Board of Governors of the Federal Reserve System

International Finance Discussion Papers

Number 1070

December 2012

Firm Characteristics and Empirical Factor Models: A Data-Mining Experiment

Leonid Kogan and Mary Tian

Firm Characteristics and Empirical Factor Models: a Data-Mining Experiment*

Leonid Kogan †
Mary Tian ‡

December 2012

Abstract

"A three-factor model using the standardized-unexpected-earnings and cashflow-to-price factors explains 15 well-known asset pricing anomalies." Our data-mining experiment provides a backdrop against which such claims can be evaluated. We construct three-factor linear pricing models that match return spreads associated with as many as 15 out of 27 commonly used firm characteristics over the 1971-2011 sample. We form target assets by sorting firms into ten portfolios on each of the chosen characteristics and form candidate pricing factors as long-short positions in the extreme decile portfolios. Our analysis exhausts all possible 351 three-factor models, consisting of two characteristic-based factors in addition to the market portfolio. 65% of the examined factor models match a larger fraction of the target return cross-sections than the CAPM or the Fama-French three-factor model. We find that the relative performance of the complete set of three-factor models is highly sensitive to the sample choice and the factor construction methodology. Our results highlight the challenges of evaluating empirical factor models.

Keywords: Anomalies, Factor Model, Data-mining, Firm Characteristic

JEL Classification: G12

*We thank seminar participants at the Finance Forum workshop at the Federal Reserve Board of Governors. The views in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System or of any other person associated with the Federal Reserve System.

†NBER and MIT Sloan School of Management, lkogan@mit.edu.

‡Division of International Finance, Federal Reserve Board of Governors, mary.tian@frb.gov.

1 Introduction

Empirical asset pricing literature has documented many examples of firm characteristics being able to predict future stock returns. When not accounted for by standard asset pricing models, such patterns are often interpreted as anomalous. It is challenging to develop meaningful theoretical explanations of the observed patterns in returns.[1] In contrast, the long-short portfolios constructed by sorting firms on various characteristics – the "c-factors", often named after the sorting variable – provide readily available inputs into empirical factor models. By searching through the firm characteristics known to be associated with large spreads is stock returns, it is relatively easy to construct seemingly successful empirical factor pricing models.

When we hear of a new c-factor model with N factors that "explains" M of the well-known anomalies, how should we evaluate such a result? Is there a quantitative threshold for the M-to-N ratio above which such a result strongly points to an economically important source of systematic risk, even without a solid theoretical foundation? The ease of construction of c-factor models and virtually unlimited freedom in selecting test assets provide fertile ground for data mining.[2] In this paper we quantify just how easy it is to generate seemingly successful empirical c-factor models. Our findings imply that it is extremely difficult to evaluate factor pricing model based solely on their pricing performance, and one must emphasize the theoretical and empirical foundation for their economic mechanism.

We systematically mine the 1971-2011 historical sample under a specific set of rules designed to be representative of the commonly used empirical procedures. We consider 27 firm characteristics proposed in the literature as predictive variables for stock returns (see section 2 and Appendix A for the list of the characteristics, with references to the relevant

[1] "Meaningful" is an important qualifier here: it is not hard to come up with an ad hoc ex-post rationalization of why a particular firm characteristic may proxy for exposure to a risk factor. A compelling theoretical explanation should identify the economic mechanism giving rise to such a factor, provide alternative testable implications of this mechanism, as well as a rationale for why other firm characteristics are correlated with firms' exposures to the proposed risk factor.

[2] Many studies in the literature warn of the dangers of data mining biases, particularly in the context of return predictability, e.g., Black (1993), Lo and MacKinlay (1990), Ferson (1996), Lewellen, Nagel, and Shanken (2006), Novy-Marx (2012).

literature). Some of these characteristics have been proposed as candidate empirical proxies for systematic risk exposures, others as likely proxies for mispricing – we do not discriminate based on the merits of the original motivation. To qualify as a contender for our data-mining exercise, a firm characteristic simply needs to be a subject of an academic publication.

We rank firms into ten portfolios based on each of the 27 characteristics and define the associated return factors as return differences between the tenth and the first decile portfolios. We then tabulate the pricing performance of all possible three- and four-factor models, each consisting of the market portfolio and two or three factors respectively, chosen out of the set of 27. We thus consider a total of 351 alternative three-factor models, and 2,925 four-factor models.

If a pricing model is not rejected by testing it against a cross-section of portfolios sorted on a particular firm characteristic, we say that this model *matches* such a cross-section. We find that it is relatively easy to construct a three-factor model that match more than half of the 25 target cross-sections of returns over the full sample (we exclude the cross-sections used to form the model factors from the set of target cross-sections).

The best-performing model over the entire sample, by the total number of matched cross-sections, includes the factors based on unexpected earnings and the cash flow-to-price ratio. It matches 15 out of 25 return cross-sections. Each of the top-twenty models reported in Table 5 matches return cross-sections based on each of 12 or more different characteristics.[3] Four-factor models achieve slightly better coverage, with the top model matching 16 out of 24 cross-sections, and the worst of the top-twenty models matching 14. For comparison, the CAPM and the Fama and French (1993) three-factor model both match eight out of 27 return cross-sections (we do not exclude any test assets when evaluating these reference models).

As expected in a data mining exercise, performance of the c-factor models tends to be fragile. It is highly sensitive to the sample period choice and the details of the factor construction. In particular, there is virtually no correlation between the relative model performance

[3]We summarize performance of all 351 models in an on-line document, `http://tinyurl.com/d43mf3h`.

in the first and the second halves of the 1971-2011 sample period. Likewise, using a two-way sort on firm stock market capitalization (size) and characteristics to construct model return factors, an often used empirical procedure, similarly scrambles the relative model rankings. Such lack of stability suggests that our data-snooping algorithm tends to pick spurious winners among the set of all possible models without revealing a robust underlying risk structure in returns. This does not mean that all of the better-performing models in our analysis are spurious and theoretically unjustifiable. Some of the many models we enumerate in this study are likely to capture economically meaningful sources of risk – we just cannot identify which of them do, based solely on the models' pricing performance.

This paper is organized as follows. Section 2 describes the data and methodology. Section 3 examines the overall factor structure of characteristic-sorted portfolios and the ability of c-factor models to capture cross-sectional differences in average returns on various characteristic-sorted portfolios. Section 4 concludes.

2 Data and Methodology

In this section, we describe the data used in our analysis and our empirical methodology.

Data on annual and quarterly firm fundamentals are from the CRSP/Compustat Merged database. Monthly data on firm-level stock returns, shares outstanding, and volume are from the CRSP database. Aggregate market liquidity data are from Pastor and Stambaugh (2003). Our sample period is 1971-2011, with subsample periods 1971-1991 and 1992-2011.

We consider a total of 27 firm characteristics, which we informally partition into seven groups:

1. valuation: size (SIZE), book to market (BM), dividend-to-price (DP), earnings-to-price (EP), cash flow-to-price (CP)

2. investment: investment-to-assets (IA), asset growth (AG), accruals (AC), abnormal investment(AI), net operating assets (NOA), investment-to-capital (IK), investment growth (IG);

3. prior returns: momentum (MOM), long-term reversal (LTR);
4. earnings: return on assets (ROA), standardized unexpected earnings (SUE), return on equity (ROE), sales growth (SG);
5. financial distress: Ohlson score (OS), market leverage (LEV);
6. external financing: net stock issues (NSI), composite issuance (CI);
7. other: organization capital (OK), liquidity risk (LIQ), turnover (TO), idiosyncratic return volatility (VOL), market beta (BETA).[4]

The definitions and construction of the characteristics are contained in Appendix A.

After dropping all firms in the financial sector (SIC 6000-6999), we sort remaining firms into ten portfolios with respect to each characteristic, thus performing 27 independent one-way sorts. We sort firms every year in June with respect to the underlying characteristic and then compute value-weighted returns of each portfolio from July to June of the next year.[5] We take the difference in value-weighted returns of the high and low portfolios (decile 10 minus decile 1) to form 27 characteristic return factors.[6] Alternatively, we also construct factors by doing a sequential double-sort on size and then the characteristic: firms are separated into either big or small firms, and subsequently within each group, sorted into ten portfolios with respect to the characteristic. Then, we construct each factor as the equal-weighted average of the high minus low portfolio within the big and small size group. Our base set of results use factors constructed from the one-way sort; we compare results using the alternative double-sort factor construction in Section 3.3.

[4]Strictly speaking, market beta is a measure of risk, and is not what is typically taken as a firm characteristic. We include market beta as one of the sorting variables because of the recent resurgence of interest in the failure of CAPM to price the market-beta sorted portfolios, (e.g., Black, Jensen, and Scholes, 1972; Frazzini and Pedersen, 2011; Baker, Bradley, and Wurgler, 2011). Similarly, idiosyncratic return volatility is a return statistic rather than a firm characteristic observable at a point in time. We include idiosyncratic volatility because of its striking ability to forecast future stock returns, e.g., Ang, Hodrick, Xing, and Zhang (2006).

[5]We perform a monthly sort for idiosyncratic volatility, following Ang et al. (2006).

[6] In particular, to be consistent, we construct the size and book-to-market factors in this manner, which we call $SIZE$ and BM, instead of using the standard Fama-French factors SMB and HML.

We create three-factor models by taking the market portfolio and choosing two factors among our 27 return factors. Overall, this generates a universe of 351 linear three-factor models. In addition to the complete list of all possible three-factor empirical models, we also consider the CAPM; the Fama-French three-factor model; and a model consisting of the market portfolio and the first two principal component vectors from the span of the 27 factor returns. While CAPM is perhaps the most commonly used theoretical benchmark, the other two models are empirical factor models.

We test each factor model's ability to match the average return differences across portfolios sorted on each characteristic using a standard time-series regression framework. In particular, following Gibbons, Ross, and Shanken (1989), for each characteristic we regress excess returns on the ten characteristic-sorted portfolios on the returns of the three factors:

$$r_n^i - r_f = \alpha^i + \beta_{n,MKT}^i (r_{MKT} - r_f) + \beta_{n,j}^i r_j + \beta_{n,k}^i r_k + \epsilon_n^i, \tag{1}$$

where $i = 1, ..., 10$ indexes the decile portfolios sorted on the characteristic number n, $n = 1, ..., 27$; j and k are the c-factors formed on characteristics j and k respectively, $j < k$. We perform the Gibbons et al. (1989) F-test of the hypothesis that $\alpha_1 = \alpha_2 = \cdots = \alpha_{10} = 0$. We say that a three-factor model using c-factors j and k is able to match, or capture, the cross-section of returns on portfolios sorted on characteristic n if the p-value of the F-test, $p_{n,j,k}^F$, exceeds ten percent.

For each three-factor model, we exclude the target cross-sections based on the two characteristics used to create the c-factor portfolios. Thus, for each three-factor model consisting of the market portfolio and two c-factors, we run the time-series regression over the remaining 25 sets of characteristic-sorted decile portfolios. We then compute a measure of the fraction of all the cross-sections that each factor model is able to match.

We consider two measures of performance, each defined as a weighted sum over the matched cross-sections:

$$\frac{\sum_{n=1, n \neq j, n \neq k}^{27} 1_{[p_{n,j,k}^F > 0.1]} w_n}{\sum_{n'=1, n' \neq j, n' \neq k}^{27} w_{n'}}.$$

For each of the measures, we define the weights w_n as:

1. (Equal-weighted) Each characteristic gets an equal weight of 1/25.

2. (Characteristic Matching Frequency) Each characteristic's weight equals 1 minus the proportion of factor models that can match the cross-section based on this characteristic,

$$w_n = 1 - \frac{\sum_{\{j=1,k=2\},j<k,j\neq n,k\neq n}^{27} 1_{[p^F_{n,j,k}>0.1]}}{\#\{j,k : 1 \leq j \leq 26, 2 \leq k \leq 27, j < k, j \neq n, k \neq n\}}$$
$$= 1 - \frac{\sum_{\{j=1,k=2\},j<k,j\neq n,k\neq n}^{27} 1_{[p^F_{n,j,k}>0.1]}}{325}.$$

In the first method, the fraction of matched return cross-sections is simply the number of return cross-sections the model can match divided by the total number of target cross-sections.

The second weighting scheme places higher weight on the "harder-to-explain" cross-sections – the cross-sections that are matched by fewer c-factor models. Our motivation for this is two-fold. First, this construction is supposed to alleviate the effect of double-counting caused by the fact that some of the return factors we consider are constructed using closely related firm characteristics, and thus may not be viewed as truly distinct. Placing a higher weight on the harder-to-match cross-sections reduces the relative performance ranking of the models that include c-factors closely related to several other characteristics. Second, c-factor models that match a number of return cross-sections that are viewed as challenging, i.e., are rarely matched by the models proposed thus far, are likely to receive more attention in the literature. Our second weighted measure places higher premium on the mechanically constructed models with such attention-grabbing potential.[7]

Unless otherwise specified, our results utilize the first weighting method.

[7] If a particular pattern in returns is firmly viewed as a true anomaly that is not supposed to be explained by systematic risk, matching such a cross-section may be seen as evidence against a proposed factor model being risk-based. We abstract from this consideration in our definition of our second performance measure.

3 Properties of Empirical Factor Models

In this section we present the summary statistics of the characteristic-based factor portfolios, examine the ability of linear factor models to capture average returns on these factors, and show which of the factors are the hardest to reconcile with empirical factor models.

3.1 Characteristic-Sorted Portfolios

We present summary statistics of 27 characteristic-based factor portfolios in Table 1. For each firm characteristic c_n, $n = 1, ..., 27$, we first form decile portfolios sorted in the order of increasing characteristic value. All portfolios are value-weighted. We then form the empirical c_n-factor, which is long the top-decile portfolio, and short the bottom-decile portfolio.

For each c-factor, we present the estimates of average returns (Panel A), CAPM alphas (Panel B), and Fama-French alphas (Panel C), together with corresponding t-statistics. All numbers are estimated with monthly data. The table contains the full sample and subsample results.

The first set of results (moving vertically down the table) covers return factors related to firm valuation. This includes the following firm characteristics: firm market capitalization (SIZE), book-to-market ratio (BM), dividend-to-price ratio (DP), earnings-to-price ratio (EP), and cash flow-to-price ratio (CP). Return factors based on BM, EP, and CP generate a statistically significant spread in average returns, which is not captured by the CAPM model.

The second set of characteristics is related to firms' investment and physical assets. This set includes return factors based on investment-to-assets ratios (IA), asset growth (AG), accruals (AC), abnormal investment (AI), net operating assets (NOA), investment over capital (IK), and investment growth (IG). Several of the investment-related characteristics forecast future stock returns. Qualitatively, firms with relatively high investment relative to assets tend to have lower future returns. Factors based on IA, AG, and AC show the strongest effects, which are not captured neither by CAPM, nor by the Fama-French model. These effects persist over both subsamples, although they are somewhat stronger in the first-half of the sample. The factors based on IK and IG have lower statistical significance. The IK factor violates the CAPM over the entire sample and each of the subsamples, while the IG factor is less robust – its return premium is captured by the CAPM in the first-half of the sample. The Fama-French model fits the average returns on both of these factors reasonably well.

The next set includes factors related to prior returns: return momentum (MOM) and long-term reversal (LTR). Returns on the MOM factor are large on average, robust across the subsamples, and not captured by the CAPM and the Fama-French model. Returns on the LTR factor are smaller on average, but violate the CAPM and Fama-French model in different subsample periods.

The next set of factors is related to firms' earnings. This covers return on assets (ROA), standardized unexpected earnings (SUE), return on equity (ROE), and sales growth (SG). Firms with high ROA or high SUE tend to have higher average returns, which is not fully captured by the CAPM and the Fama-French model. For ROA, the patterns are robust across the subsamples, while the patterns for SUE have higher statistical significance in the first subsample. ROE produces weaker patterns of the same sign. Sales growth predicts stock returns with the opposite sign to the other earnings-based characteristics. SG returns violate the CAPM over the entire sample, but are captured by the Fama-French model.

The next set of factors is related to financial distress, sorting firms on their Ohlson score (OS) and market leverage (LEV). OS predicts returns with a negative sign. The magnitude of the average returns of this factor is large, with statistically significant CAPM and Fama-French alphas of -1% per month over the entire and subsample periods. LEV predicts returns with a positive sign and a weakly-significant CAPM alpha of 0.5% per month. The Fama-French model captures the returns on the LEV factor.

The next two factors are related to external financing: net stock issues (NSI) and composite issuance (CI). Both characteristics predict returns negatively, and the resulting factor returns violate both the CAPM and the Fama-French model in both sub-samples and over the entire sample.

The last group contains several firm characteristics that are not immediately related to each other nor to the characteristics covered above. These include organizational capital (OK), liquidity risk (LIQ), turnover (TO), idiosyncratic return volatility (VOL), and market beta (BETA). VOL factor returns are negative, extremely large (-1.4% monthly), and violate both models in both sub-samples. BETA factor has insignificant average returns but weakly significant CAPM alphas.

3.2 Factor Structure of Characteristic-Sorted Portfolios

After observing the average return patterns, we next examine to what extent return factors are related to each other, via principal component analyses (Tables 2 through 4) and factor correlation maps (Figure 1).

Table 2 presents results from a principal component analysis on the 27 return factors. The table shows the proportion of cumulative variation in factor returns that the first n principal components can capture. Over the whole sample period 1971-2011, the first three principal components together can capture 63% of total variation in the 27 return factors; this increases slightly to 69% in the second subsample period. The marginal effects of increasing the number of principal components decrease as we look down the table, adding no more than 5% in explanatory power for each additional component.

Another way to observe the factor correlation structure is through a heatmap representation in Figure 1. Figure 1 shows the matrix of return factor correlations, as well as correlations of individual factor returns with the market portfolio and the first three principal components extracted from the return factors. Darker areas represent higher correlation.

Certain blocks of factors stand out with high within-block correlations. For instance, over the full sample period, 1971-2011, valuation-related factors are highly correlated with each other, as are investment-related, earnings-related, and issuance-related factors. Factors are generally more correlated with each other in the second-half of the sample than in the first. This is consistent with better performance of empirical pricing models in the second-half of the sample, which we discuss below. Some factors stand out as having relatively low correlation with all other factors. These include accruals (AC), momentum (MOM), standardized unexpected earnings (SUE), and liquidity risk (LIQ).

Overall, we conclude that there is a substantial degree of comovement among the 27 factors, indicated both by the high amount of total variance explained by the first three principal components of the covariance matrix, and by the correlation patterns among economically related groups of factors.

Table 3 shows the factor loadings for the first three principal components extracted from the set of 27 factor returns. Over the whole sample period 1971-2011, we observe that the first principal component (PC1) has the highest loading from the idiosyncratic volatility (VOL) factor, followed by market beta (BETA), and Ohlson score (OS). The second principal component (PC2) captures the valuation-related factors (SIZE, BM, DP), asset growth (AG), investment-to-capital (IK), long-term reversal (LTR), market leverage (LEV), turnover (TO), and market beta (BETA). The third principal component (PC3) has a very high loading from the momentum (MOM) factor, especially for the second subsample period.

To see how closely each of the characteristic-based factors is spanned by the leading principal components in the entire cross-section of 27 factors, we regress each factor on a benchmark three-factor model consisting of the market portfolio excess returns and the first two principal components. In Table 4, we present the intercept coefficient, t-statistic, and R^2 from the regression for the whole sample 1971-2011 and subsamples 1971-1991 and 1992-2011.

Over the full sample period, there is a significant degree of heterogeneity in the properties of characteristic-based factors. For some, such as IK, ROA, ROE, OS, TO, VOL, BETA, the benchmark three-factor explains over 70% of variance. Among these, only TO and VOL have economically and statistically significant alphas with respect to the benchmark model.

A few factors are practically uncorrelated with all the components of the benchmark model. Regressions of AC, AI, MOM, SUE, and LIQ on the benchmark model have R^2 of ten percent or less. All of these except AI have significant alphas with respect to the benchmark model. The results in Table 4 are largely robust over the two subsamples.

In summary, our analysis of factor correlation suggests that certain groups of characteristic-based factors can be effectively related to a low-dimensional factor model, but the overall pattern of results indicates that there is significant remaining heterogeneity among the factors that a parsimonious model may not be able to capture. In the following section we further quantify these observations.

3.3 Pricing Performance of Empirical Factor Models

In this section we evaluate the empirical performance of all possible c-factor models constructed based on our set of 27 characteristics. As we show in the previous section, the corresponding 27 c-factors exhibit a non-trivial factor structure. Therefore, several of the three-factor models may potentially account for the observed average returns differences within many of the 27 characteristic-sorted portfolio cross-sections. Moreover, since we do not impose any prior theoretical restrictions on the admissible models, mining through all of 351 possible three-factor models is likely to unearth a few with particularly good in-sample performance. Thus, while some of the empirical relations between the 27 c-factors are due to the fundamental economic links and therefore the observed performance of certain c-factor models can be grounded in standard theory, it is also clear that the best observed in-sample performance of c-factor models benefits from a positive bias introduced by data-mining.

Our data-mining exercise is explicit and exhaustive across the space of the 27 characteristics we consider. One can therefore get a sense of the level of performance that can be achieved by a mechanical search across all candidate models. Evaluating the empirical c-factor models proposed in the literature is a lot harder because of the lack of information on how the c-factors and the test portfolios have been chosen among all the possible alternatives. This is not necessarily a targeted critique of specific studies – data snooping is a well known and hard-to-control side-effect of the research process dynamics at the community level.

Table 5 lists twenty best-performing c-factor models, where performance is measured by the equal-weighted performance measure defined in Section 2. Over the full sample period, the most successful model uses standardized unexpected earnings (SUE) and cashflow-to-price (CP) factors, and captures return differences associated with 60% of the considered characteristics (a total of 15 out of 25 test cross-sections). The model ranked in the twentieth place includes asset growth (AG) and earnings-over-price (EP) factors, fitting 48% of the characteristic-sorted cross-sections. In comparison, both the single-factor CAPM and the Fama-French three-factor model, span 30% of the characteristics (a total of eight), placing them behind 65% of all possible three-factor models in this universe. The bottom line is that over the 1971-2011 sample period, many randomly constructed empirical three-factor models comfortably "outperform" both the CAPM or the Fama-French model by capturing average return differences among portfolios sorted on as many as fifteen characteristics on our list.

Over the second half of the sample, three-factor models fit average returns on the characteristic-sorted portfolios much better than over the full sample, with the best-performing models matching as many as 84% of the test cross-sections. This compares to 80% for the first-half of the sample. The relatively high "success" rate over shorter samples is to be expected, given the lower statistical power to reject the null of zero model alphas in shorter samples. What is informative is whether the same models tend to exhibit high success rates over the sub-samples – we investigate such model stability below.

Figure 2 displays the distribution of performance across the c-factor models over the full sample and the two subsample periods. We use both the equal-weighted method and the characteristic matching frequency method to measure model performance (see the definitions in Section 2). For comparison, we indicate the relative performance ranking of the CAPM and the Fama-French three-factor model relative to all the three-factor models we consider. Over the full sample (panel (a)), the median-performing three-factor model is able to match 32% of the 25 target portfolio cross-sections, while the median factor model in the first and second-half sample (panel (c) and (e)) matches 44% and 56% respectively. The Fama-French model outperforms the CAPM model over the first half of the full sample while substantially underperforming the CAPM over the second half.

Figure 3 provides a more detailed graphical illustration of the performance of various three-factor models. The models are ordered along the horizontal axis in order of increasing performance (based on the proportion of characteristic-sorted cross–sections matched); characteristics are ordered along the vertical axis in order of increasing matching difficulty (measured as the fraction of all three-factor models able to match the return cross-section generated by sorting stocks on a given characteristic). Both the performance measure, and the frequency with which three-factor models match each cross-section are listed in parentheses along each axis. Each cell (i, j) on the figure is shaded black if the c-factor model i is able to match the cross-section based on characteristic j; shaded gray if the c-factor model i is unable to match the cross-section based on characteristic j, and shaded white if factor model i includes a factor constructed using characteristic j.

A few patterns are apparent. Return-forecasting ability of several characteristics, including SG, BETA, ROE, TO, OK, DP, LEV, BM, and LTR, is relatively easy to capture using empirical c-factor models – most of the randomly constructed three-factor models fit the average returns of decile portfolios sorted on these characteristics. A few characteristics generated particularly challenging cross-sections of test portfolios, matched only by the few highest-ranked models. These include AI, ROA, and IK. Several characteristics are virtually impossible to reconcile with empirical three-factor models constructed using our procedure. These are VOL, IG, CI, OS, IA, and MOM.

Return momentum (MOM) is the most challenging characteristic according to our measure: none of the three-factor models (that don't include a MOM factor) can capture it in the full sample or in the first half of the sample period, and only 11% of the models can capture it in the second half of the sample. The other characteristics seem to be more or less difficult to span depending on the subsample. For instance, while only 7% of the three-factor models match the OS cross-section in the first half of the sample period, 71% of all models can match it in the second half. Such lack of stability is consistent with the spurious nature of performance of many of the randomly constructed c-factor models.

3.4 Model Stability and Robustness

Table 7 quantifies the (in)stability of c-factor models' performance across the two subsamples: the correlation between model performance in the two subsamples ranges between 11%

and 16%, depending on the characteristic weighting method and the notion of correlation statistic. The low degree of correlation in relative model performance across the two subsamples is partly due to the sampling errors, but it also suggests that performance of many models in our set may be spurious.

Another possibility for data-mining is associated with the choice of the empirical procedure for return factor construction. Thus far we have used a straightforward procedure for constructing return factors as long-short portfolios of the top and bottom deciles of stocks sorted on each characteristic. One popular alternative approach, following Fama and French (1993), prescribes a two-dimensional sort: first on firm size and then on a characteristic (in case of Fama and French (1993), the characteristic is the book-to-market ratio). We apply a conceptually similar approach in our setting. Specifically, for each characteristic, we first sort firms into big and small (big firms are above the median in market capitalization, small firms are below), form 10-1 long-short portfolios within each size class, and then average the returns on the two long-short portfolios to construct a return factor.

In Table 9, we report cross-sectional correlations of performance between the 351 empirical factor models formed using our univariate factor construction method and the corresponding models with factors formed via the double-sorting procedure. While there is no strong theoretical rationale for using one method of factor construction over the other, the correlation in empirical model performance across the two methods of forming return factors is strikingly low – in the range of 30% to 35% over the full sample. In Tables 10 and 11 we report very different top-twenty and bottom-twenty factor model lists compared to Tables 5 and 6. As an example, the model using net stock issues (NSI) and liquidity (LIQ) is the top twenty performing factor models in our original full-sample analysis (Table 5), but it is one of the worst-performing models over the full sample under the double-sorting method (Table 11).

We can also compare overall factor model performance using the original one-dimensional sort factor construction (Figure 3 panel A) and the double-sort factor construction (Figure 4). While we observed in Table 9 a low correlation in model performance across the two factor construction methods, the relative predictability of characteristics is very similar.

Characteristics that were captured by a large proportion of factor models in Figure 3 are also captured by a significant number of models in Figure 4 – these range from the return on assets (ROA) characteristic at 35% to the organization capital (OK) characteristic at 62%. Similarly, investment-to-capital (IK) also appears to be spanned only by the highest-ranked models. Finally, the same list of characteristics remain the most difficult to span: CI, IG, VOL, NSI, AC, AG, MOM, and IA all remain at 5% or less.

Finally, we examine the improvement in model performance caused by moving from three to four factors in the pricing models. We repeat our analysis by considering the universe of 2,925 four-factor models, consisting of the market portfolio and three c-factors based on our list of 27 firm characteristics. We present the results for four-factor models in Appendix B.

The best-performing four-factor model in Table B.1 is able to match 67% of the 24 target cross-sections, only 7% higher than the best performing three-factor model in Table 5. Many of the twenty best-performing four-factor models add factors constructed on momentum (MOM), standardized unexpected earnings (SUE), investment over assets (IA), and asset growth (AG) to one of the top-performing three-factor models. All of these additions are based on characteristics that present the most challenge to the three-factor c-models, as we show in Figure 3. Adding such factors to the three-factor models produces a slight mechanical improvement in performance by excluding the corresponding cross-section from the set of test portfolios. Beyond that, the improvement is minimal: most challenging cross-sections have little correlation with each other or with other c-factors, and therefore it is not possible to capture many additional cross-sections by introducing a fourth c-factor.

4 Conclusion

The potential hazards of data-mining are well known. Our findings show just how difficult it is to judge the performance of empirically constructed factor pricing models when both the return factors and the target cross-sections of assets are chosen in a virtually unrestricted manner. Starting with a set of 27 commonly used firm characteristics, we show that randomly constructed characteristic-based factor models can match as many as 60% of the target return cross-sections over the 1971-2011 sample period.

While the impressive performance of some of the models we consider is spurious, some models must indeed capture economically meaningful sources of risk. Distinguishing one set from the other purely based on empirical performance seems difficult – if the factors included in a theoretically grounded risk-factor model are some of the many possible c-factors, such a model is likely to be defeated in a pure performance horse-race by the spuriously picked champions. The winner in such a horse-race is not necessarily a superior risk model. For example, the momentum factor (MOM) appears in at least one of the three best-performing three-factor models for the full sample, and each of the half-samples. Yet, without a convincing attribution of the return spread on the momentum-sorted portfolios to a well-understood source of risk, it is difficult to interpret momentum as a primitive risk factor of first-order economic importance.

Other situations may be more ambiguous, and one may be able to offer at least a tentative ex-post theoretical justification for the top-performing model. Such theory-mining can add a patina of false legitimacy to the spurious pricing models, exacerbating the effects of data-mining. For example, the top-performing model based on the standardized-unexpected-earnings (SUE) and the cashflow-to-price (CP) factors suggests some tantalizing possibilities for straddling the behavioral and neoclassical asset pricing paradigms to "motivate" a hybrid pricing model with empirical performance that is literally second to none. Needless to say, a superficial theory adds no more value than a spurious empirical result.

In summary, our analysis lends further support to the notion that to distinguish meaningful pricing models from the spurious ones, we should place less weight on the number of seemingly anomalous return cross-sections the models are able to match, and instead closely scrutinize the theoretical plausibility and empirical evidence in favor or against their main economic mechanisms.

References

Ang, A., R. Hodrick, Y. Xing, and X. Zhang (2006). The cross-section of volatility and expected returns. *Journal of Finance 61 (1)*, 259–299.

Baker, M., B. Bradley, and J. Wurgler (2011). Benchmarks as limits to arbitrage: understanding the low-volatility anomaly. *Financial Analysts Journal 67*, 40–54.

Ball, R. and P. Brown (1968). An empirical evaluation of accounting income numbers. *Journal of Accounting Research 6 (2)*, 159–178.

Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics 9*, 3–18.

Basu, S. (1977). Investment performance of common stocks in relation to their price-earnings ratios: a test of the efficient market hypothesis. *Journal of Finance 32 (3)*, 663–682.

Basu, S. (1983). The relationship between earnings' yield, market value and return for nyse common stocks. *Journal of Financial Economics 12*, 129–156.

Bernard, V. L. and J. K. Thomas (1989). Post-earnings-announcement drift: delayed price response or risk premium? *Journal of Accounting Research 27*, 1–36.

Bhandari, L. C. (1988). Debt/equity ratio and expected common stock returns: empirical evidence. *Journal of Finance 43 (2)*, 507–528.

Black, F. (1993). Beta and return. *Journal of Political Economy 20*, 8–18.

Black, F., M. C. Jensen, and M. S. Scholes (1972). *Studies in the Theory of Capital Markets*, Chapter The Capital Asset Pricing Model: Some Empirical Tests. Praeger Publishers Inc.

Chan, L. K. C., Y. Hamao, and J. Lakonishok (1991). Fundamentals and stock returns in japan. *Jouranl of Finance 46 (5)*, 1739–1764.

Chan, L. K. C., N. Jegadeesh, and J. Lakonishok (1996). Momentum strategies. *Journal of Finance 51 (5)*, 1681–1713.

Chen, L., R. Novy-Marx, and L. Zhang (2010). An alternative three-factor model. working paper.

Cohen, R. B., P. A. Gompers, and T. Vuolteenaho (2002). Who underreacts to cash-flow news? evidence from trading between individuals and institutions. *Journal of Financial Economics 66*, 409–462.

Cooper, M. J., H. Gulen, and M. J. Schill (2008). Asset growth and the cross-section of stock returns. *Journal of Finance 63 (4)*, 1609–1651.

Daniel, K. and S. Titman (2006). Market reactions to tangible and intangible information. *Journal of Finance 61 (4)*, 1605–1643.

DeBondt, W. F. and R. Thaler (1985). Does the stock market overreact? *Journal of Finance 40 (3)*, 793–805.

Eisfeldt, A. and D. Papanikolaou (2012). Organization capital and the cross-section of expected returns. *Journal of Finance*. working paper.

Fairfield, P. M., J. S. Whisenant, and T. L. Yohn (2003). Accrued earnings and growth: implications for future profitability and market mispricing. *The Accounting Review 78 (1)*, 353–371.

Fama, E. F. and K. French (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics 33 (1)*, 3–56.

Fama, E. F. and K. R. French (1992). The cross-section of expected stock returns. *Journal of Finance 47 (2)*, 427–465.

Fama, E. F. and K. R. French (1996). Multifactor explanations of asset pricing anomalies. *Journal of Finance 51 (1)*, 55–84.

Fama, E. F. and K. R. French (2006). Profitability, investment and average returns. *Journal of Financial Economics 82*, 491–518.

Fama, E. F. and K. R. French (2008). Dissecting anomalies. *Journal of Finance 63 (4)*, 1653–1678.

Ferson, W. (1996). Warning: Attribute-sorted portfolios can be hazardous to your research. In S. Saitou, K. Sawaki, and K. Kubota (Eds.), *Modern Finance Theory and its Applications*, pp. 21–32. Center for Academic Societies, Osaka, Japan.

Frazzini, A. and L. Pedersen (2011, October). Betting against beta. working paper.

Gibbons, M. R., S. A. Ross, and J. Shanken (1989). A test of the efficiency of a given portfolio. *Econometrica 57 (5)*, 1121–1152.

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

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

Ikenberry, D., J. Lakonishok, and T. Vermaelen (1995). Market underreaction to open market share repurchases. *Journal of Financial Economics 39*, 181–208.

Jegadeesh, N. and S. Titman (1993). Returns to buying winners and selling losers: implications for stock market efficiency. *Journal of Finance 48 (1)*, 65–91.

Lakonishok, J., A. Shleifer, and R. W. Vishny (1994). Contrarian investment, extrapolation, and risk. *Journal of Finance 49 (5)*, 1541–1578.

Lee, C. M. and B. Swaminathan (2000). Price momentum and trading volume. *Journal of Finance 55 (5)*, 2017–2069.

Lewellen, J. (2011). The cross-section of expected stock returns. working paper.

Lewellen, J., S. Nagel, and J. Shanken (2006, July). A skeptical appraisal of asset-pricing tests. NBER Working Papers 12360, National Bureau of Economic Research, Inc.

Litzenberger, R. H. and K. Ramaswamy (1982). The effects of dividends on common stock prices: tax effects or information effects? *Journal of Finance 37 (2)*, 429–443.

Lo, A. W. and A. C. MacKinlay (1990). Data-snooping biases in tests of financial asset pricing models. *Review of Financial Studies 3 (3)*, 431–467.

Loughran, T. and J. R. Ritter (1995). The new issues puzzle. *Journal of Finance 50 (1)*, 23–51.

Lyandres, E., L. Sun, and L. Zhang (2008). The new issues puzzle: testing the investment-based explanation. *Review of Financial Studies 21 (6)*, 2825–2855.

Miller, M. H. and M. S. Scholes (1982). Dividends and taxes: some empirical evidence. *Journal of Political Economy 90 (6)*, 1118–1141.

Novy-Marx, R. (2012, May). Pseudo-predictability in conditional asset pricing tests: explaining anomaly performance with politics, the weather, global warming, sunspots, and the stars. working paper.

Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research 18 (1)*, 109–131.

Pastor, L. and R. F. Stambaugh (2003). Liquidity risk and expected stock returns. *Journal of Political Economy 111 (3)*, 642–685.

Piotroski, J. (2000). Value investing: the use of historical financial statement information to separate winners from losers. *Journal of Accounting Research 38*, 1–51.

Pontiff, J. and A. Woodgate (2008). Share issuance and cross-sectional returns. *Journal of Finance 63 (2)*, 921–945.

Rosenberg, B., K. Reid, and R. Lanstein (1985). Persuasive evidence of market inefficiency. *Journal of Portfolio Management 11*, 9–17.

Sloan, R. G. (1996). Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review 71 (3)*, 289–315.

Titman, S., K. J. Wei, and F. Xie (2004). Capital investments and stock returns. *Journal of Financial and Quantitative Analysis 39 (4)*, 677–700.

Xing, Y. (2008). Interpreting the value effect through the q-theory: an empirical investigation. *Review of Financial Studies 21 (4)*, 1767–1795.

Tables and Figures

Table 1 contains the monthly value-weighted average returns, CAPM alphas, and Fama-French alphas for the 27 characteristic-based return factors, over the whole and subsample periods. Factors are the high minus low portfolio from sorting firms into ten portfolios with respect to the underlying firm characteristic. Abbreviations are as follows: size (SIZE), book-to-market (BM), dividend-to-price (DP), earnings-to-price (EP), cash flow-to-price (CP), investment-to-assets (IA), asset growth (AG), accruals (AC), abnormal investment (AI), net operating assets (NOA), investment-to-capital (IK), investment growth (IG), momentum (MOM), long-term reversal (LTR), return on assets (ROA), standardized unexpected earnings (SUE), return on equity (ROE), sales growth (SG), Ohlson score (OS), market leverage (LEV), net stock issues (NSI), composite issuance (CI), organization capital (OK), liquidity risk (LIQ), turnover (TO), idiosyncratic volatility (VOL), and market beta (BETA).

Table 1: Characteristics Factors: Summary Statistics

Panel A: Average Returns

	1971-2011		1971-1991		1992-2011	
Characteristic	ret	t stat	ret	t stat	ret	t stat
SIZE	-0.004	-1.68	-0.003	-0.86	-0.006	-1.48
BM	0.004	2.18	0.005	1.53	0.004	1.60
DP	0.002	0.87	0.001	0.33	0.003	0.91
EP	0.007	2.71	0.010	2.95	0.004	1.00
CP	0.009	3.49	0.014	4.53	0.004	0.95
IA	-0.007	-4.37	-0.008	-3.26	-0.007	-2.91
AG	-0.007	-3.15	-0.006	-2.19	-0.007	-2.27
AC	-0.006	-2.75	-0.004	-1.40	-0.008	-2.43
AI	-0.001	-0.84	0.001	0.30	-0.003	-1.40
NOA	0.001	0.29	0.003	0.96	-0.002	-0.35
IK	-0.003	-1.11	-0.005	-1.57	-0.002	-0.35
IG	-0.004	-2.33	-0.002	-1.12	-0.006	-2.05
MOM	0.014	4.08	0.018	5.06	0.009	1.62
LTR	-0.003	-1.27	-0.001	-0.44	-0.005	-1.35
ROA	0.008	2.59	0.007	2.16	0.009	1.69
SUE	0.006	2.96	0.007	3.18	0.006	1.59
ROE	0.004	1.14	0.006	1.64	0.001	0.27
SG	-0.002	-1.14	-0.003	-1.04	-0.002	-0.58
OS	-0.007	-2.01	-0.007	-1.71	-0.007	-1.22
LEV	0.004	1.91	0.003	0.97	0.005	1.78
NSI	-0.005	-3.70	-0.005	-3.07	-0.005	-2.32
CI	-0.005	-2.57	-0.004	-1.54	-0.006	-2.07
OK	-0.001	-0.20	-0.001	-0.20	-0.001	-0.11
LIQ	0.004	1.78	0.003	1.10	0.004	1.41
TO	0.000	-0.07	-0.001	-0.15	0.000	0.03
VOL	-0.014	-3.31	-0.019	-4.69	-0.008	-1.12
BETA	0.000	-0.06	0.000	-0.02	0.000	-0.06

Panel B: CAPM alpha						
	1971-2011		1971-1991		1992-2011	
Characteristic	alpha	t stat	alpha	t stat	alpha	t stat
SIZE	-0.004	-1.44	-0.003	-0.82	-0.005	-1.17
BM	0.005	2.48	0.006	1.91	0.004	1.50
DP	0.004	2.18	0.003	1.37	0.005	1.69
EP	0.009	3.43	0.012	3.19	0.006	1.66
CP	0.011	3.98	0.015	4.41	0.006	1.53
IA	-0.008	-4.19	-0.008	-3.06	-0.007	-2.84
AG	-0.007	-3.16	-0.007	-2.21	-0.008	-2.28
AC	-0.006	-3.15	-0.005	-2.24	-0.007	-2.30
AI	-0.001	-0.67	0.001	0.42	-0.003	-1.52
NOA	0.003	1.01	0.005	1.29	0.001	0.27
IK	-0.006	-2.12	-0.006	-2.68	-0.006	-1.14
IG	-0.005	-2.38	-0.002	-1.15	-0.007	-2.17
MOM	0.015	4.69	0.018	4.87	0.012	2.43
LTR	-0.003	-1.19	-0.002	-0.48	-0.005	-1.10
ROA	0.010	3.25	0.008	2.24	0.013	2.66
SUE	0.007	2.23	0.007	2.95	0.008	1.33
ROE	0.006	1.73	0.006	1.64	0.006	1.17
SG	-0.004	-1.97	-0.004	-1.44	-0.004	-1.37
OS	-0.009	-2.63	-0.009	-1.83	-0.011	-2.00
LEV	0.005	1.99	0.004	1.06	0.006	1.66
NSI	-0.006	-3.61	-0.005	-3.00	-0.007	-2.65
CI	-0.007	-3.99	-0.005	-2.36	-0.009	-3.29
OK	-0.003	-0.92	-0.002	-0.60	-0.004	-0.76
LIQ	0.004	1.66	0.004	1.34	0.003	0.97
TO	-0.004	-1.51	-0.004	-1.25	-0.004	-1.01
VOL	-0.018	-4.58	-0.021	-5.12	-0.015	-2.41
BETA	-0.006	-1.78	-0.004	-1.08	-0.009	-1.70

Panel C: Fama-French alpha						
	1971-2011		1971-1991		1992-2011	
Characteristic	alpha	t stat	alpha	t stat	alpha	t stat
SIZE	-0.001	-0.64	-0.001	-0.21	-0.003	-0.74
BM	0.000	-0.11	-0.002	-0.97	0.001	0.36
DP	0.000	0.10	-0.002	-1.25	0.002	0.80
EP	0.009	3.99	0.014	4.35	0.005	1.72
CP	0.008	3.19	0.012	3.14	0.005	1.68
IA	-0.005	-3.07	-0.004	-2.07	-0.006	-2.27
AG	-0.003	-1.72	-0.002	-0.77	-0.005	-1.57
AC	-0.006	-2.98	-0.005	-2.10	-0.008	-2.18
AI	0.000	-0.28	0.002	0.67	-0.002	-1.15
NOA	0.001	0.49	0.004	1.54	-0.002	-0.45
IK	0.000	-0.22	-0.002	-0.98	0.000	0.12
IG	-0.002	-1.28	0.000	-0.19	-0.004	-1.58
MOM	0.017	5.74	0.020	5.87	0.014	2.80
LTR	0.002	1.12	0.005	1.79	0.000	0.16
ROA	0.012	3.85	0.013	4.51	0.013	3.02
SUE	0.008	2.67	0.009	4.58	0.007	1.36
ROE	0.007	2.25	0.012	4.27	0.005	1.04
SG	-0.001	-0.42	0.000	-0.15	-0.001	-0.72
OS	-0.010	-3.84	-0.013	-4.24	-0.009	-2.56
LEV	-0.001	-0.85	-0.004	-1.69	0.001	0.37
NSI	-0.005	-3.08	-0.004	-2.23	-0.006	-2.66
CI	-0.005	-3.48	-0.004	-2.00	-0.007	-2.89
OK	0.000	-0.16	-0.001	-0.38	0.000	-0.17
LIQ	0.003	1.40	0.002	0.74	0.003	1.08
TO	0.000	0.14	-0.001	-0.40	0.001	0.33
VOL	-0.018	-6.09	-0.023	-7.68	-0.014	-3.32
BETA	-0.002	-0.79	-0.001	-0.22	-0.005	-1.27

Table 2: Variation Explained: Principal-Component Analysis of Return Factors

PC	1971-2011	1971-1991	1992-2011
1	0.41	0.31	0.51
2	0.57	0.52	0.62
3	0.63	0.60	0.69
4	0.68	0.66	0.74
5	0.72	0.71	0.77
6	0.75	0.75	0.80
7	0.78	0.78	0.83
8	0.80	0.80	0.85
9	0.83	0.83	0.86
10	0.84	0.85	0.88
11	0.86	0.87	0.89
12	0.88	0.88	0.91
13	0.89	0.90	0.92
14	0.90	0.91	0.93
15	0.92	0.92	0.94
16	0.93	0.93	0.95
17	0.94	0.94	0.96
18	0.95	0.95	0.96
19	0.95	0.96	0.97
20	0.96	0.96	0.98
21	0.97	0.97	0.98
22	0.98	0.98	0.99
23	0.98	0.98	0.99
24	0.99	0.99	0.99
25	0.99	0.99	0.99
26	1.00	1.00	1.00
27	1	1	1

Table 2 presents results from a principal component analysis on the 27 characteristic-based return factors. Factors are the high minus low portfolio from sorting firms into ten portfolios with respect to the underlying firm characteristic. The table shows the proportion of cumulative variation that the first n principal components can capture. Results are presented over the whole sample period 1971-2011 and subsamples 1971-1991 and 1992-2011.

Table 3: Principal-Component Factor Loadings

	1971-2011			1971-1991			1992-2011		
	PC1	PC2	PC3	PC1	PC2	PC3	PC1	PC2	PC3
SIZE	0.16	-0.27	0.16	0.30	-0.07	0.14	0.12	0.31	0.09
BM	0.03	0.27	0.07	-0.14	0.29	0.33	0.05	-0.17	0.06
DP	0.12	0.24	0.05	0.01	0.34	0.16	0.12	-0.19	0.03
EP	0.21	-0.10	-0.04	0.25	0.06	0.25	0.19	0.13	-0.03
CP	0.18	0.03	0.04	0.02	0.16	0.45	0.20	0.14	0.08
IA	0.00	-0.18	0.08	0.08	-0.16	0.03	-0.01	0.18	0.08
AG	0.00	-0.29	0.08	0.16	-0.20	0.08	-0.02	0.32	0.05
AC	-0.01	-0.07	0.02	-0.09	-0.14	0.15	0.03	0.07	-0.04
AI	0.04	-0.07	0.13	0.08	0.00	-0.02	0.03	0.09	0.13
NOA	0.23	0.02	0.25	0.26	0.15	0.00	0.21	-0.07	0.24
IK	-0.24	-0.26	0.02	-0.04	-0.28	-0.10	-0.26	0.27	-0.02
IG	-0.04	-0.15	0.10	0.04	-0.08	0.00	-0.05	0.22	0.09
MOM	0.08	-0.13	-0.84	0.12	-0.09	0.15	0.08	0.11	-0.86
LTR	0.06	-0.33	-0.02	0.24	-0.18	0.00	0.03	0.38	0.04
ROA	0.29	-0.20	-0.03	0.29	-0.09	0.11	0.30	0.19	-0.01
SUE	0.06	-0.08	-0.16	0.09	-0.06	0.04	0.06	0.04	-0.14
ROE	0.27	-0.21	0.03	0.31	-0.11	0.03	0.28	0.17	0.01
SG	-0.13	-0.16	-0.01	0.01	-0.22	0.04	-0.15	0.10	-0.07
OS	-0.34	0.22	-0.19	-0.44	-0.01	-0.08	-0.31	-0.21	-0.18
LEV	0.04	0.31	0.10	-0.18	0.26	0.33	0.07	-0.26	0.10
NSI	-0.08	-0.05	0.00	0.00	-0.06	-0.11	-0.10	0.03	-0.03
CI	-0.14	-0.09	0.05	-0.09	-0.18	-0.17	-0.13	0.07	0.06
OK	-0.25	-0.05	-0.15	-0.21	-0.16	-0.11	-0.25	0.06	-0.19
LIQ	-0.02	0.00	0.08	-0.01	0.11	-0.24	-0.05	0.11	0.06
TO	-0.25	-0.24	0.02	-0.13	-0.33	0.27	-0.25	0.25	0.00
VOL	-0.42	0.13	0.07	-0.37	-0.11	0.15	-0.41	-0.18	0.11
BETA	-0.37	-0.30	0.19	-0.14	-0.43	0.43	-0.39	0.25	0.17

Table 3 presents factor loadings for the first three principal components extracted from the set of 27 factor returns. Loadings are shown for the whole sample period 1971-2011 and subsamples 1971-1991 and 1992-2011.

Characteristic abbreviations are as follows: size (SIZE), book-to-market (BM), dividend-to-price (DP), earnings-to-price (EP), cash flow-to-price (CP), investment-to-assets (IA), asset growth (AG), accruals (AC), abnormal investment (AI), net operating assets (NOA), investment-to-capital (IK), investment growth (IG), momentum (MOM), long-term reversal (LTR), return on assets (ROA), standardized unexpected earnings (SUE), return on equity (ROE), sales growth (SG), Ohlson score (OS), market leverage (LEV), net stock issues (NSI), composite issuance (CI), organization capital (OK), liquidity risk (LIQ), turnover (TO), idiosyncratic return volatility (VOL), and market beta (BETA). Details on characteristic definitions and construction is in Appendix A.

Table 4: Factor Regression on the Principal-Component Model

	1971-2011			1971-1991			1992-2011		
factor	alpha	t stat	R^2	alpha	t stat	R^2	alpha	t stat	R^2
SIZE	-0.008	-4.20	0.59	-0.011	-4.16	0.69	-0.008	-2.73	0.59
BM	0.002	1.19	0.52	0.002	0.90	0.65	0.001	0.40	0.40
DP	0.000	-0.15	0.55	-0.003	-1.20	0.63	0.000	0.16	0.51
EP	0.004	2.00	0.55	0.006	2.04	0.46	0.002	0.71	0.62
CP	0.004	1.85	0.39	0.010	2.91	0.15	0.001	0.31	0.64
IA	-0.007	-4.30	0.33	-0.007	-3.33	0.38	-0.006	-2.63	0.28
AG	-0.005	-3.09	0.54	-0.006	-2.98	0.56	-0.005	-1.97	0.52
AC	-0.006	-2.76	0.03	-0.002	-0.73	0.25	-0.007	-2.06	0.04
AI	-0.002	-1.17	0.10	0.000	-0.19	0.12	-0.004	-1.95	0.11
NOA	-0.004	-2.15	0.57	-0.004	-1.85	0.67	-0.006	-2.13	0.57
IK	0.003	1.86	0.77	0.000	-0.14	0.56	0.004	2.01	0.85
IG	-0.002	-1.41	0.24	-0.001	-0.31	0.14	-0.004	-1.77	0.35
MOM	0.014	3.64	0.10	0.018	4.88	0.11	0.013	2.16	0.10
LTR	-0.002	-1.29	0.59	-0.003	-1.08	0.58	-0.002	-0.89	0.66
ROA	0.003	1.84	0.76	0.004	1.85	0.63	0.006	2.44	0.82
SUE	0.006	2.12	0.10	0.006	2.85	0.19	0.007	1.25	0.07
ROE	-0.001	-0.51	0.75	0.002	1.05	0.72	-0.001	-0.35	0.77
SG	0.001	0.83	0.51	0.000	-0.04	0.36	0.002	1.00	0.65
OS	0.000	-0.07	0.81	0.002	1.07	0.84	-0.002	-0.70	0.81
LEV	0.000	0.30	0.62	0.001	0.29	0.63	0.001	0.38	0.59
NSI	-0.003	-2.16	0.29	-0.004	-2.13	0.08	-0.004	-1.76	0.44
CI	-0.003	-1.83	0.44	0.000	-0.10	0.39	-0.005	-2.08	0.51
OK	0.006	2.97	0.65	0.006	2.52	0.57	0.005	1.73	0.70
LIQ	0.005	2.04	0.01	0.003	0.84	0.08	0.005	1.46	0.12
TO	0.004	2.50	0.77	0.004	1.64	0.66	0.005	2.46	0.85
VOL	-0.007	-3.11	0.82	-0.011	-4.31	0.73	-0.004	-1.30	0.85
BETA	0.004	1.80	0.78	0.005	1.59	0.64	0.002	0.51	0.85

Table 4 presents results from regressing the characteristic-based return factors on the benchmark three-factor model, consisting of the market portfolio and the first two principal component vectors of the return factors. Factors are the high minus low portfolio from sorting firms into ten portfolios with respect to the underlying firm characteristic. The alpha coefficient, t-statistic, and R^2 from the regression is shown in the table for the whole sample period 1971-2011 and subsamples 1971-1991 and 1992-2011.

Characteristic abbreviations are as follows: size (SIZE), book-to-market (BM), dividend-to-price (DP), earnings-to-price (EP), cash flow-to-price (CP), investment-to-assets (IA), asset growth (AG), accruals (AC), abnormal investment (AI), net operating assets (NOA), investment-to-capital (IK), investment growth (IG), momentum (MOM), long-term reversal (LTR), return on assets (ROA), standardized unexpected earnings (SUE), return on equity (ROE), sales growth (SG), Ohlson score (OS), market leverage (LEV), net stock issues (NSI), composite issuance (CI), organization capital (OK), liquidity risk (LIQ), turnover (TO), idiosyncratic return volatility (VOL), and market beta (BETA). Details on characteristic definitions and construction is in Appendix A.

Table 5: Top 20 Performing Factor Models

	1971-2011			1971-1991			1992-2011		
	C1	C2	prop	C1	C2	prop	C1	C2	prop
1	SUE	CP	0.60	MOM	CP	0.80	AG	EP	0.84
2	MOM	CP	0.56	MOM	IA	0.72	AG	CP	0.84
3	AG	CP	0.56	MOM	IK	0.72	MOM	NSI	0.80
4	AI	CP	0.56	IA	SUE	0.72	MOM	CI	0.80
5	CP	LIQ	0.56	IA	EP	0.72	ROA	AG	0.80
6	SIZE	VOL	0.52	OS	AG	0.72	AG	SUE	0.80
7	BM	MOM	0.52	IA	OS	0.68	AG	CI	0.80
8	BM	SUE	0.52	AC	CP	0.68	AG	VOL	0.80
9	BM	CP	0.52	AI	CP	0.68	SUE	CI	0.80
10	EP	IG	0.52	NOA	CP	0.68	CI	LIQ	0.80
11	ROE	CP	0.52	NOA	IK	0.68	CP	IG	0.80
12	NOA	CP	0.52	CP	IG	0.68	SIZE	VOL	0.76
13	CP	IG	0.52	CP	LIQ	0.68	MOM	SG	0.76
14	MOM	EP	0.48	MOM	AG	0.64	NSI	SUE	0.76
15	LTR	CP	0.48	IA	ROA	0.64	EP	IG	0.76
16	ROA	CP	0.48	ROA	CP	0.64	IG	VOL	0.76
17	OS	AG	0.48	DP	CP	0.64	BM	MOM	0.72
18	OS	CP	0.48	AG	SUE	0.64	BM	SUE	0.72
19	NSI	LIQ	0.48	AG	EP	0.64	MOM	DP	0.72
20	AG	EP	0.48	AC	IK	0.64	MOM	LEV	0.72

Table 5 lists the characteristic-based factors that constitute the top twenty linear factor models, in terms of the proportion of remaining characteristics they can capture, via the equal-weighted method. We say that a factor model M captures, or spans, a characteristic C, if the p-value from the Gibbons et al. (1989) F-test of joint significance of abnormal average return with respect to M across the ten sorted portfolios on C is above 10%. Top factor models are shown for the whole sample period 1971-2011 and subsamples 1971-1991 and 1992-2011.

The universe of factor models is all three-factor models consisting of the market portfolio and two characteristic return factors (C1, C2) from our list of 27. Characteristic abbreviations are as follows: size (SIZE), book-to-market (BM), dividend-to-price (DP), earnings-to-price (EP), cash flow-to-price (CP), investment-to-assets (IA), asset growth (AG), accruals (AC), abnormal investment (AI), net operating assets (NOA), investment-to-capital (IK), investment growth (IG), momentum (MOM), long-term reversal (LTR), return on assets (ROA), standardized unexpected earnings (SUE), return on equity (ROE), sales growth (SG), Ohlson score (OS), market leverage(LEV), net stock issues (NSI), composite issuance (CI), organization capital (OK), liquidity risk (LIQ), turnover (TO), idiosyncratic return volatility (VOL), and market beta (BETA). Details on characteristic definitions and construction is in Appendix A.

Table 6: Bottom 20 Performing Factor Models

	1971-2011			1971-1991			1992-2011		
	C1	C2	prop	C1	C2	prop	C1	C2	prop
1	ROE	SG	0.20	LEV	BETA	0.28	IA	LIQ	0.40
2	IK	VOL	0.20	CI	VOL	0.28	IA	TO	0.40
3	VOL	SG	0.20	EP	VOL	0.28	IA	SG	0.40
4	SIZE	BM	0.16	AC	OK	0.28	IA	BETA	0.40
5	SIZE	MOM	0.16	AI	VOL	0.28	LTR	DP	0.40
6	SIZE	IA	0.16	IK	VOL	0.28	DP	BETA	0.40
7	SIZE	LTR	0.16	VOL	SG	0.28	AC	AI	0.40
8	SIZE	AI	0.16	BM	CI	0.24	SMB	HML	0.37
9	SIZE	LIQ	0.16	LTR	AC	0.24	SIZE	LTR	0.36
10	IA	LTR	0.16	ROA	VOL	0.24	SIZE	DP	0.36
11	IA	AC	0.16	CI	OK	0.24	SIZE	AC	0.36
12	ROA	VOL	0.16	AC	VOL	0.24	SIZE	OK	0.36
13	NSI	VOL	0.16	OK	VOL	0.24	SIZE	LIQ	0.36
14	DP	VOL	0.16	LIQ	VOL	0.24	SIZE	BETA	0.36
15	CI	VOL	0.16	VOL	BETA	0.24	IA	AI	0.36
16	AC	OK	0.16	SIZE	ROA	0.20	LTR	IK	0.36
17	SIZE	LEV	0.12	SIZE	SUE	0.20	LTR	BETA	0.36
18	SIZE	AC	0.12	SIZE	ROE	0.20	SIZE	IA	0.32
19	IA	SG	0.12	DP	VOL	0.20	IA	DP	0.32
20	SIZE	SUE	0.08	VOL	TO	0.20	SIZE	AI	0.28

Table 6 lists the characteristic-based factors that constitute the bottom twenty linear factor models, in terms of the proportion of remaining characteristics they can capture, via the equal-weighted method. We say that a factor model M captures, or spans, a characteristic C, if the p-value from the Gibbons et al. (1989) F-test of joint significance of abnormal average return with respect to M across the ten sorted portfolios on C is above 10%. Bottom factor models are shown for the whole sample period 1971-2011 and subsamples 1971-1991 and 1992-2011.

The universe of factor models is all three-factor models consisting of the market portfolio and two characteristic return factors (C1, C2) from our list of 27. Characteristic abbreviations are as follows: size (SIZE), book-to-market (BM), dividend-to-price (DP), earnings-to-price (EP), cash flow-to-price (CP), investment-to-assets (IA), asset growth (AG), accruals (AC), abnormal investment (AI), net operating assets (NOA), investment-to-capital (IK), investment growth (IG), momentum (MOM), long-term reversal (LTR), return on assets (ROA), standardized unexpected earnings (SUE), return on equity (ROE), sales growth (SG), Ohlson score (OS), market leverage (LEV), net stock issues (NSI), composite issuance (CI), organization capital (OK), liquidity risk (LIQ), turnover (TO), idiosyncratic return volatility (VOL), and market beta (BETA). Details on characteristic definitions and construction is in Appendix A.

Table 7: Model Performance Correlation: First versus Second Half of the Sample

Method	Rank Corr	Corr
Equal-weighted	0.11	0.13
Characteristic Freq	0.12	0.16

Table 7 shows the rank correlation and correlation of factor model performance for the first subsample period (1971-1991) versus the second subsample period (1992-2011). The universe of factor models is all three-factor models consisting of the market portfolio and two characteristic return factors from our list of 27. The rank correlation is Spearman's rank correlation coefficient from the ranking of factor models, based on the percentage of characteristics matched. The correlation is the correlation coefficient of factor models' percentage of characteristics matched.

Correlations are shown for two characteristic weighting methods: equal-weighted method and characteristic matching frequency method. The "equal-weighted" method gives an equal weight to each characteristic matched. The "characteristic matching frequency" method gives each characteristic a weight of 1 minus the proportion of factor models that can match the cross-section of returns based on the characteristic under consideration.

Table 8: Model Performance Correlation: Characteristic Weighting Methods

Sample	Rank Corr	Corr
1971-2011	0.92	0.92
1971-1991	0.95	0.96
1992-2011	0.97	0.95

Table 8 shows the rank correlation and correlation of factor model performance across the two characteristic weighting methods used to compute the proportion of characteristics explained. The "equal-weighted" method gives an equal weight to each characteristic matched. The "characteristic matching frequency" method gives each characteristic a weight of 1 minus the proportion of factor models that can match the cross-section of returns based on the characteristic under consideration.
The universe of factor models is all three-factor models consisting of the market portfolio and two characteristic return factors from our list of 27. The rank correlation is Spearman's rank correlation coefficient from the ranking of factor models, based on the percentage of characteristics matched. Results are shown for the whole sample period 1971-2011 and subsamples 1971-1991 and 1992-2011.

Table 9: Model Performance Correlation: Factor Construction

	1971-2011		1971-1991		1992-2011	
Method	rank corr	corr	rank corr	corr	rank corr	corr
Equal-weighted	0.32	0.35	0.42	0.43	0.32	0.34
Characteristic Freq	0.33	0.37	0.40	0.42	0.25	0.19

Table 9 shows the rank correlation and correlation of factor model performance across the two different methods to construct characteristic-based return factors. The default method is to construct the factor as the high minus low portfolio of a one-way sort. The second method is to construct the factor as the equal-weighed average of the high minus low portfolio within the big and small size group, from a double-sort first on size and then the characteristic.

The universe of factor models is all three-factor models consisting of the market portfolio and two characteristic return factors from our list of 27. The rank correlation is Spearman's rank correlation coefficient from the ranking of factor models, based on the percentage of characteristics matched. The correlation is the correlation coefficient of factor models' percentage of characteristics matched.

Correlations are shown for two characteristic weighting methods, equal-weighted method and characteristic matching frequency method, as well as for the whole sample period 1971-2011 and subsamples 1971-1991 and 1992-2011. The "equal-weighted" method gives an equal weight to each characteristic matched. The "characteristic matching frequency" method gives each characteristic a weight of 1 minus the proportion of factor models that can match the cross-section of returns based on the characteristic under consideration.

Table 10: Top 20 Performing Factor Models - Double Sort

	1971-2011			1971-1991			1992-2011		
	C1	C2	prop	C1	C2	prop	C1	C2	prop
1	BM	VOL	0.48	LEV	CP	0.80	MOM	SUE	0.80
2	MOM	EP	0.48	CP	IG	0.80	DP	SUE	0.76
3	BM	LEV	0.44	BM	CI	0.76	EP	AC	0.76
4	OS	CP	0.44	BM	CP	0.76	ROA	AC	0.72
5	OS	TO	0.44	LTR	CP	0.76	AC	BETA	0.72
6	LEV	LIQ	0.44	NSI	TO	0.76	SUE	AI	0.68
7	AC	CP	0.44	AG	ROE	0.76	BM	MOM	0.64
8	AC	TO	0.44	ROE	CP	0.76	MOM	OS	0.64
9	AC	BETA	0.44	CP	SG	0.76	ROA	ROE	0.64
10	CP	VOL	0.44	BM	ROE	0.72	CI	AC	0.64
11	CP	BETA	0.44	MOM	CP	0.72	AC	ROE	0.64
12	LIQ	SG	0.44	IA	CP	0.72	AC	TO	0.64
13	BM	MOM	0.40	LTR	EP	0.72	BM	SUE	0.60
14	BM	EP	0.40	OS	IK	0.72	BM	AI	0.60
15	BM	LIQ	0.40	NSI	DP	0.72	MOM	EP	0.60
16	BM	TO	0.40	NSI	AG	0.72	MOM	BETA	0.60
17	MOM	CP	0.40	NSI	IK	0.72	ROA	SUE	0.60
18	MOM	VOL	0.40	NSI	VOL	0.72	ROA	AI	0.60
19	MOM	TO	0.40	AG	CP	0.72	OS	CP	0.60
20	MOM	BETA	0.40	ROE	IK	0.72	SUE	CP	0.60

Table 10 lists the characteristic-based factors that constitute the top twenty linear factor models, in terms of the proportion of remaining characteristics they can capture, via the equal-weighted method. We say that a factor model M captures, or spans, a characteristic C, if the p-value from the Gibbons et al. (1989) F-test of joint significance of abnormal average return with respect to M across the ten sorted portfolios on C is above 10%. Factors are constructed as the equal-weighed average of the high minus low portfolio within the big and small size group, from a double-sort first on size and then the characteristic. Top factor models are shown for the whole sample period 1971-2011 and subsamples 1971-1991 and 1992-2011.

The universe of factor models is all three-factor models consisting of the market portfolio and two characteristic return factors (C1, C2) from our list of 27. Characteristic abbreviations are as follows: size (SIZE), book-to-market (BM), dividend-to-price (DP), earnings-to-price (EP), cash flow-to-price (CP), investment-to-assets (IA), asset growth (AG), accruals (AC), abnormal investment (AI), net operating assets (NOA), investment-to-capital (IK), investment growth (IG), momentum (MOM), long-term reversal (LTR), return on assets (ROA), standardized unexpected earnings (SUE), return on equity (ROE), sales growth (SG), Ohlson score (OS), market leverage (LEV), net stock issues (NSI), composite issuance (CI), organization capital (OK), liquidity risk (LIQ), turnover (TO), idiosyncratic return volatility (VOL), and market beta (BETA). Details on characteristic definitions and construction is in Appendix A.

Table 11: Bottom 20 Performing Factor Models - Double Sort

	1971-2011			1971-1991			1992-2011		
	C1	C2	prop	C1	C2	prop	C1	C2	prop
1	IA	VOL	0.08	OS	SUE	0.20	SIZE	SUE	0.20
2	IA	SG	0.08	SUE	AC	0.20	SIZE	SG	0.20
3	ROA	TO	0.08	SUE	IG	0.20	MOM	IA	0.20
4	NSI	LEV	0.08	SUE	LIQ	0.20	IA	DP	0.20
5	NSI	AI	0.08	NOA	OK	0.20	IA	AI	0.20
6	NSI	LIQ	0.08	SIZE	MOM	0.16	IA	IG	0.20
7	NSI	TO	0.08	SIZE	LTR	0.16	SIZE	DP	0.16
8	AG	CI	0.08	ROA	SUE	0.16	SIZE	AG	0.16
9	SUE	ROE	0.08	ROA	AC	0.16	SIZE	AC	0.16
10	CI	OK	0.08	SUE	EP	0.16	SIZE	AI	0.16
11	AC	LIQ	0.08	SUE	ROE	0.16	SIZE	IG	0.16
12	SIZE	MOM	0.04	SUE	OK	0.16	SIZE	LIQ	0.16
13	SIZE	SUE	0.04	SUE	VOL	0.16	IA	AG	0.16
14	BM	IA	0.04	SUE	BETA	0.16	IA	CI	0.16
15	IA	DP	0.04	SIZE	ROE	0.12	IA	LIQ	0.16
16	ROA	IK	0.04	DP	SUE	0.12	AG	AC	0.16
17	NSI	SUE	0.04	SUE	CI	0.12	AG	LIQ	0.16
18	SUE	IK	0.04	SUE	NOA	0.12	SIZE	MOM	0.12
19	CI	AC	0.04	SIZE	ROA	0.08	SIZE	IA	0.12
20	ROA	NSI	0	SIZE	SUE	0.04	IA	AC	0.12

Table 11 lists the characteristic-based factors that constitute the bottom twenty linear factor models, in terms of the proportion of remaining characteristics they can capture, via the equal-weighted method. We say that a factor model M captures, or spans, a characteristic C, if the p-value from the Gibbons et al. (1989) F-test of joint significance of abnormal average return with respect to M across the ten sorted portfolios on C is above 10%. Factors are constructed as the equal-weighed average of the high minus low portfolio within the big and small size group, from a double-sort first on size and then the characteristic. Bottom factor models are shown for the whole sample period 1971-2011 and subsamples 1971-1991 and 1992-2011.

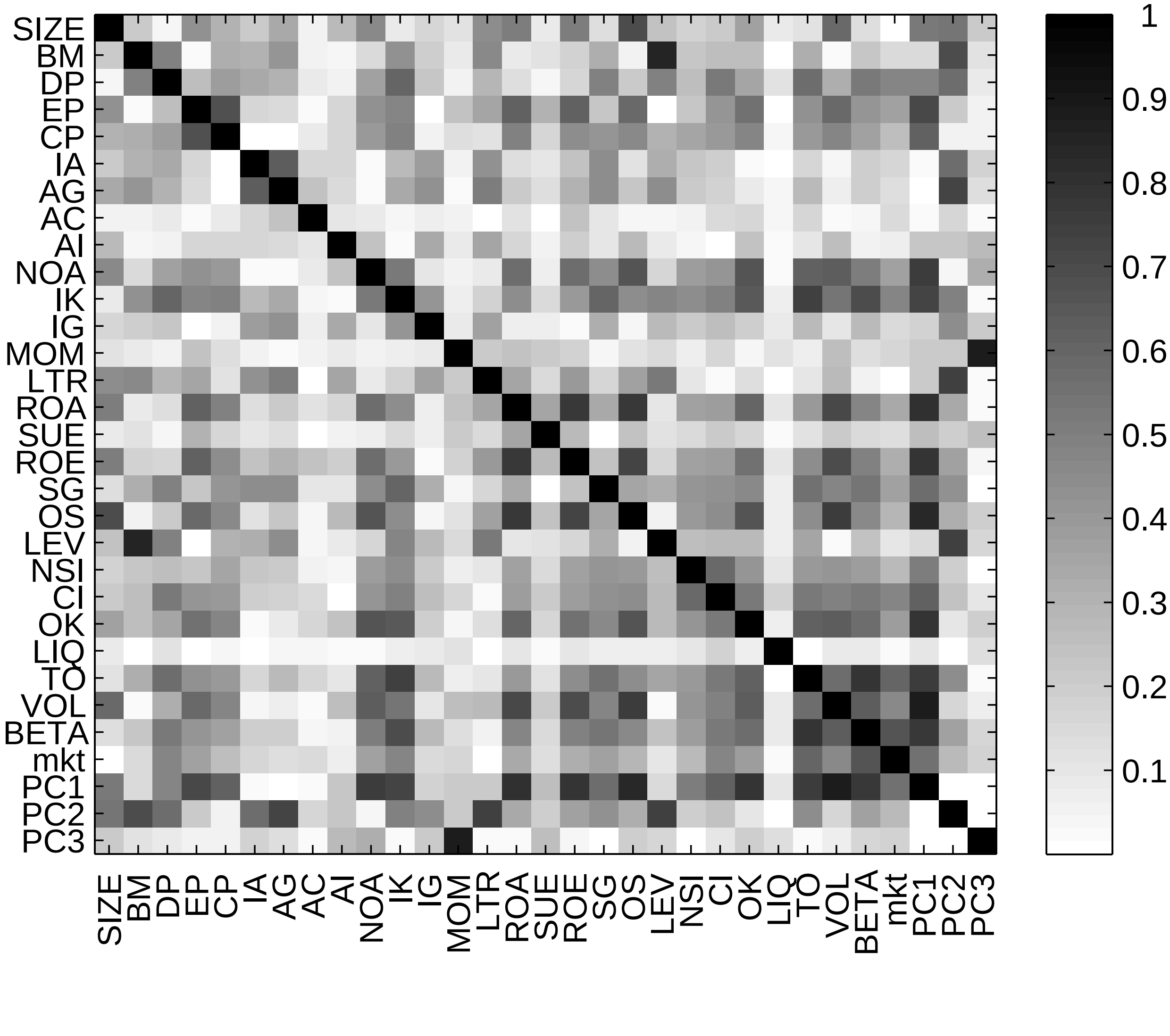
The universe of factor models is all three-factor models consisting of the market portfolio and two characteristic return factors (C1, C2) from our list of 27. Characteristic abbreviations are as follows: size (SIZE), book-to-market (BM), dividend-to-price (DP), earnings-to-price (EP), cash flow-to-price (CP), investment-to-assets (IA), asset growth (AG), accruals (AC), abnormal investment (AI), net operating assets (NOA), investment-to-capital (IK), investment growth (IG), momentum (MOM), long-term reversal (LTR), return on assets (ROA), standardized unexpected earnings (SUE), return on equity (ROE), sales growth (SG), Ohlson score (OS), market leverage (LEV), net stock issues (NSI), composite issuance (CI), organization capital (OK), liquidity risk (LIQ), turnover (TO), idiosyncratic return volatility (VOL), and market beta (BETA). Details on characteristic definitions and construction is in Appendix A.

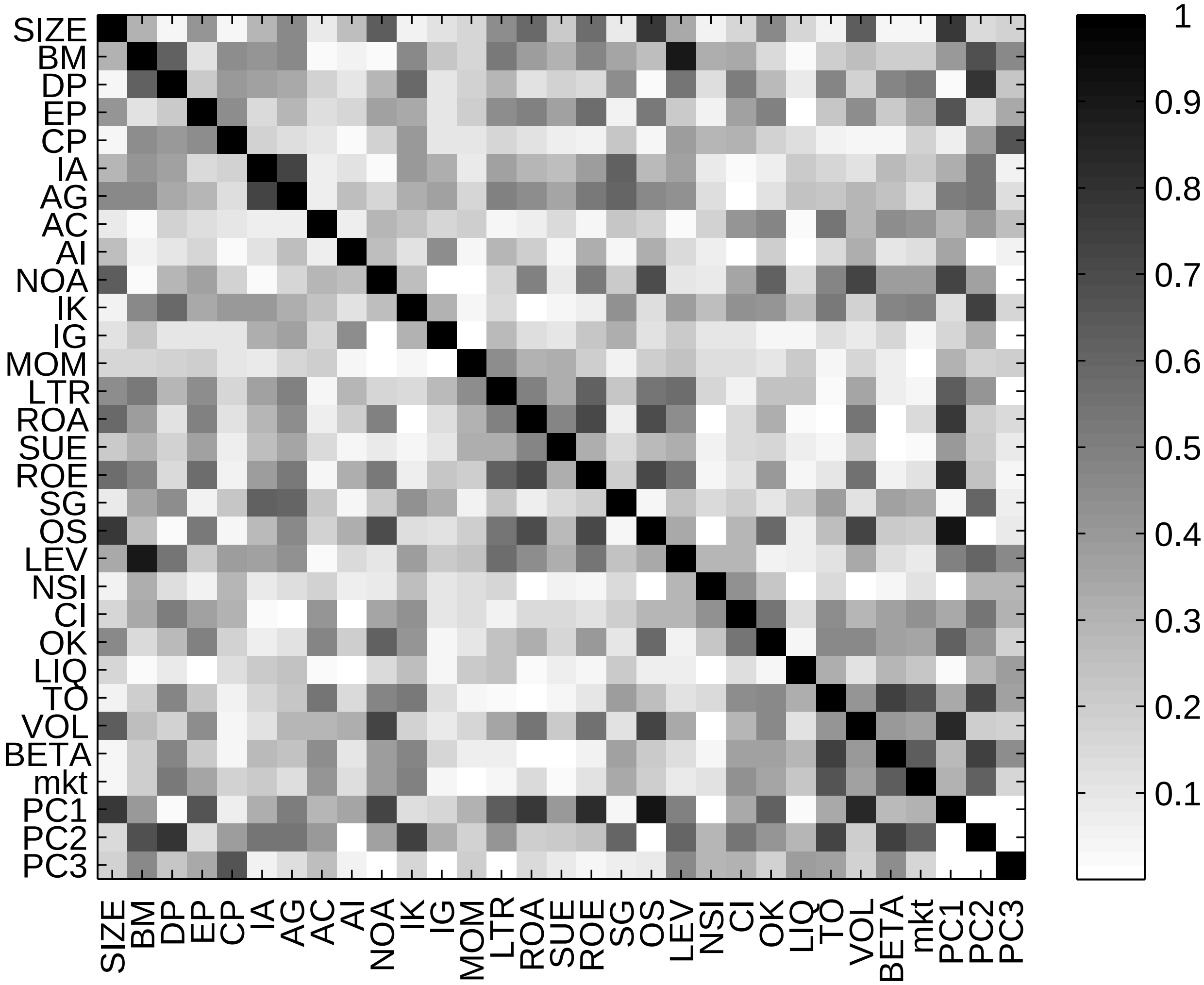
Figure 1: Factor Correlation

Figure 1 shows a heatmap representation of the correlation matrix for the 27 characteristic-based factors, the market portfolio, and the first three principal components extracted from the return factors. The magnitude of correlations is represented in the figure, with darker areas representing higher correlation.

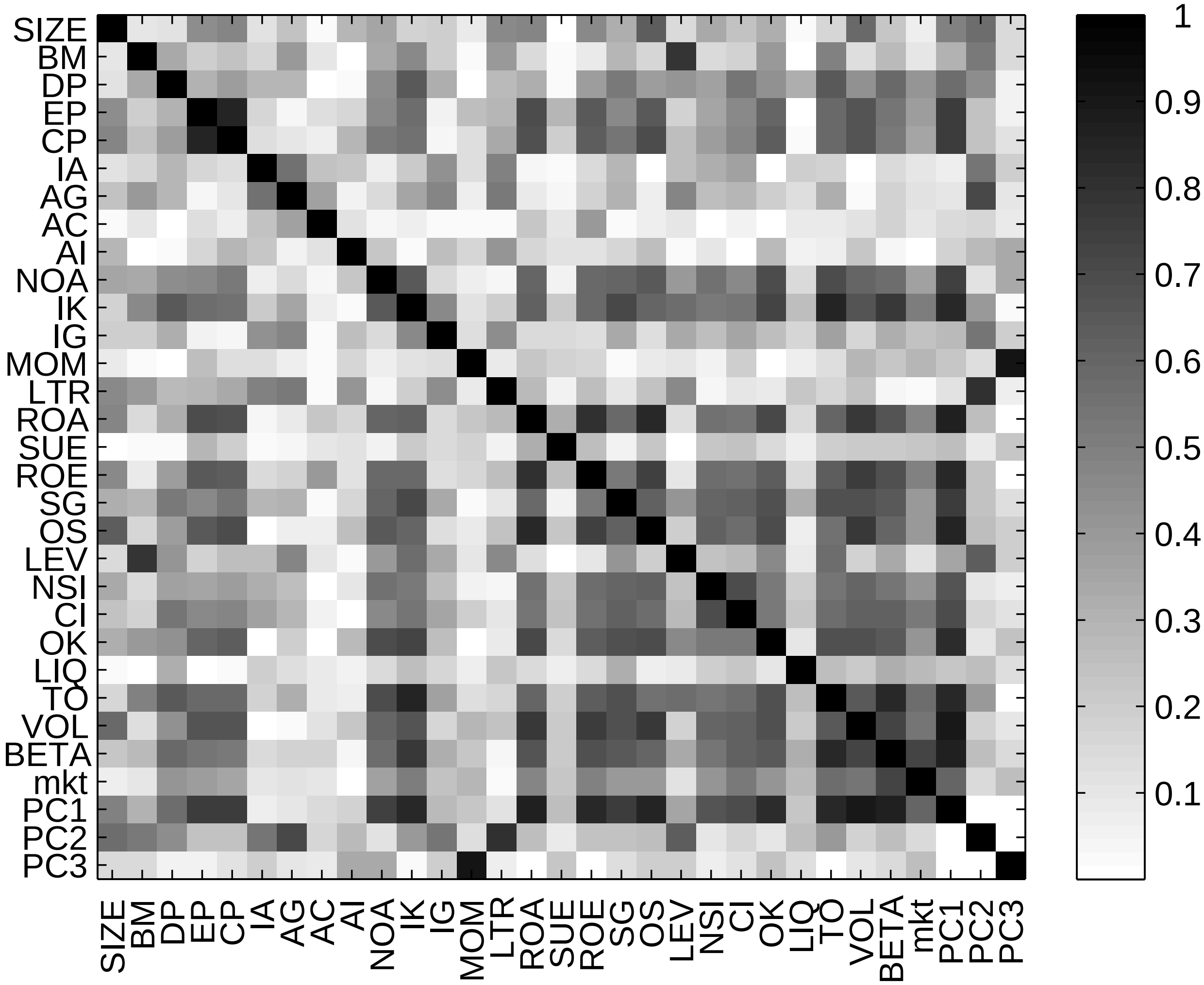
Factors are the high minus low portfolio from sorting firms into ten portfolios with respect to the underlying firm characteristic. Characteristic abbreviations are as follows: size (SIZE), book-to-market (BM), dividend-to-price (DP), earnings-to-price (EP), cash flow-to-price (CP), investment-to-assets (IA), asset growth (AG), accruals (AC), abnormal investment (AI), net operating assets (NOA), investment-to-capital (IK), investment growth (IG), momentum (MOM), long-term reversal (LTR), return on assets (ROA), standardized unexpected earnings (SUE), return on equity (ROE), sales growth (SG), Ohlson score (OS), market leverage(LEV), net stock issues (NSI), composite issuance (CI), organization capital (OK), liquidity risk (LIQ), turnover (TO), idiosyncratic return volatility (VOL), and market beta (BETA). Details on characteristic definitions and construction is in Appendix A.

(a) 1971-2011




SIZE
BM
DP
EP
CP
IA
AG
AC
AI
NOA
IK
IG
MOM
LTR
ROA
SUE
ROE
SG
OS
LEV
NSI
CI
OK
LIQ
TO
VOL
BETA
mkt
PC1
PC2
PC3
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1

(b) 1971-1991



SIZE
BM
DP
EP
CP
IA
AG
AC
AI
NOA
IK
IG
MOM
LTR
ROA
SUE
ROE
SG
OS
LEV
NSI
CI
OK
LIQ
TO
VOL
BETA
mkt
PC1
PC2
PC3
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1

(c) 1992-2011

Figure 2: Factor Model Performance

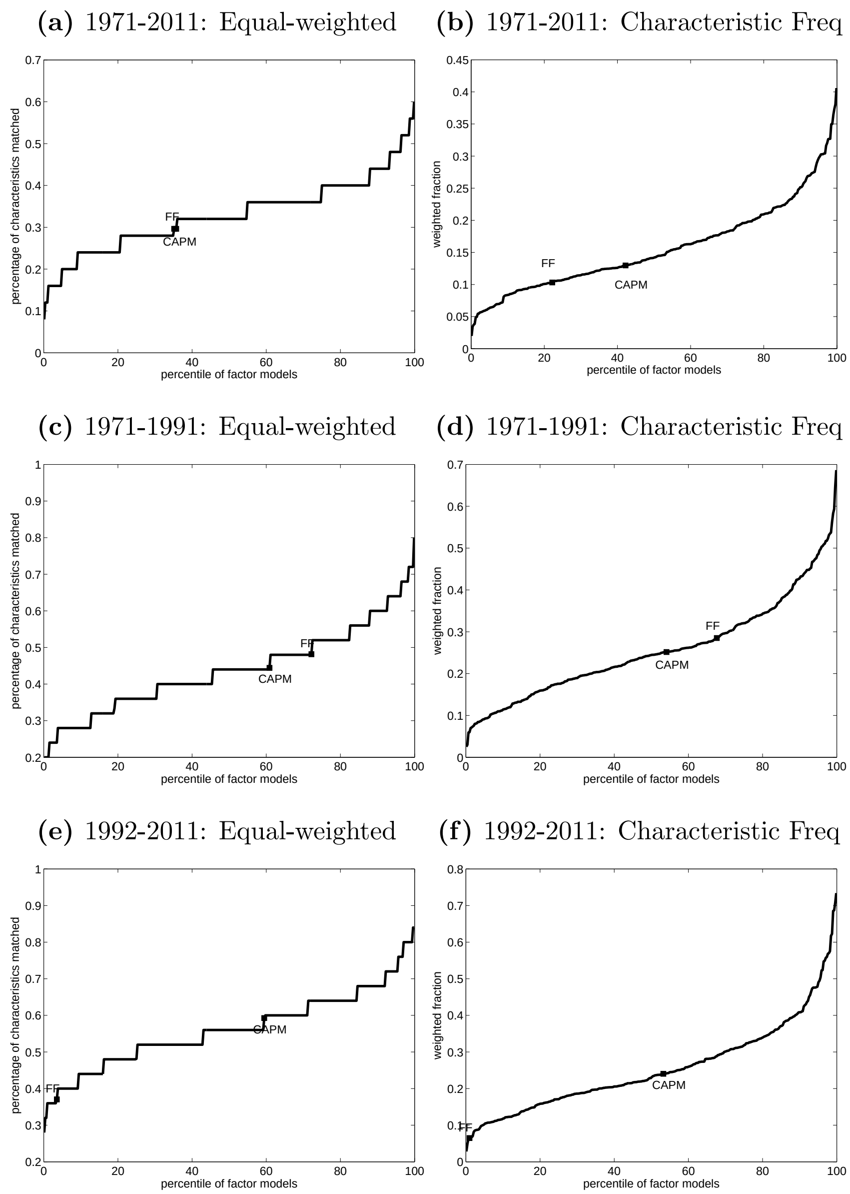


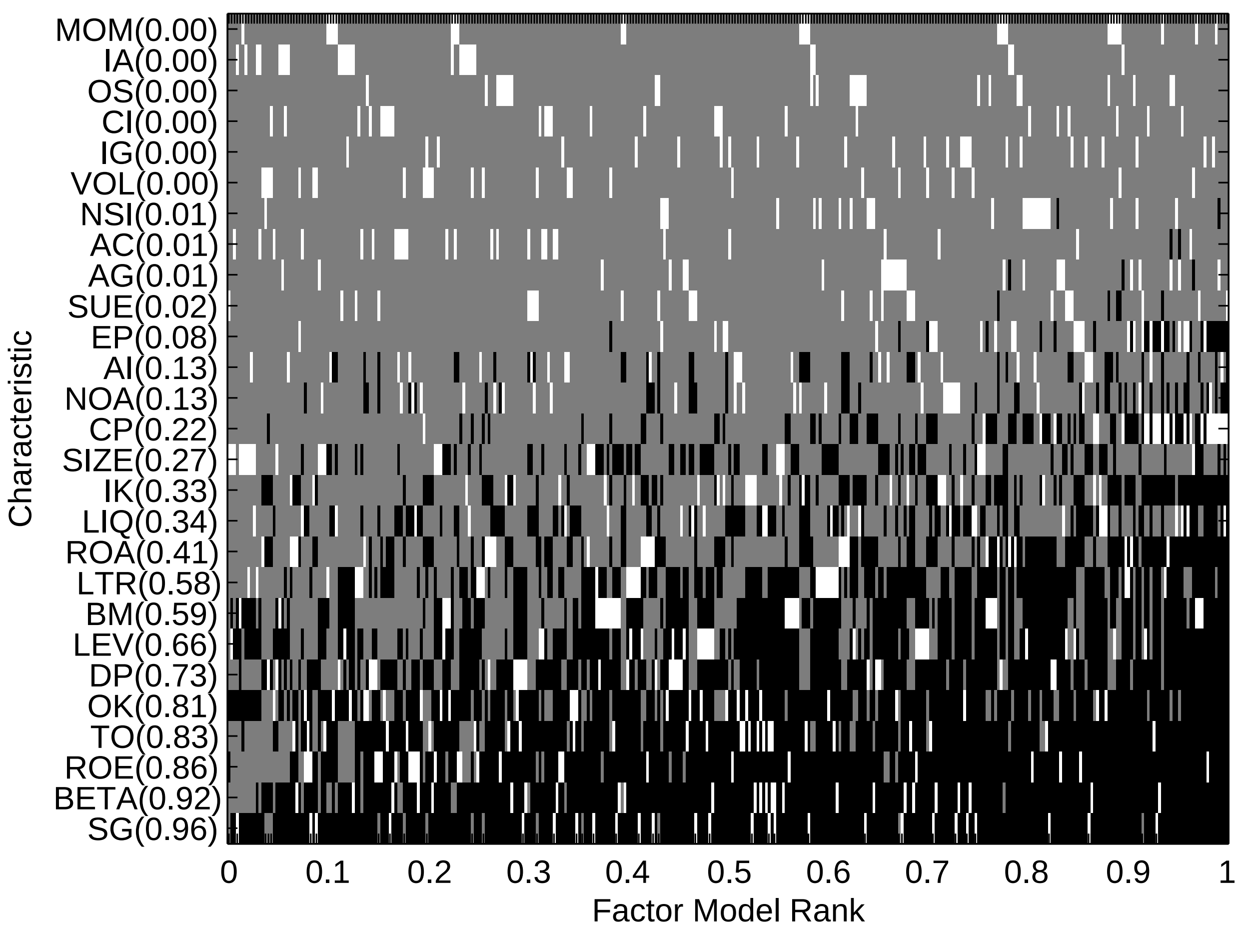
Figure 2 displays the distribution of factor model performance, as measured by the percentage of characteristics matched, over the whole sample period 1971-2011 and subsamples 1971-1991 and 1992-2011. The universe of factor models is all three-factor models consisting of the market portfolio and two characteristic return factors from our list of 27. The percentage of characteristics matched is computed using two characteristic weighting methods: equal-weighted method and characteristic matching frequency method. The "equal-weighted" method gives an equal weight to each characteristic matched. The "characteristic matching frequency" method gives each characteristic a weight of 1 minus the proportion of factor models that can match the cross-section of returns based on the characteristic under consideration. For comparison, the figures also show the rankings of the CAPM and the Fama-French three-factor model.

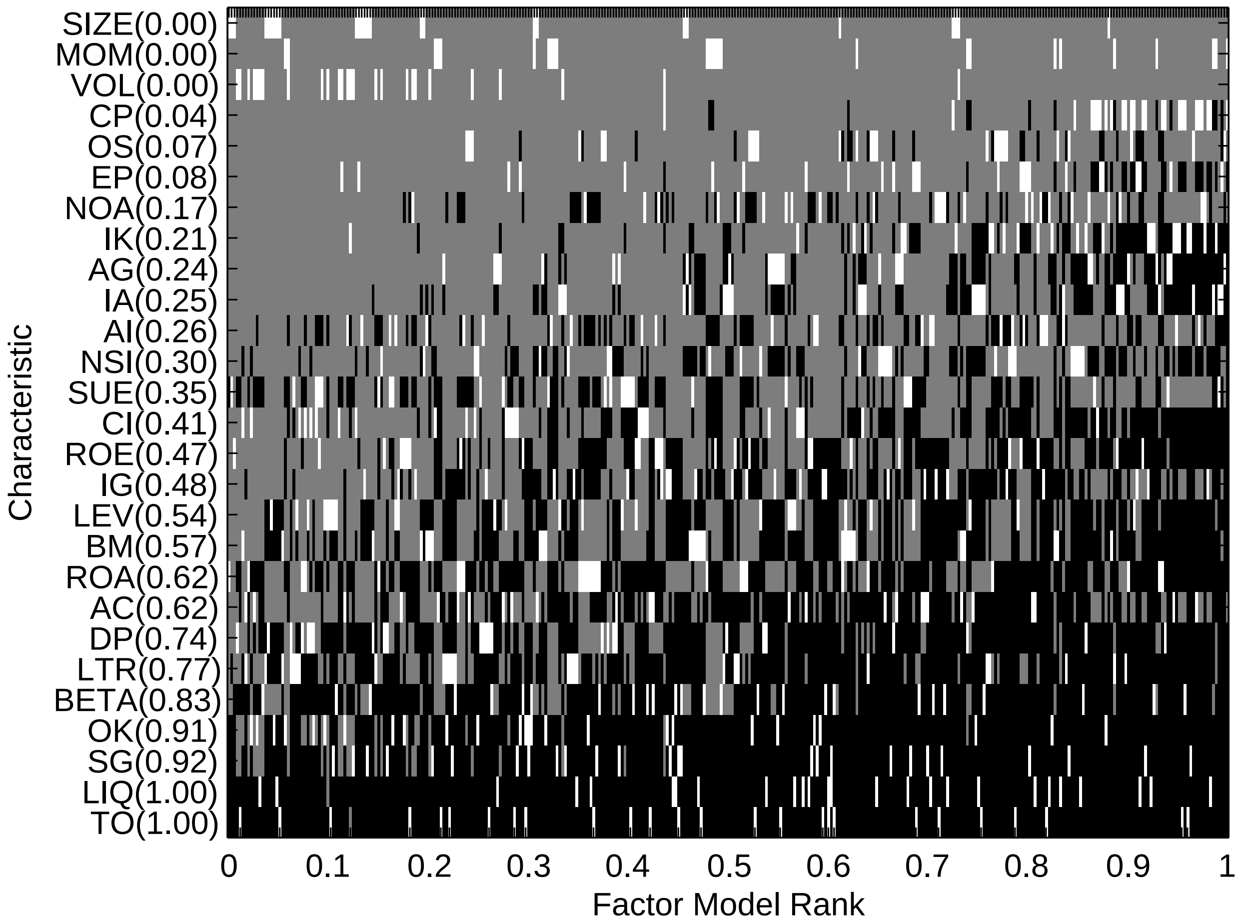
Figure 3: Factor Model Performance

Figure 3 shows a heatmap matrix representation of overall factor model performance. The universe of factor models is all three-factor models consisting of the market portfolio and two characteristic return factors from our list of 27. Factors are the high minus low portfolio from sorting firms into ten portfolios with respect to the underlying firm characteristic. Factor models are ordered along the x-axis in increasing proportion of characteristics matched; characteristics are ordered along the y-axis in decreasing frequency matched (listed in parentheses). Cell (i, j) is shaded black if factor model i is able to match characteristic j, shaded gray if factor model i is unable to match characteristic j, and shaded white if factor model i comprises of a factor constructed from characteristic j. We present figures for the whole sample period 1971-2011 and subsamples 1971-1991 and 1992-2011.

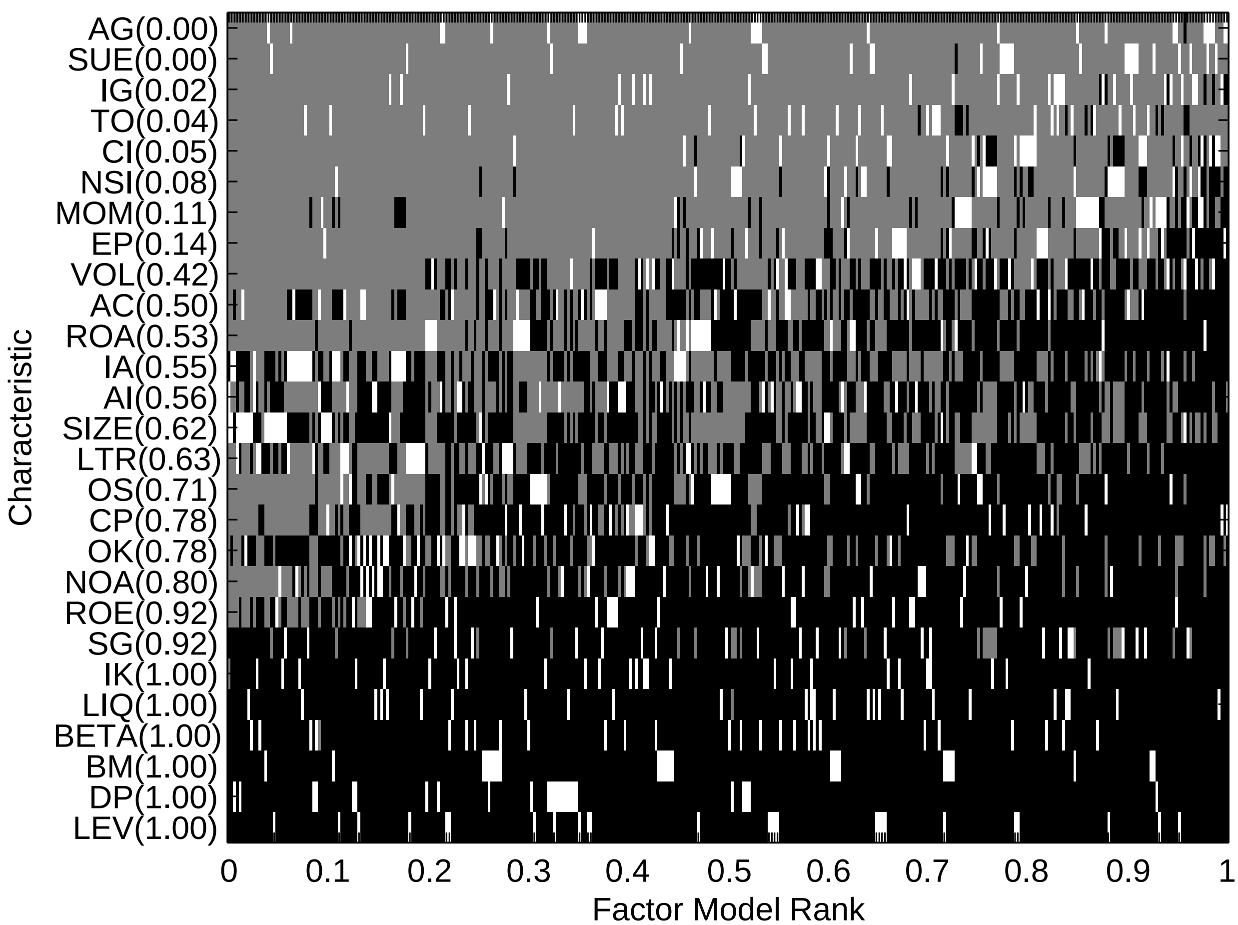
Characteristic abbreviations are as follows: size (SIZE), book-to-market (BM), dividend-to-price (DP), earnings-to-price (EP), cash flow-to-price (CP), investment-to-assets (IA), asset growth (AG), accruals (AC), abnormal investment (AI), net operating assets (NOA), investment-to-capital (IK), investment growth (IG), momentum (MOM), long-term reversal (LTR), return on assets (ROA), standardized unexpected earnings (SUE), return on equity (ROE), sales growth (SG), Ohlson score (OS), market leverage (LEV), net stock issues (NSI), composite issuance (CI), organization capital (OK), liquidity risk (LIQ), turnover (TO), idiosyncratic return volatility (VOL), and market beta (BETA). Details on characteristic definitions and construction is in Appendix A.

(a) 1971-2011




SIZE(0.00)
MOM(0.00)
VOL(0.00)
CP(0.04)
OS(0.07)
EP(0.08)
NOA(0.17)
IK(0.21)
AG(0.24)
IA(0.25)
AI(0.26)
NSI(0.30)
SUE(0.35)
CI(0.41)
ROE(0.47)
IG(0.48)
LEV(0.54)
BM(0.57)
ROA(0.62)
AC(0.62)
DP(0.74)
LTR(0.77)
BETA(0.83)
OK(0.91)
SG(0.92)
LIQ(1.00)
TO(1.00)
Characteristic
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Factor Model Rank

(b) 1971-1991

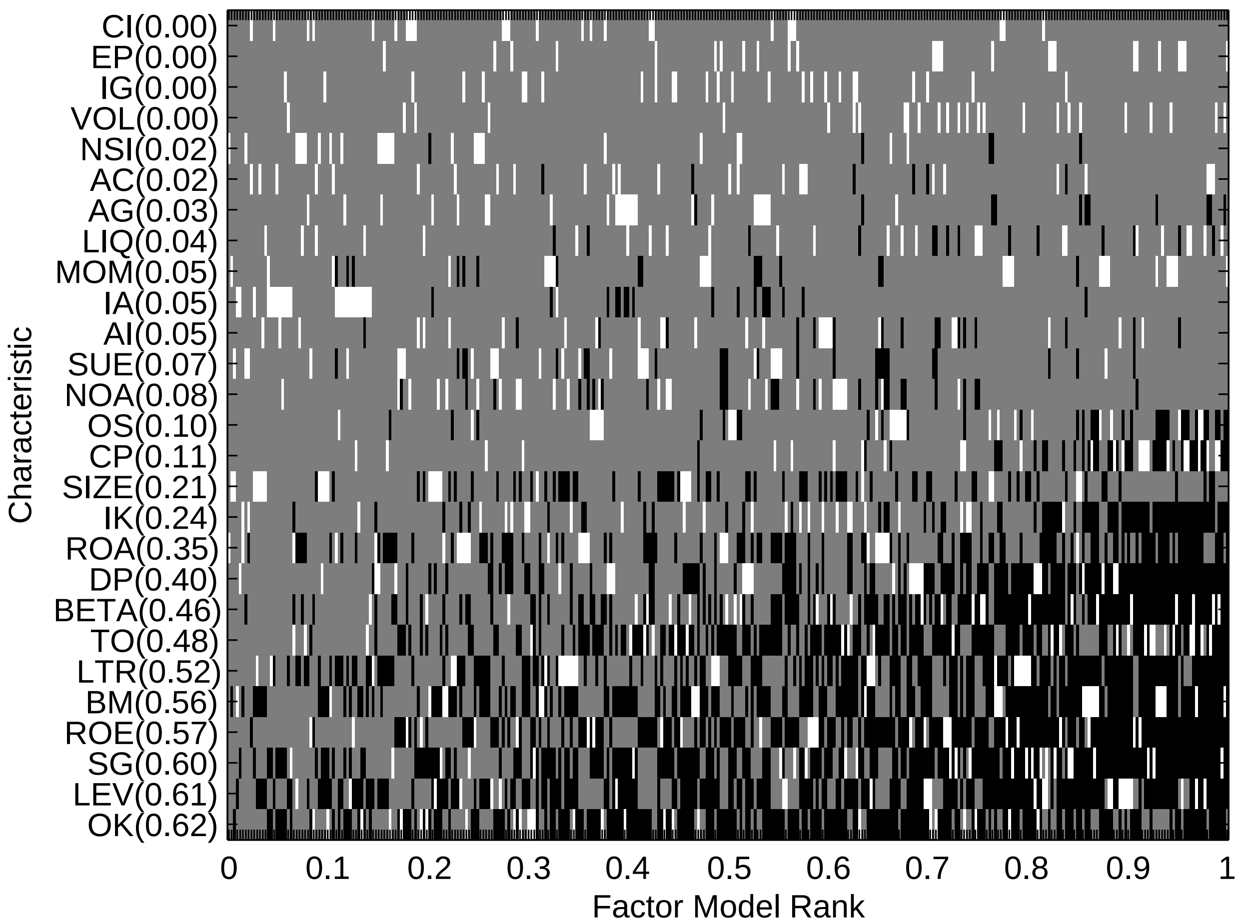

AG(0.00)
SUE(0.00)
IG(0.02)
TO(0.04)
CI(0.05)
NSI(0.08)
MOM(0.11)
EP(0.14)
VOL(0.42)
AC(0.50)
ROA(0.53)
IA(0.55)
AI(0.56)
SIZE(0.62)
LTR(0.63)
OS(0.71)
CP(0.78)
OK(0.78)
NOA(0.80)
ROE(0.92)
SG(0.92)
IK(1.00)
LIQ(1.00)
BETA(1.00)
BM(1.00)
DP(1.00)
LEV(1.00)
Characteristic
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Factor Model Rank

(c) 1992-2011

Figure 4: Factor Model Performance - Double Sort

Figure 4 shows a heatmap matrix representation of overall factor model performance. The universe of factor models is all three-factor models consisting of the market portfolio and two characteristic return factors from our list of 27. Factors are constructed as the equal-weighed average of the high minus low portfolio within the big and small size group, from a double-sort first on size and then the characteristic. Factor models are ordered along the x-axis in increasing proportion of characteristics matched; characteristics are ordered along the y-axis in decreasing frequency matched (listed in parentheses). Cell (i, j) is shaded black if factor model i is able to match characteristic j, shaded gray if factor model i is unable to match characteristic j, and shaded white if factor model i comprises of a factor constructed from characteristic j. We present the figure for the whole sample period 1971-2011.

Characteristic abbreviations are as follows: size (SIZE), book-to-market (BM), dividend-to-price (DP), earnings-to-price (EP), cash flow-to-price (CP), investment-to-assets (IA), asset growth (AG), accruals (AC), abnormal investment (AI), net operating assets (NOA), investment-to-capital (IK), investment growth (IG), momentum (MOM), long-term reversal (LTR), return on assets (ROA), standardized unexpected earnings (SUE), return on equity (ROE), sales growth (SG), Ohlson score (OS), market leverage (LEV), net stock issues (NSI), composite issuance (CI), organization capital (OK), liquidity risk (LIQ), turnover (TO), idiosyncratic return volatility (VOL), and market beta (BETA). Details on characteristic definitions and construction is in Appendix A.



A Appendix: Construction of Characteristics

We provide details on the definitions and construction of 27 firm characteristics.

A.1 Valuation

Size (SIZE)

Stocks with low market capitalization have abnormally high average returns (Banz (1981), Fama and French (1992)). Size is defined to be the log of market capitalization.

Book-to-Market (BM)

Stocks with high book-to-market have abnormally high average returns (Rosenberg, Reid, and Lanstein (1985), Chan, Hamao, and Lakonishok (1991), Fama and French (1992)). The effect remains after controlling for many other variables and is strongest among smaller stocks (Fama and French (1993), Fama and French (2008)).

Dividend-to-Price (DP)

There is a positive association between stock returns and dividend yield (Litzenberger and Ramaswamy (1982), Miller and Scholes (1982)). However, more recently, it has been shown that dividend yield has little predictive power for future returns (Lewellen (2011)).

Earnings-to-Price (EP)

Stocks with high earnings-to-price have abnormally high average returns (Basu (1977), Basu (1983)). The effect seems to be subsumed by size and book-to-market (Fama and French (1992), Fama and French (1996)). The earnings measure is total earnings before extraordinary items.

Cash Flow-to-Price (CP)

Stocks with high cash flow-to-price ratios have abnormally high average returns. Cash flow is total earnings before extraordinary items, plus equity's share of depreciation, plus deferred taxes if available.

A.2 Investment

Investment-to-Assets (IA)

Stocks with low investment-to-assets ratios have abnormally high average returns (Lyandres, Sun, and Zhang (2008), Chen, Novy-Marx, and Zhang (2010)). Following Chen et al. (2010), we define investment-to-assets as the annual change in property, plant, and equipment (Compustat item PPEGT) plus annual change in total inventories (Compustat item INVT) divided by lagged total assets (Compustat item AT).

Asset Growth (AG)

Stocks with low asset growth have abnormally high average returns (Cooper, Gulen, and Schill (2008)). The effect is not very robust to sorting within different size groups and is absent for large stocks (Fama and French (2008)). Asset growth is the percentage change in total assets (Compustat item AT).

Accruals (AC)

Stocks with low accruals have abnormally high average returns (Sloan (1996)). Accruals is the change in current assets (Compustat item ACT) minus the change in cash and short-term investments (Compustat item CASH) minus the change in current total liabilities (Compustat item LCT) plus the change in debt in current liabilities (Compustat item DLC) plus the change in income taxes payable (Compustat item TXP) minus depreciation and amortization (Compustat item DP). All of this is divided by the average of total assets (Compustat item AT) over fiscal year $t-1$ and $t-2$.

Abnormal Investment (AI)

Stocks with low abnormal investment have abnormally high average returns (Fairfield, Whisenant, and Yohn (2003), Titman, Wei, and Xie (2004)). Abnormal investment is the deviation of current investment from the past three year moving average. Investment is defined to be the ratio of capital expenditure (Compustat item CAPX) over the net sales turnover ratio (Compustat item SALE).

Net Operating Assets (NOA)

Stocks with low net operating assets have abnormally high average returns (Hirshleifer, Hou, Teoh, and Zhang (2004)). Net operating assets is defined as follows:

$$\begin{aligned} NOA(t) = & [AT(t-1) - CHE(t-1)] - [AT(t-1) - DLC(t-1) \\ & -DLTT(t-1) - MIB(t-1) - PSTK(t-1) - CEQ(t-1)] \end{aligned}$$

where AT is total assets, CHE is cash and short-term investments, DLC is debt in current liabilities, $DLTT$ is long term debt, MIB is non-controlling interest, $PSTK$ is preferred capital stock, and CEQ is common equity.

Investment-to-Capital (IK)

Stocks with low investment-to-capital ratios have abnormally high average returns (Xing (2008)). Investment to capital is the ratio of capital expenditure (Compustat item CAPX) over property, plant, and equipment (Compustat item PPENT).

Investment Growth (IG)

Stocks with low investment growth rates have abnormally high average returns (Xing (2008)). Investment growth is the percentage change in capital expenditure (Compustat item CAPX).

A.3 Prior Returns

Momentum (MOM)

Stocks with high returns over the last year have abnormally high average returns for the next few months (Jegadeesh and Titman (1993), Chan, Jegadeesh, and Lakonishok (1996)). The effect is robust to sorting within different size groups (Fama and French (2008)). Momentum in month t is defined as the cumulated continuously compounded stock return from month $t-12$ to month $t-2$.

Long-term Reversal (LTR)

Stocks with low returns over the past 3-5 years have abnormally high average returns (DeBondt and Thaler (1985)). The effect is not present after accounting for the Fama French factors (Fama and French (1996)). Long-term reversal in month t is defined as the cumulated continuously compounded stock return from month $t-60$ to month $t-13$.

A.4 Earnings

Return on Assets (ROA)

Stocks with high return on assets have abnormally high average returns (Chen et al. (2010)). Return on assets is defined to be the ratio of income before extraordinary items (Compustat item IBQ) over total assets (Compustat item ATQ).

Standardized unexpected earnings (SUE)

Post-earnings announcement drift is the tendency for a stock's returns to drift in the direction of an earnings surprise for several weeks after an earnings announcement. Stocks with high SUE have abnormally high average returns (Ball and Brown (1968), Bernard and Thomas (1989)). SUE is defined to be the change in the most recently announced quarterly earnings per share (Compustat item EPSPIQ) from its announced value four quarters ago divided by the standard deviation of the change in quarterly earnings over the prior eight quarters.

Return on Equity (ROE)

More profitable firms have abnormally high average returns (Haugen and Baker (1996), Cohen, Gompers, and Vuolteenaho (2002), Piotroski (2000), Fama and French (2006)). The effect is not as robust as there is little evidence that unprofitable firms have unusually low returns (Fama and French (2008)). Return on equity is defined to be the ratio of equity income over book value of equity. Equity income is income before extraordinary items (Compustat item IB) minus preferred dividends (Compustat item DVP) plus deferred income taxes (Compustat item TXDI), if available.

Sales Growth (SG)

Stocks with low past sales growth have abnormally high average returns (Lakonishok, Shleifer, and Vishny (1994)). Sales growth is the percent change in net sales over turnover (Compustat item SALE).

A.5 Financial Distress

Ohlson Score (OS)

Stocks with lower Ohlson score (lower probability of default) have abnormally high average returns. OS is computed using Model One Table 4 of Ohlson (1980).

Market Leverage (LEV)

Stocks with higher market leverage have abnormally high average returns (Bhandari (1988)). The predictive power of leverage is subsumed by the book to market effect in returns (Fama and French (1992)). Market leverage is the ratio of total assets (Compustat item AT) over the market value of equity.

A.6 External Financing

Net Stock Issues (NSI)

Stocks with low net stock issues have abnormally high average returns (Fama and French (2008), Pontiff and Woodgate (2008)), where returns after stock repurchases are high (Ikenberry, Lakonishok, and Vermaelen (1995)) and returns after stock issues are low (Loughran and Ritter (1995)). Net stock issues is the log of the ratio of split-adjusted shares outstanding at fiscal year end $t-1$ and $t-2$. Split-adjusted shares outstanding is the product of common shares outstanding (Compustat item CSHO) and the cumulative adjustment factor (Compustat item ADJEXC).

Composite Issuance (CI)

Stocks with low composite issuance have abnormally high average returns (Daniel and Titman (2006)). The five year composite issuance measure is defined as:

$$\iota(t-\tau) = log(\frac{ME_t}{ME_{t-\tau}}) - r(t-\tau, t)$$

where $r(t-\tau, t)$ is the cumulative log return on the stock from the last trading day of calendar year $t-6$ to the last trading day of calendar year $t-1$, and $ME(t)$ ($ME(t-\tau)$) is total market equity on the last trading day of calendar year t ($t-6$).

A.7 Other

Organization Capital (OK)

Eisfeldt and Papanikolaou (2012) find that firms with more organization capital relative to industry peers outperform firms with less organization capital. The stock of organization capital is (1-depreciation rate) of organization capital from one period before plus the deflated value of selling, general, and administrative expenses (Compustat item XSGA). Following the original paper, we sort on the ratio of organization capital to physical capital.

Liquidity Risk (LIQ)

Firms with high liquidity betas have higher returns than firms with low liquidity betas (Pastor and Stambaugh (2003)). Liquidity beta is measured as the loading on innovations in aggregate liquidity, in a regression of excess returns on the Fama French three factors and aggregate liquidity innovation.

Turnover (TO)

Average turnover over the past 3-12 months is negatively related to subsequent returns (Lee and Swaminathan (2000)). Turnover is defined to be the ratio of shares traded over shares outstanding.

Idiosyncratic Return Volatility (VOL)

Ang et al. (2006) find that firms with high idiosyncratic return volatility have abnormally low returns. Idiosyncratic volatility is measured as the standard deviation of residuals from a regression of daily excess returns on the Fama French three factor model.

Market Beta (BETA)

Frazzini and Pedersen (2011) find that a portfolio long on assets with high market betas and short on assets with low market betas exhibits significantly negative risk-adjusted returns. Market beta is estimated as the sum of the coefficients from regressing an asset's daily excess returns on current and lagged excess returns of the market portfolio, with lags up to 5 trading days.

B Appendix: Performance of Four-Factor Models

Table B.1: Top 20 Performing Four-Factor Models

	1971-2011				1971-1991				1992-2011			
	C1	C2	C3	prop	C1	C2	C3	prop	C1	C2	C3	prop
1	MOM	OS	AG	0.67	MOM	IA	NOA	0.83	MOM	ROA	AG	0.92
2	SIZE	MOM	VOL	0.63	MOM	LTR	NOA	0.83	MOM	NSI	AG	0.92
3	SIZE	LIQ	VOL	0.63	MOM	NOA	CP	0.83	MOM	NSI	LIQ	0.92
4	BM	MOM	CP	0.63	MOM	CP	LIQ	0.83	MOM	CI	LIQ	0.92
5	MOM	IA	EP	0.63	MOM	CP	TO	0.83	ROA	AG	NOA	0.92
6	MOM	NSI	LIQ	0.63	SIZE	MOM	NOA	0.79	ROA	AG	IK	0.92
7	MOM	AG	CP	0.63	MOM	IA	CP	0.79	ROA	AG	TO	0.92
8	MOM	CI	LIQ	0.63	MOM	LTR	OS	0.79	AG	SUE	CP	0.92
9	SUE	AI	CP	0.63	MOM	AG	NOA	0.79	MOM	AG	SUE	0.88
10	SUE	CP	IG	0.63	MOM	AC	CP	0.79	MOM	AG	CP	0.88
11	SUE	CP	LIQ	0.63	MOM	AI	CP	0.79	MOM	SUE	CI	0.88
12	CP	IG	LIQ	0.63	MOM	CP	IK	0.79	ROA	AG	AI	0.88
13	SIZE	VOL	BETA	0.58	MOM	CP	BETA	0.79	ROA	AG	LIQ	0.88
14	BM	SUE	CP	0.58	IA	OS	CP	0.79	ROA	AG	SG	0.88
15	MOM	IA	CP	0.58	IA	NSI	EP	0.79	AG	SUE	CI	0.88
16	MOM	AG	EP	0.58	IA	SUE	NOA	0.79	AG	SUE	EP	0.88
17	MOM	AI	CP	0.58	IA	SUE	CP	0.79	AG	SUE	ROE	0.88
18	MOM	CP	IG	0.58	IA	EP	IG	0.79	AG	SUE	VOL	0.88
19	MOM	CP	LIQ	0.58	IA	EP	LIQ	0.79	AG	LIQ	VOL	0.88
20	MOM	CP	SG	0.58	BM	MOM	NOA	0.75	AG	VOL	TO	0.88

Table B.1 lists the characteristic-based factors that constitute the top twenty linear four-factor models, in terms of the proportion of remaining characteristics they can capture, via the equal-weighted method. We say that a factor model M captures, or spans, a characteristic C, if the p-value from the Gibbons et al. (1989) F-test of joint significance of abnormal average return with respect to M across the ten sorted portfolios on C is above 10%. Top factor models are shown for the whole sample period 1971-2011 and subsamples 1971-1991 and 1992-2011.

The universe of factor models is all four-factor models consisting of the market portfolio and three characteristic return factors (C1, C2, C3) from our list of 27. Characteristic abbreviations are as follows: size (SIZE), book-to-market (BM), dividend-to-price (DP), earnings-to-price (EP), cash flow-to-price (CP), investment-to-assets (IA), asset growth (AG), accruals (AC), abnormal investment (AI), net operating assets (NOA), investment-to-capital (IK), investment growth (IG), momentum (MOM), long-term reversal (LTR), return on assets (ROA), standardized unexpected earnings (SUE), return on equity (ROE), sales growth (SG), Ohlson score (OS), market leverage(LEV), net stock issues (NSI), composite issuance (CI), organization capital (OK), liquidity risk (LIQ), turnover (TO), idiosyncratic return volatility (VOL), and market beta (BETA). Details on characteristic definitions and construction is in Appendix A.

Figure B.1: Four-Factor Model Performance

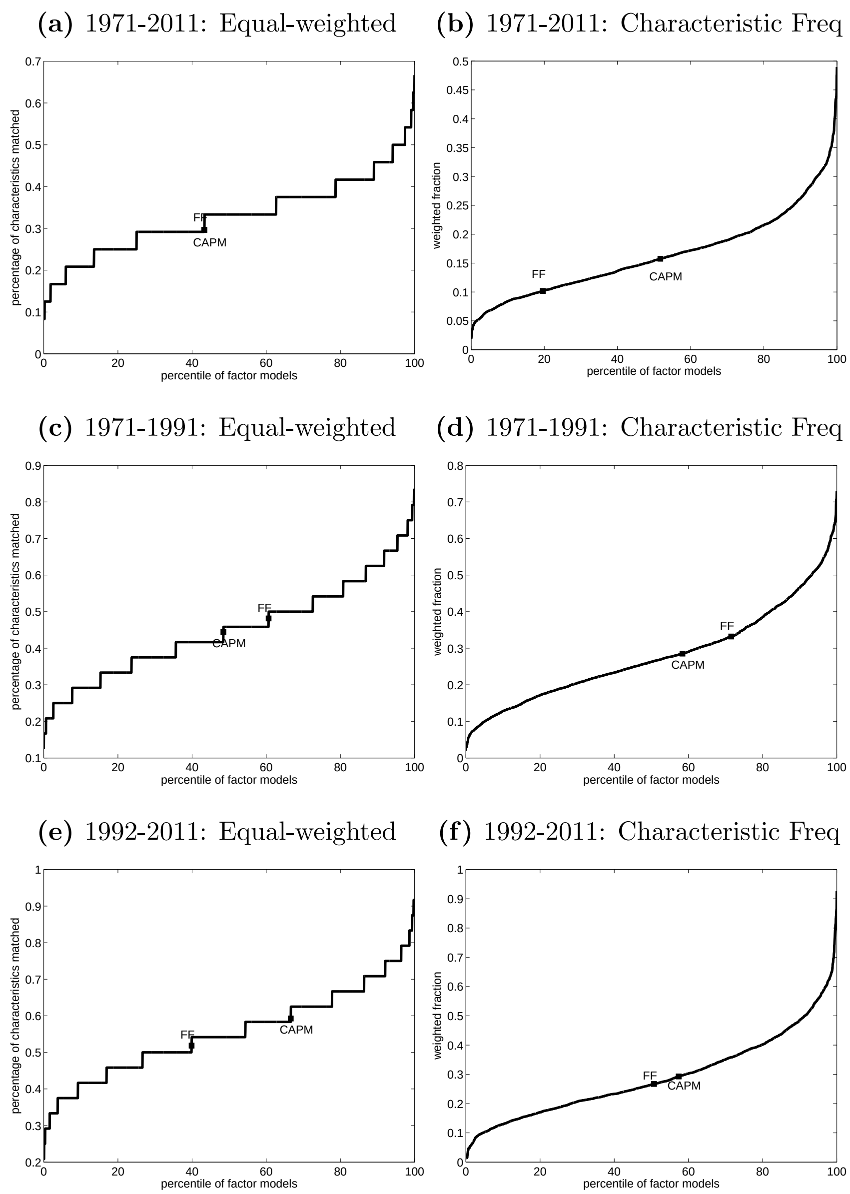


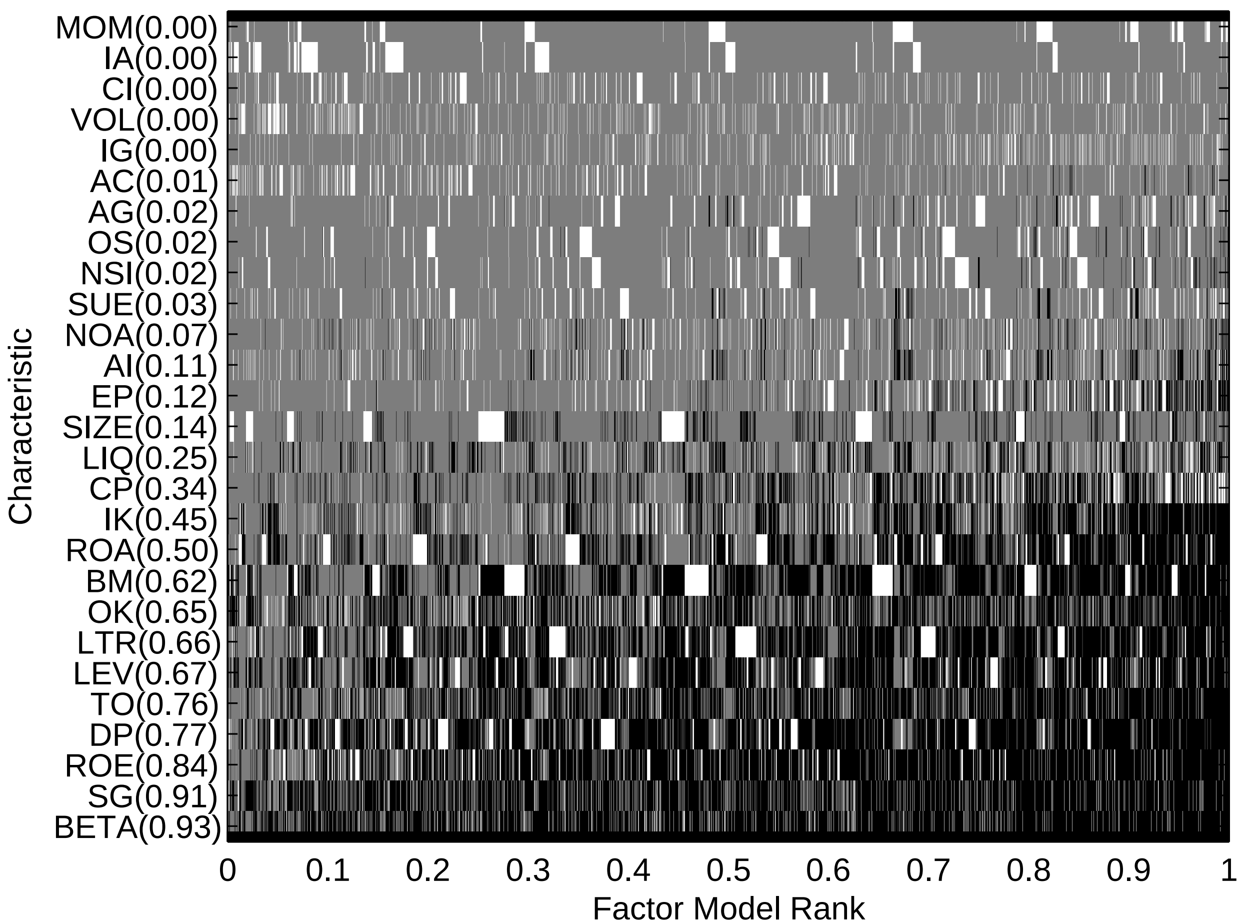
Figure B.1 displays the distribution of four-factor model performance, as measured by the percentage of characteristics matched, over the whole sample period 1971-2011 and subsamples 1971-1991 and 1992-2011. The universe of factor models is all four-factor models consisting of the market portfolio and three characteristic return factors from our list of 27. The percentage of characteristics matched is computed using two characteristic weighting methods: equal-weighted method and characteristic matching frequency method. The "equal-weighted" method gives an equal weight to each characteristic matched. The "characteristic matching frequency" method gives each characteristic a weight of 1 minus the proportion of factor models that can match the cross-section of returns based on the characteristic under consideration. For comparison, the figures also show the rankings of the CAPM and the Fama-French-Carhart four-factor model (consisting of the market, SMB, HML, and MOM).

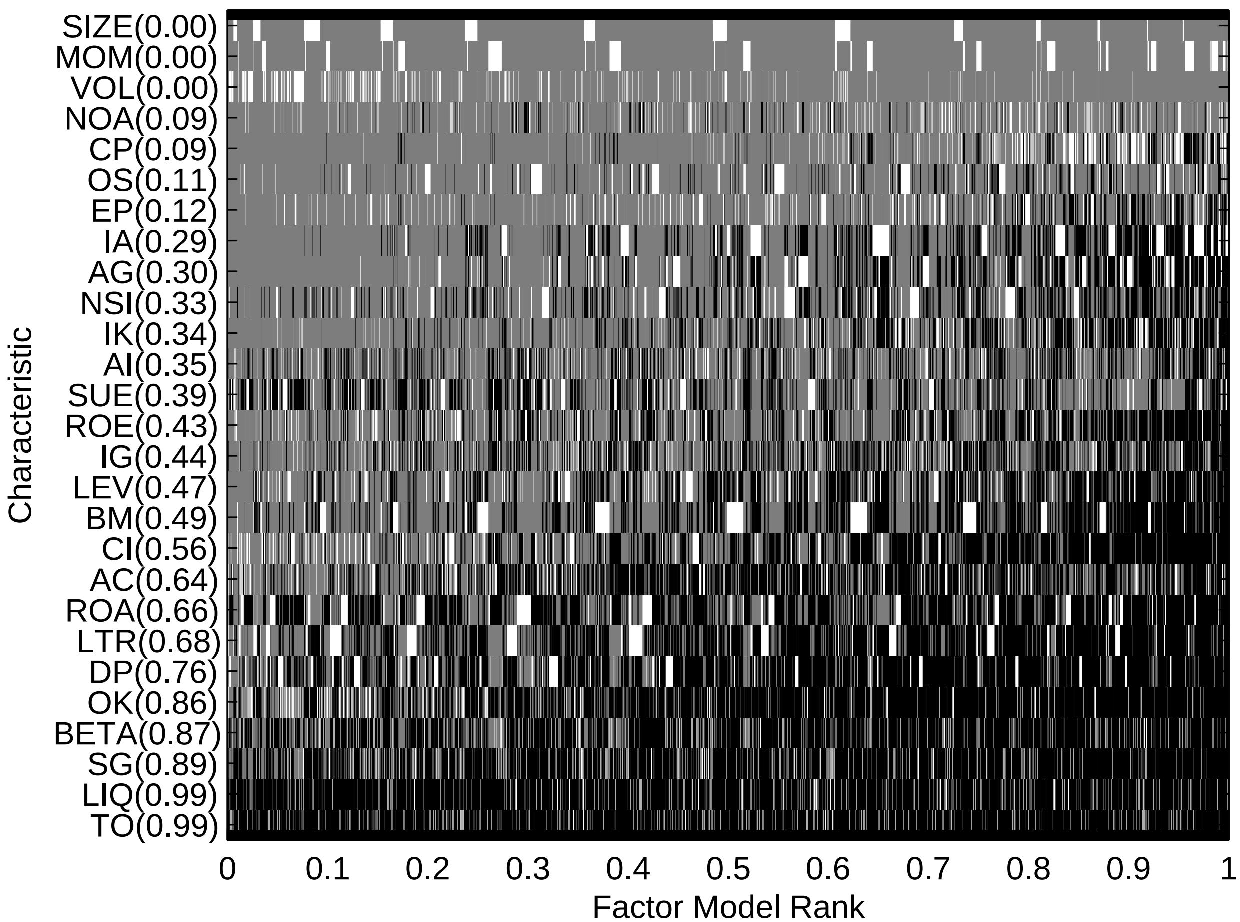
Figure B.2: Four-Factor Model Performance

Figure B.2 shows a heatmap matrix representation of overall factor model performance. The universe of factor models is all four-factor models consisting of the market portfolio and three characteristic return factors from our list of 27. Factors are the high minus low portfolio from sorting firms into ten portfolios with respect to the underlying firm characteristic. Factor models are ordered along the x-axis in increasing proportion of characteristics matched; characteristics are ordered along the y-axis in decreasing frequency matched (listed in parentheses). Cell (i, j) is shaded black if factor model i is able to match characteristic j, shaded gray if factor model i is unable to match characteristic j, and shaded white if factor model i comprises of a factor constructed from characteristic j. We present figures for the whole sample period 1971-2011 and subsamples 1971-1991 and 1992-2011.

Characteristic abbreviations are as follows: size (SIZE), book-to-market (BM), dividend-to-price (DP), earnings-to-price (EP), cash flow-to-price (CP), investment-to-assets (IA), asset growth (AG), accruals (AC), abnormal investment (AI), net operating assets (NOA), investment-to-capital (IK), investment growth (IG), momentum (MOM), long-term reversal (LTR), return on assets (ROA), standardized unexpected earnings (SUE), return on equity (ROE), sales growth (SG), Ohlson score (OS), market leverage (LEV), net stock issues (NSI), composite issuance (CI), organization capital (OK), liquidity risk (LIQ), turnover (TO), idiosyncratic return volatility (VOL), and market beta (BETA). Details on characteristic definitions and construction is in Appendix A.

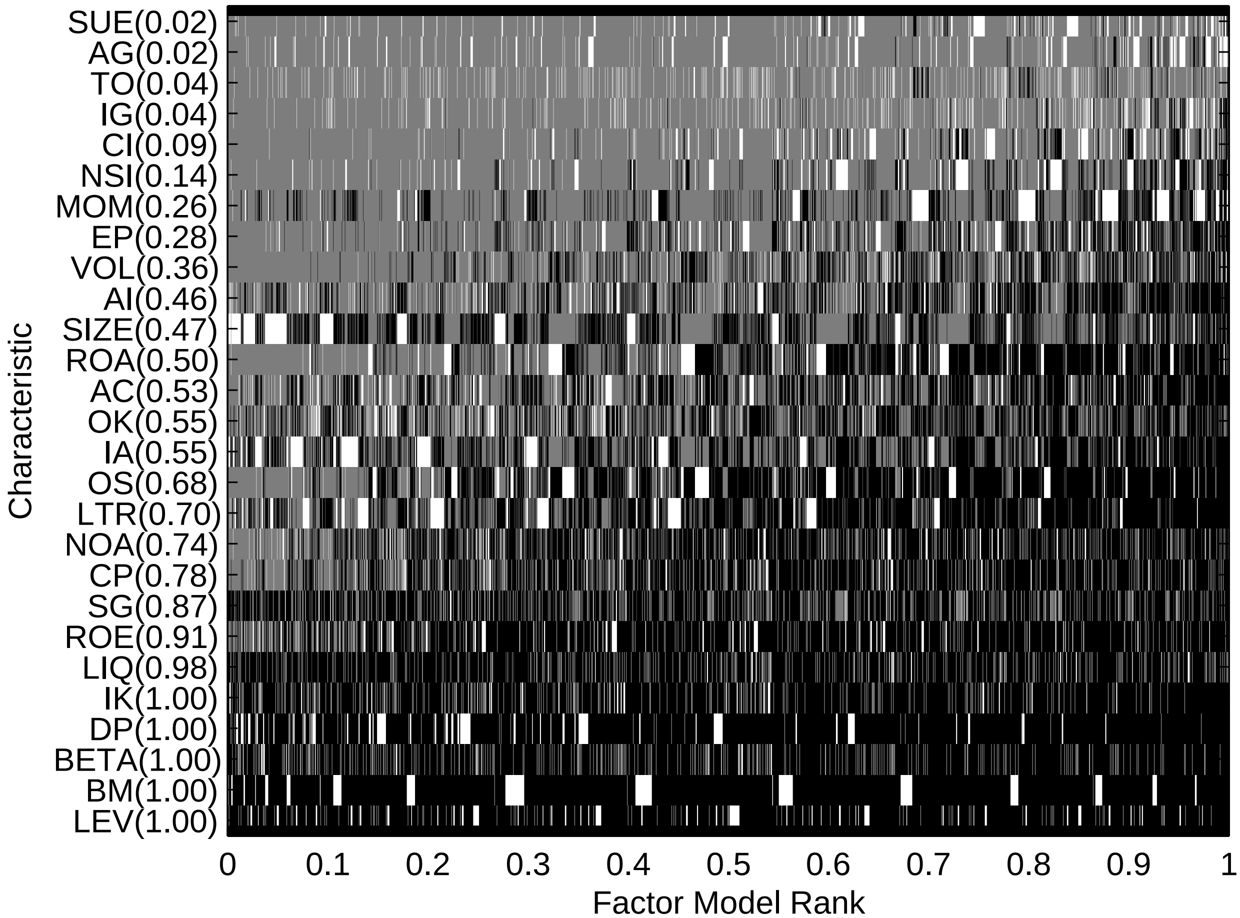
(a) 1971-2011





SIZE(0.00)
MOM(0.00)
VOL(0.00)
NOA(0.09)
CP(0.09)
OS(0.11)
EP(0.12)
IA(0.29)
AG(0.30)
NSI(0.33)
IK(0.34)
AI(0.35)
SUE(0.39)
ROE(0.43)
IG(0.44)
LEV(0.47)
BM(0.49)
CI(0.56)
AC(0.64)
ROA(0.66)
LTR(0.68)
DP(0.76)
OK(0.86)
BETA(0.87)
SG(0.89)
LIQ(0.99)
TO(0.99)
Characteristic
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Factor Model Rank

(b) 1971-1991


SUE(0.02)
AG(0.02)
TO(0.04)
IG(0.04)
CI(0.09)
NSI(0.14)
MOM(0.26)
EP(0.28)
VOL(0.36)
AI(0.46)
SIZE(0.47)
ROA(0.50)
AC(0.53)
OK(0.55)
IA(0.55)
OS(0.68)
LTR(0.70)
NOA(0.74)
CP(0.78)
SG(0.87)
ROE(0.91)
LIQ(0.98)
IK(1.00)
DP(1.00)
BETA(1.00)
BM(1.00)
LEV(1.00)
Characteristic
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Factor Model Rank

(c) 1992-2011