panel B: Size and book-to-market portfolios

Specification		γ_0	γ_m	$\gamma_{\Delta c}$	$\gamma_{\Delta l}$	γ_{DS}	γ_{SMB}	γ_{HML}	ln MV	BM	$\bar{R}^2$
(1)	Coeff.	2.793	−0.058								−0.043
	t-stat	3.840	−0.088								
(2)	Coeff.	9.300	−2.150						−0.332	0.138	0.837
	t-stat	9.488	−5.209						−7.833	3.245	
(3)	Coeff.	1.820		0.264							0.163
	t-stat	4.554		2.382							
(4)	Coeff.	5.955		−0.146					−0.228	0.247	0.650
	t-stat	4.676		−0.187					−3.377	4.466	
(5)	Coeff.	2.402	−1.456		0.397	43.530					0.200
	t-stat	3.653	−1.877		0.833	2.993					
(6)	Coeff.	9.401	−2.160		0.013	−0.953			−0.336	0.137	0.820
	t-stat	6.963	−4.735		0.057	−0.336			−5.807	3.019	
(7)	Coeff.	3.519	−1.822				1.367	0.923			0.762
	t-stat	2.389	−0.818				2.223	1.523			
(8)	Coeff.	10.706	−4.075				0.454	−0.602	−0.343	0.163	0.828
	t-stat	3.651	−1.989				0.623	−0.853	−2.435	2.808	

A7 This table presents results of cross-sectional regressions. Coefficients, t-statistics, and adjusted R^2 are presented for a number of specifications: (1) market beta; (2) market beta plus average log size and book-to-market; (3) consumption beta; (4) consumption beta plus average log size and book-to-market; (5) conditional CAPM, with market beta, labor beta, and default spread beta, calculated as the slope coefficient of a regression of returns on the difference in Baa-rated and Aaa-rated bond yields; (6) conditional CAPM plus average log size and book-to-market; (7) the Fama–French (1993) three-factor model; and (8) the Fama–French model plus average log size and book-to-market. In Panel A, we present results for 25 cluster-sorted portfolios. In Panel B, we present analogous results for 25 size- and book-to-market-sorted portfolios. Data are sampled at the quarterly frequency from 1959:2 through 2003:4 and are converted to real quantities using the PCE deflator from the NIPA tables.

concern is that we fail to find a significant role for the size and book-to-market characteristics in pricing the cluster portfolios due to insufficient dispersion in these characteristics. We examine this issue in the next section.

Perhaps the most striking performance difference in the two sets of basis assets is the conditional CAPM of Jagannathan and Wang (1996); see regression 5 in both panels. The adjusted R^2 of the model for the cluster portfolios exceeds 40%, the coefficient on market beta is positive, and the coefficient on default spread is negative. The prices of market beta and default spread beta are both statistically different than zero. Thus, the results indicate that exposure to the market portfolio return and the default spread have reasonable explanatory power for the cross-section of cluster-sorted portfolio returns. Furthermore, these conclusions remain after controlling for size and book-to-market characteristics, which do not have significant explanatory power for the cluster portfolio returns. In contrast, when tested on the characteristic-sorted portfolios, the conditional CAPM has roughly half of this explanatory power and generates a negative and statistically insignificant market beta coefficient, as well as a significantly positive coefficient on default spread (which does not survive the inclusion of portfolio size and book-to-market variables).

In general, we find large differences in the results of the cross-sectional regression tests across the two sets of assets. Traditional measures of risk, such as the CAPM and consumption CAPM betas, are positively and significantly related to the returns of cluster portfolios, even when firm characteristics such as size and book-to-market are included in the regression. In addition, the magnitude of the market risk premia implied by the estimate of γ_m for cluster portfolios is economically plausible, implying an annualized market risk premium of 4.15%–7.71% per year. There are also significant differences in the relation of other variables, such as the default spread beta, to the portfolio returns.

Because the Fama–French portfolio returns are designed to have maximum dispersion along size and book-to-market dimensions, the $\bar{R}^2$ of the cross-sectional regressions is much higher for these portfolios than that for the cluster portfolios when the characteristics or characteristic-based risk measures are included in the regression. However, the inclusion of size and book-to-market factors results in prices of CAPM and consumption CAPM beta risk that are either insignificant, or significant and negative.

The differences in the results of our cross-sectional tests across the two sets of portfolios are consistent with the arguments of Daniel and Titman (2005). The authors suggest that sorting on size and book-to-market maximizes dispersion in these characteristics at the expense of masking dispersion in exposures to other fundamental factors. The dispersion in size, book-to-market, and their associated factor exposures is obviously much higher for the characteristic-sorted portfolios than for the cluster portfolios. However, the dispersion of market beta is only slightly higher for the characteristic portfolios than for the cluster portfolios. Despite this fact, the price of cluster portfolio beta risk is positive and significant, in contrast to the price of characteristic portfolio beta risk. Since our clusters are formed on the basis of comovement among securities, these portfolios may be more related to fundamental risk factors in the economy. Consequently, rejecting a particular asset pricing model because of its inability to explain the returns of (some) cluster portfolios may provide additional information about the misspecification, or the specific identity of omitted risk factors, in the model.

5. Characterizing Clusters

In the cross-sectional tests above, we find evidence that the cluster portfolios are positively and significantly related to both betas and consumption betas; these results are robust to the inclusion of other characteristics, such as firm size and book-to-market equity. One possible explanation for these results is that the cluster portfolios do not have sufficient dispersion along the dimension of the firm characteristics which we include in the tests. In contrast, of course, the 25 size- and book-to-market-sorted portfolios are designed to have maximum dispersion along these dimensions. In this section, we explore the relation between the cluster portfolios we form and a number of characteristics. These results are useful not only because they address the question of statistical power in the cross-sectional tests, but because they help to characterize the nature of the cluster portfolios.

The clustering algorithm is silent on the nature of the factors that influence correlation between securities.[11] However, there are well-known techniques that can be used to characterize, or profile, the groups that result from a cluster analysis. For example, one could profile the clusters using discriminant analysis, with the cluster variable as the dependent variable in an analysis that asks what factors influence cluster or group membership.

Although we do not perform an exhaustive search of the possible factors, we examine whether or to what extent well-known and well-used factors or characteristics relate to the clusters formed using this grouping algorithm. Since our cluster portfolios have positive (value) weights on the individual securities in each portfolio, we are easily able to measure portfolio characteristics such as size and book-to-market. Specifically, in Table 9, we characterize the cluster portfolios' average (log) market capitalization, book-to-market equity, and market beta; we also test whether there is significant dispersion among the cluster portfolios along these dimensions. For each characteristic, we calculate the value at the end of the 60-month period of returns used to generate the clusters; this value is thus taken immediately before the 12 months over which the cluster identification is held constant. Similarly, the beta is calculated over

[11] Despite the fact that the clustering algorithm does not identify factors, we note that any econometric method that seeks to estimate factors from the covariance matrix of returns, such as factor analysis, would benefit from the superior conditioning of the cluster portfolios' covariance matrix.

Table 9
Cluster portfolio characteristics

Portfolio	lnMV	BM	Beta	MV Dec.
C1	14.533	0.873	1.117	6.2
C2	14.784	0.971	0.806	7.1
C3	15.427	0.773	0.900	7.1
C4	15.443	1.000	1.014	7.3
C5	15.588	0.713	1.092	7.4
C6	15.335	1.024	1.030	7.3
C7	14.680	1.015	1.002	6.3
C8	15.297	0.686	0.995	6.9
C9	15.185	0.785	1.086	6.9
C10	15.252	0.784	1.032	7.2
Std. dev.	0.358	0.131	0.094	
t(Max–Min)	12.645	8.682	18.286	

This table depicts characteristics of the 10 cluster assets examined in the paper. We present the average log market value, book-to-market ratio, beta, and NYSE market capitalization decile of each of the 10 portfolios formed on clusters. At the bottom of each of the first three columns, we present the standard deviation in the average size, book-to-market, and beta across portfolios, and the t-statistic for the difference in the average maximum and minimum. Data are sampled at the monthly frequency from July 1959 through December 2003.

the 60-month window preceding the cluster identification in the table. These portfolio characteristics are then averaged over all subperiods.

Clearly, there is significant dispersion along these well-known firm characteristics between the cluster portfolios. For example, the average betas in the cluster portfolios vary from 0.81 to 1.12, and this range is highly significant. The differences among portfolios' market capitalization and book-to-market are significant as well. Overall, these results suggest that clusters are related to firm characteristics and that the evidence in the cross-sectional tests that betas are priced, while characteristics are not, does not seem to be driven by a lack of significant dispersion in the characteristics.[12]

Clearly, there are other methods of forming portfolios, such as principal components analysis, which also seek to capture the information in a large matrix of returns by forming a small number of (well-chosen) portfolios. In Appendix C, we examine the relation between cluster portfolios and the principal components extracted from security returns. We present evidence that, despite the binding constraints applied to the clustering method, the cluster portfolios have similar advantages to principal components and these portfolios are easier to form and characterize. Consequently, they may provide more useful information to identify economically important sources of comovement than principal components analysis.

[12] Note that although the clustering algorithm sorts on covariances between individual securities, this does not imply that we are sorting on betas. The covariances between the returns of individual securities may reflect many other factors besides their correlation with a common market return. For comparison, we also examined the relation between the returns of principal component-mimicking portfolios and characteristics. The correlations between "average" characteristics and subsequent returns were not as strong as the relation for the cluster portfolios. For example, the correlation between market value and returns was 0.04, while the relation between book-to-market and returns was both smaller and negative (−0.38). The relation between beta and returns was also negative, at −0.32.

Table 10
Industry portfolio correlations

	NoDur	Durbl	Manuf	Enrgy	HiTec	Telcm	Shops	Hlth	Utils	Other
NoDur	1.00									
Durbl	0.66	1.00								
Manuf	0.83	0.80	1.00							
Enrgy	0.51	0.46	0.61	1.00						
HiTec	0.57	0.63	0.74	0.41	1.00					
Telcm	0.63	0.59	0.64	0.41	0.61	1.00				
Shops	0.84	0.77	0.84	0.46	0.70	0.67	1.00			
Hlth	0.78	0.51	0.74	0.45	0.64	0.57	0.69	1.00		
Utils	0.63	0.47	0.54	0.57	0.29	0.53	0.50	0.49	1.00	
Other	0.83	0.77	0.89	0.60	0.71	0.69	0.85	0.74	0.61	1.00

This table presents correlations of monthly returns on 10 industry-sorted portfolios. Industry definitions and data are taken from Kenneth French's Web site. Data cover the period July 1959–December 2003.

5.1 Clusters and industries

As mentioned above, King (1966) presents evidence that firms in similar industries have higher cross-correlations. We compare the ability of industry and cluster portfolios to group on the basis of comovement. The cross-correlations of the industry portfolios are presented in Table 10. The average cross-correlations of the ten industry portfolios, at 0.63, is higher than the cross-correlations of the cluster portfolios, and the dispersion in subsequent returns, at 27 basis points per month, is significantly lower as well.

The lower cross-correlations in the cluster portfolios suggest that these portfolios may be more similar to "pure plays" on fundamental factors than are industry portfolios. To explore this further, we run the following sets of regressions. First, we examine the increase in explanatory power in a regression of cluster portfolio returns as we sequentially add additional industry portfolios; we then reverse the experiment and regress industry portfolio returns on cluster portfolios. If cluster portfolios are more similar to a single fundamental factor, then the increase in explanatory power in industry returns as we add cluster portfolios should be higher than that observed when we add additional industry portfolios to explain cluster portfolio returns.[13]

Specifically, we begin with individual cluster portfolios as the dependent variable, and the single industry portfolio on the right-hand side with which the cluster is most closely associated. Taking the residual from this regression, we examine the increase in explanatory power as we add industry portfolios as independent variables one at a time. We then repeat this exercise, exactly reversing the role of cluster and industry portfolios. The results are presented in Table 11 and Figure 5.

[13] With N underlying factors, N cluster portfolios, and N industry portfolios, utilizing all N portfolios will result in regressions for industry portfolios on cluster portfolios, and vice versa, with the same R^2. If, however, we examine a subset of these factors, this will not necessarily be the case. We thank an anonymous referee for pointing out this issue.

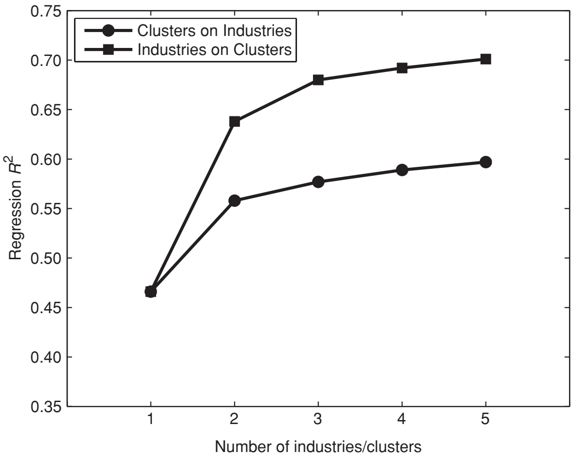


Figure 5
Contribution to explanatory power: Industries vs. clusters

In Table 11, we present the increase in explanatory power as we add the first five portfolios to each regression. In Figure 5, we present the average (across ten cluster [industry] portfolios) of the R^2 obtained as we add all ten industry (cluster) portfolios to the regression. The increase in explanatory power as we add cluster portfolios is larger; the slope of the top line in Figure 5 is steeper than the bottom line, which represents adding industry portfolios to the regression. Moreover, the total explanatory power with ten cluster portfolios is higher than that with ten industry portfolios.[14] Overall, these results suggest that cluster portfolios are related to industry portfolios, but industry portfolios are more cross-correlated, have lower incremental explanatory power, and explain a lower cumulative fraction of returns than cluster portfolios.

Finally, we report evidence on the stability of the clusters, which one would expect to find if the determinants of cluster membership are related to economic characteristics of the firms, which changed only slowly over time. Specifically, we estimate the propensity of firms in a particular cluster i at time t to have been in the same cluster in the previous period. For the set of ten cluster portfolios, the average propensity of firms to remain in similar clusters for overlapping samples (nonoverlapping five-year) samples is 20.3% (9.8%) per year. The stability of clusters is lower than that of the portfolios formed on the basis of firm characteristics, but there are also large differences in stability across the characteristics. For example, while market capitalization is quite stable, with persistence across one-year periods (five-year periods) of 64.9% (36.5%), the stability of book-to-market portfolios is much lower,

[14] In fact, the results with cluster portfolios are quite similar to the cumulative percentage of returns explained with principal components; see Appendix C for details.

Table 11
Contribution to explanatory power: Industries vs. clusters

Panel A: Clusters on industries

Cluster	Industry	R_1^2	R_2^2	R_3^2	R_4^2	R_5^2
1	Telecom	0.249	0.401	0.438	0.459	0.473
2	Utilities	0.428	0.532	0.549	0.579	0.585
3	Healthcare	0.310	0.425	0.437	0.449	0.458
4	Manufacturing	0.616	0.671	0.671	0.671	0.671
5	High tech.	0.603	0.638	0.654	0.662	0.667
6	Energy	0.381	0.571	0.579	0.585	0.592
7	Durables	0.389	0.457	0.462	0.468	0.476
8	Nondurables	0.512	0.568	0.603	0.612	0.627
9	Retail/wholesale	0.551	0.660	0.711	0.728	0.739
10	Other	0.619	0.657	0.669	0.677	0.685

Panel B: Industries on clusters

Cluster	Industry	R_1^2	R_2^2	R_3^2	R_4^2	R_5^2
1	Telecom	0.249	0.455	0.533	0.539	0.543
2	Utilities	0.428	0.583	0.615	0.627	0.636
3	Healthcare	0.310	0.503	0.543	0.553	0.570
4	Manufacturing	0.616	0.785	0.829	0.855	0.860
5	High Tech.	0.603	0.767	0.791	0.798	0.801
6	Energy	0.381	0.500	0.527	0.544	0.549
7	Durables	0.389	0.547	0.583	0.595	0.600
8	Nondurables	0.512	0.682	0.754	0.758	0.764
9	Retail/wholesale	0.551	0.746	0.777	0.792	0.812
10	Other	0.619	0.809	0.850	0.864	0.875

This table presents the explanatory power of regressions of industry (cluster) returns on returns on clusters (industries). Regressions are performed stepwise, starting with the industry and cluster pairings in the first two columns of the table, retrieving the residual, and incrementally adding clusters (industries) that most increase the regression R^2. We utilize 10 cluster portfolios and 10 industry portfolios in the regressions. Panel A presents results for regressions of cluster portfolio returns on industry portfolio returns, and Panel B presents results for regressions of industry portfolio returns on cluster portfolio returns. Data are sampled at the monthly frequency and cover the period July 1959–December 2003.

at 36.1% in one-year samples and only 15.9% in five-year samples, and is not dramatically higher over five years than the persistence in the ten cluster portfolios.[15]

Overall, these results suggest that the use of the clustering algorithm to generate a set of basis assets yields significant cross-sectional dispersion in firm characteristics as well as returns.

6. Conclusion

In this paper, we focus on an alternative construction of a set of basis assets that characterizes an investor's opportunity set. The formation of basis assets underlies much of the empirical literature on asset pricing, including asset pricing model tests, inferences regarding profitable trading opportunities, and the investigation of omitted pricing model factors. We propose a new method of forming basis assets, which is not subject to potential data-snooping biases. This method explicitly seeks to minimize (maximize) inter-(intra-)group correlations. This approach contrasts sharply with the goal of traditional sorting exercises, which seek to maximize variation in *ex post* sample mean returns across portfolios.

The resulting set of basis assets seems to perform well along several dimensions. First, the method generates significant dispersion in future returns; indeed, despite the fact that the clustering algorithms uses correlations, rather than *ex post* mean returns, to form portfolios, the dispersion in out-of-sample returns is similar to that generated using single firm characteristics such as size or book-to-market, or sorts along multiple characteristics. Second, the set of portfolios constructed seems to generate a relatively well-conditioned covariance matrix of returns. This fact implies increased precision in the calculation of efficient frontiers and related Sharpe ratios. Moreover, while it has been shown in previous papers that measurement error leads to an upward bias in the Sharpe ratio associated with the investor's opportunity set, we show that this bias increases sharply as the covariance matrix becomes less well-conditioned. Thus, the conditioning of the covariance matrix generated by a particular set of basis assets plays an important role in the reliability with which we can draw inferences regarding the performance of asset pricing models, the composition of frontier portfolios, and the measurement of expected, and hence abnormal, returns. These results highlight a little-emphasized issue in basis asset

[15] Recall that we identify cluster numbers through time by maximizing the extent to which the membership in a cluster remains constant. Consequently, the persistence in cluster membership reported here is, by construction, the highest we can attain from year to year when forming clusters annually. Empirically, if we use other rules for assigning cluster numbers, we still find significant persistence in the clusters. For example, when we use Sharpe ratios to identify clusters through time, the persistence in the ten portfolios declines slightly, to 19.1% in the overlapping sample, and is significantly higher than one would expect by random chance. Interestingly, when we used Sharpe ratios to identify clusters, the probability that firms would remain in the same clusters in the nonoverlapping sample was higher, at 13.5%. This suggests that the factors associated with a firm's observed Sharpe ratio are stable over longer horizons, possibly due to persistence in volatility.

construction: the focus on generating dispersion in *ex post* mean returns may come at the cost of a reduction in the precision of inferences.

In specification tests of the CAPM and the three-factor asset pricing model, we find that there appear to be significant components of cluster portfolio returns that are not explained by the three-factor model; in cross-sectional regression tests, the cluster portfolio returns are positively and significantly related to CAPM and CCAPM betas, and this result is robust to the inclusion of characteristic variables such as size and book-to-market. In contrast, size- and book-to-market-sorted portfolio returns are significantly related to characteristics, and the returns of factor-mimicking portfolio returns related to those characteristics, while they are negatively related to other risk measures.

Finally, the cluster portfolios are correlated with firm characteristics such as size, book-to-market, and beta. This result suggests that the clustering algorithm, with its focus on returns and return correlations, is capable of sorting firms into groups with important economic differences. Equally importantly, the algorithm is able to sort securities into groups without requiring the researcher to *ex ante* identify the firm characteristics of interest, perhaps generating a data-snooping bias. Rather, the clustering algorithm uses only returns data; more specifically, it uses only the correlation between the returns of individual securities. And, in contrast to the sorting methods currently in use, the algorithm does not appear to favor an increased number of basis portfolios, with its attendant increased risk of data snooping.

Despite the fact that the clustering algorithm does not identify factors, the increased precision with which the cluster portfolios allow the investor's opportunity set to be identified has several advantages. These advantages include increased precision in measuring abnormal returns, testing asset pricing models, generating efficient portfolio weights, and generating the covariance matrix from which, in future work, "real" economic factors can be estimated.

Appendix A: Implications of Conditioning for Pricing Inferences

The well-known constrained minimization problem faced by a rational, risk-averse single-period investor with quadratic utility making a portfolio choice is

$$\begin{aligned} &\min_{\omega} \quad \omega'\Sigma\omega \\ &s.t. \quad \omega'\iota = 1, \\ &\qquad\;\; \omega'\mu = e, \end{aligned} \tag{A1}$$

where μ is a vector of expected returns, ι is a conforming vector of ones, ω is the vector of portfolio weights (and the solution to the problem), e is a scalar portfolio expected return, and the covariance matrix of returns is denoted by Σ. The solution is given by

$$\omega = \Sigma^{-1}(\lambda\iota + \theta\mu), \tag{A2}$$

where λ and θ are the Lagrange multipliers for the constraints in the minimization problem given above.

Now, consider the situation in which mean returns μ are measured with error; denote the measurement error by $\Delta\mu$. Through the first-order condition to the minimization problem above, the measurement error in μ will lead to a related solution error in ω. Call this solution error $\Delta\omega$. This error $\Delta\omega$ is given by

$$\Delta\omega = \Sigma^{-1}\theta\Delta\mu.$$

Note that the change in the solution is related to the (inverse of the) covariance matrix, as well as the measurement error in μ. It can be shown that the relative size of this solution error is bounded by a function of the relative size of the measurement error $\Delta\mu$ and a property of the covariance matrix:

$$\frac{||\Delta\omega||}{||\omega||} \leq \kappa \frac{||\Delta\mu||}{||\mu||}, \tag{A3}$$

where κ is the condition number of the matrix Σ.[16] The condition number can be calculated as

$$\text{cond}(\Sigma) = \kappa(\Sigma) = ||\Sigma|| \cdot ||\Sigma^{-1}||. \tag{A4}$$

Thus, the relative importance of the solution error is related to the conditioning of the covariance matrix of returns.

Equation (A3) highlights the role of the conditioning of the covariance matrix in gauging the influence of measurement error in the data. Intuitively, the condition number of the covariance matrix of returns is related to the collinearity of the data. For example, a matrix which is not of full rank has an infinite condition number. At the other extreme, the minimum condition number is 1; the identity matrix has a condition number of 1. The higher the condition number of the matrix, the more collinearity in the data and the less precise the solution obtained using the matrix. If the condition number is high enough, the solution may be meaningless.[17]

The clustering algorithm we use seeks to reduce the correlation, or the collinearity, between the portfolio groups in the data. Consequently, one measure of the method's success is to examine the condition number of the covariance matrix Σ, and compare this result with the condition number of alternative constructions of Σ, formed from different groups of securities.

Appendix B: Simulation Design

In our primary simulation experiment, we add a measurement error with a zero mean and standard deviation of σ_μ to the two sets of basis asset returns μ_C and μ_S, where μ_C is the vector of mean returns of the cluster portfolios and μ_S is the vector of mean returns for the characteristic-sorted portfolios. For both sets of basis assets, σ_μ is set to 1, 5, and 10 basis points. The magnitude of the simulated perturbations in mean returns is designed to be similar to, or smaller than, common market frictions. For example, MacKinlay uses σ_μ estimates of 7 and 10 basis points and argues that these magnitudes are plausible, even conservative, and are consistent with spreads that could

[16] There are many norms that one can use to calculate condition numbers; as Heath (2002) notes, while the numerical value of the condition number depends on the particular norm used, because they differ from each other only by a fixed constant, they are "equally useful as quantitative measures of conditioning." Throughout the paper, we calculate the condition number using the 2-norm of the matrix. This condition number can also be calculated as the ratio of the maximum and minimum eigenvalues of Σ.

[17] The sensitivity of the solution to measurement error in the covariance matrix, rather than the vector of mean returns, is also driven by the condition number of the covariance matrix. More generally, if there is measurement error $\Delta\Sigma$ in the covariance matrix, as well as measurement error $\Delta\mu$ in the vector of mean returns, the sensitivity of the solution to these two measurement errors is given by

$$\frac{||\Delta\omega||}{||\omega||} \leq f(\kappa_2(\Sigma)) \left[\frac{||\Delta\mu||}{||\mu||} + \frac{||\Delta\Sigma||}{||\Sigma||} \right].$$

arise from data snooping. As another benchmark, if we bootstrap the mean returns in our sample, the average standard deviation across 5000 iterations, at 23 basis points, is larger than any of our three simulated experiments.

Five thousand runs of the simulations are performed for each value of σ_{μ} and for each set of portfolios; in the body of the paper, we primarily discuss the simulation performed under the assumption that the simulated measurement errors are cross-sectionally independent. In other experiments, we allow the measurement errors to be correlated across portfolios. Specifically, we begin by assuming that measurement errors are positively correlated, and, in fact, assume that measurement errors have the same correlation structure as the raw portfolio returns. For this experiment, we examine only the case where σ_{μ} is set equal to 5 basis points and the number of portfolios is 25.

Assuming that the measurement error is highly positively correlated across portfolios is equivalent to assuming that there is a systematic component to the error; if we assume that the correlation structure in the measurement errors across portfolios is identical to the correlation structure in the raw returns, in the MacKinlay framework, we are inserting a direct link between the measurement error in, and consequently the value of, the observed portfolio return and its variance. Not surprisingly, the reimposition of this link reduces the bias in the observed Sharpe ratio in the simulation, and consequently reduces the conditioning advantage of the cluster portfolios. Specifically, if we set the correlation level to be the same as we observe in the raw cluster portfolio returns, the increase in the mean Sharpe ratio (and hence the increase in the bias) of the characteristic-sorted portfolios decreases to 0.02, from the increase of 0.037 observed when measurement errors are assumed to be uncorrelated. If we set the correlation level higher, at the level observed in the raw size- and book-to-market-sorted portfolio returns, the increase in the mean Sharpe ratio for the characteristic-sorted portfolios is essentially zero.

Note that if we impose a correlated error structure that does not require that the mean and variance of the measurement error are linked, the magnitude of the correlation in the error terms has a less dramatic effect on the Sharpe ratio of the characteristic-sorted portfolios. For example, if we set the correlation level to be 0.8, the increase in the mean Sharpe ratio of the characteristic-sorted portfolios falls from 0.037 to 0.006. If we set the correlation level to be 0.4, the increase in the mean Sharpe ratio falls from 0.037 to 0.021. In all cases, the mean increase in the Sharpe ratio of the better-conditioned cluster portfolios is much smaller, never rising above 0.004.

Lo and MacKinlay (1990) argue that data-snooping biases may result in a sorting of measurement errors as well; in this case, measurement errors may be negatively correlated across basis asset portfolios. Not surprisingly, if we set the correlation in the measurement errors to be negative, the effect of conditioning differences increases. For example, if the correlation is set to −0.04, the increase in the Sharpe ratio for the characteristic-sorted portfolios increases from 0.037 to 0.039.[18] Our results suggest that, while conditioning continues to matter, the correlation in, and consequently the source of, measurement error in returns data is an important determinant of the advantage that better conditioning provides.

Appendix C: Alternative Data Reduction Techniques

We compare the ability of these clustering algorithms and principal components analysis to capture important features of the investor's *ex ante* opportunity set. Both principal components and cluster portfolios are linear combinations of individual securities. However, the principal components extracted from returns data are designed to account for the maximum variation in returns (in-sample), with each component being orthogonal to all others. Hence, the first component will be along the dimension of greatest return variation in the sample, the second (orthogonal) component will capture the next largest variation, and so on. The principal components can be thought of

[18] Note that reducing the correlation level below −0.04 creates, for some iterations of the simulation, a covariance matrix that is not positive definite unless we impose further structure on the measurement errors.

as weighted combinations of individual securities, where the weights are unconstrained, and all securities can (and generally do) play a role in each principal component.

In contrast, while the clustering algorithm is designed to minimize the in-sample correlation between clusters, these clusters will not be orthogonal to one another. In addition, an individual security will be placed in only one cluster, and the weight given to an individual security is always positive (and is given by the value-weight of the security in the portfolio). Thus, the cluster portfolios are much more constrained than the principal components in their membership, and their weighting scheme.[19]

There is another difference between our clustering technique and principal components analysis. Our clusters are based on correlations, while principal components analysis uses the covariance matrix. Since principal components analysis is not scale invariant, the components formed are different, and there is no convenient correspondence between the two decompositions.

In the end, it is an empirical question as to which data reduction technique does a better job of capturing investors' opportunity sets, or which lends itself better to an understanding of fundamental economic risks. We compare the opportunity set formed using principal components analysis with that formed using our clustering algorithm. Specifically, we use the method of Connor and Korajczyk (1988), who extract principal components from a $T \times T$ covariance matrix formed from monthly individual security returns over five-year subperiods. Following their method, we require that a security have all 60 months of returns data available to be included in the sample; this requirement is applied to both the clustering algorithm and the principal component analysis for the purposes of this comparison. Our in-sample results are similar to theirs. For example, we find that when we regress the equal-weighted market return on the first factor, we obtain R^2 values above 0.98 for each of the subperiods in our sample. Thus, as Brown (1989) shows, the first principal component is virtually identical to the equal-weighted index, regardless of the true underlying structure of returns.[20]

We are interested in the ability of these methods to capture the investor's *ex ante* opportunity set. Consequently, we wish to examine the out-of-sample frontier constructed from the extracted principal components, with the out-of-sample frontier constructed from cluster portfolios. To construct the out-of-sample frontier generated by principal components, we construct mimicking portfolio weights for each of the principal components. We then apply these weights to the assets in months t through $t + 11$ to form K factor-mimicking portfolios. We form frontiers from these portfolios, for increasing K.

We find that the factor-mimicking portfolios of the first ten principal components generate an out-of-sample dispersion in returns of 43 basis points per month, and an efficient frontier with a Sharpe ratio of 0.24. The dispersion in returns is 4 basis points less than the dispersion in returns observed with ten cluster portfolios, and the Sharpe ratio is larger than the Sharpe ratio generated by the cluster portfolios of 0.196. In addition, the condition number of the covariance matrix generated by the factor-mimicking portfolios, at 9, is not significantly different from the condition number observed using ten cluster portfolios. However, the weights attached to individual securities in these factor-mimicking portfolios vary substantially across securities, particularly compared to the value-weights used in the cluster portfolios. For example, on average, more than a third of securities in all of the mimicking portfolios have negative weights, and the average standard deviation of

[19] Note that weight constraints on the cluster portfolios may be an advantage, rather than a burden. For example, Jagannathan and Ma (2003) argue that constraining portfolio weights to be nonnegative may actually help when forming optimal portfolios, since extreme weights may reflect large estimation errors. They show that such constraints can be equivalent to a shrinkage estimator applied to the sample covariance matrix.

[20] In January 1992, one stock in our sample (OCG Technologies) has a 2400% return. This one observation affects the principal components analysis; essentially, it causes the second principal component (and in one month the third principal component) to be significantly more important during the five-year window that includes this month. However, when we regress the equal-weighted market return on the first two components in this subperiod, we obtain R^2 that are quite similar to those obtained from using the first principal component only in all other subperiods. This effect occurs only in our in-sample results. Out-of-sample results are not affected by this single large return in one small stock.

the weight given to a particular firm through time is almost three times the standard deviation of the weights in the cluster portfolios. As Brown, Goetzmann, and Grinblatt (1997) point out, these differences make the principal components' portfolios difficult to form, or characterize.[21] Thus, despite the binding constraints imposed on the cluster portfolios, they appear capable of generating an opportunity set that is comparable with principal components. In addition, these constraints make the portfolios easier to form, and may make the sources of comovement, and hence the underlying economic factors, easier to identify.

References

Berk, J. B. 2000. Sorting Out Sorts. *Journal of Finance* 55:407–27.

Brown, S. J. 1989. The Number of Factors in Security Returns. *Journal of Finance* 44:1247–62.

Brown, S. J., W. N. Goetzmann, and M. Grinblatt. 1997. Positive Portfolio Factors. Working Paper F-57, Yale School of Management.

Cass, D., and J. Stiglitz. 1970, The Structure of Investor Preference and Asset Returns, and Separability in Portfolio Allocation: A Contribution to the Pure Theory of Mutual Funds. *Journal of Economic Theory* 2:122–60.

Cochrane, J. H. 2001. *Asset Pricing*. Princeton, NJ: Princeton University Press.

Connor, G., and R. A. Korajczyk. 1988. Risk and Return in an Equilibrium APT: Application of a New Test Methodology. *Journal of Financial Economics* 21:255–89.

Conrad, J., M. Cooper, and G. Kaul. 2003. Value vs. Glamour. *Journal of Finance* 58:1969–95.

Daniel, K., and S. Titman. 1997. Evidence on the Characteristics of Cross Sectional Variation in Stock Returns. *Journal of Finance* 52:1–33.

Daniel, K., and S. Titman. 2005. Testing Factor-Model Explanations of Market Anomalies. Working Paper, Kellogg School of Management, Northwestern University.

Edelman, A., and B. D. Sutton. 2005. Tails of Condition Number Distributions. *SIAM Journal of Matrix Analysis and Applications* 27:547–60.

Fama, E., and K. French. 1993. Common Risk Factors in the Returns on Stocks and Bonds. *Journal of Financial Economics* 33:3–56.

Fama, E., and K. French. 1996. Multifactor Explanations of Asset Pricing Anomalies. *Journal of Finance* 51:55–84.

Fama, E., and J. MacBeth. 1973. Risk, Return, and Equilibrium: Empirical Tests. *Journal of Political Economy* 81:607–36.

Fama, E. F., J. L. Davis, and K. R. French. 2000. Characteristics, Covariances, and Average Returns: 1929 to 1997. *Journal of Finance* 55:389–406.

Gibbons, M., S. Ross, and J. Shanken. 1989. A Test of the Efficiency of a Given Portfolio. *Econometrica* 57:1121–52.

Heath, M. T. 2002. *Scientific Computing: An Introductory Survey*, 2nd ed. New York: McGraw-Hill.

Jagannathan, R., and T. Ma. 2003. Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps. *Journal of Finance* 58:1651–83.

Jagannathan, R., and Z. Wang. 1996. The Conditional CAPM and the Cross-Section of Expected Returns. *Journal of Finance* 51:3–54.

[21] In the out-of-sample period, the first component continues to be significantly and positively correlated with the equal-weighted market portfolio, although the correlation declines sharply, to 0.41.

Kan, R. 1999. On the Explanatory Power of Asset Pricing Models Across and Within Portfolios. Working Paper, University of Toronto.

King, B. 1966. Market and Industry Factors in Stock Price Behavior. *Journal of Business* 39:139–90.

Lo, A., and A. C. MacKinlay. 1990. Data-Snooping Biases in Tests of Financial Asset Pricing Models. *Review of Financial Studies* 3:431–68.

MacKinlay, A. C. 1995. Multifactor Models Do Not Explain Deviations from the CAPM. *Journal of Financial Economics* 38:3–28.

Markowitz, H. 1952. Portfolio Selection. *Journal of Finance* 7:77–91.

Ormerod, P., and C. Mounfield. 2000. Localised Structures in the Temporal Evolution of Asset Prices. In *New Approaches to Financial Economics*, Santa Fe Conference.

Ross, S. 1978. Mutual Fund Separation in Financial Theory: The Separating Distributions. *Journal of Economic Theory* 17:254–86.

Ward, J. H. 1963. Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association* 58:236–44.