Characteristics and the Cross-Section of Covariances

Charles Clarke[*] Matthew Linn[†‡]

December 12, 2022

Abstract

We model firm-level, stock return covariances as a function of firm characteristics. Flexible panel regressions allow us to estimate the marginal predictive power associated with characteristics in a multi-dimensional setting where portfolio sorts are infeasible. We use the model to identify characteristics that proxy for priced factors, unpriced factors, and near-arbitrages while circumventing the need to identify underlying risk factors. Cyclical variation in how characteristics are related to covariances shows that many well-known characteristics proxy for exposure to business cycle risk while few proxy for market sentiment.

JEL classification: G12

[*]University of Kentucky, Gatton School of Business, Department of Finance and Quantitative Methods, Lexington, KY 40506; charlie.clarke@uky.edu

[†]University of Massachusetts, Isenberg School of Business, Department of Finance; mlinn@isenberg.umass.edu

[‡]We thank Robert Dittmar, Joseph Engelberg, Mark Kamstra (discussant), Andrew Karolyi, Robert Korajczyk, Jonathan Lewellen, Ben Matthies (discussant), Seth Pruitt, Sophie Shive, participants at the 2018 Northern Finance Association annual meeting and the 2019 Financial Management Association annual meeting for helpful comments and suggestions.

1 Introduction

We study the relationship between stock covariances and firm characteristics in order to gain insight into the relationship between characteristics and systematic risk exposure. We introduce a simple and transparent regression model that allows us to directly test whether a characteristic proxies for systematic risk without having to pre-specify or extract latent factors. This approach leverages the rich cross-section of pairwise, firm covariances to shed new light on the characteristics versus covariances debate. We are able to classify characteristics into proxies for priced risk, unpriced risk, and statistical arbitrage. We show much of this characteristic associated covariance varies cyclically with the business cycle.

A main stream of the recent literature tackles the multivariate relation between characteristics and the cross-section of expected returns. As discussed in Harvey et al. (2016), decades of academic research has produced hundreds of characteristics that are correlated with average returns. (Cochrane, 2011) points out that in addition to understanding relationships between characteristics and average returns, it is important to understand the relationships between characteristics and covariances:

> Asset pricing really is about the equality of two *functions*: The function relating means to characteristics should be proportional to the function relating covariances to characteristics (Cochrane, 2011).

Cochrane (2011) calls for progress in understanding the function mapping characteristics to covariances and how it relates to the function mapping characteristics to mean returns.

We adopt a very general linear factor model with both priced and unpriced factors as the return generating process underlying our analysis. Daniel et al. (2020) study this process and stress the importance of including unpriced factors in the return generating process of stocks. Within this family of linear models, the function relating expected returns to characteristics depends upon loadings on *priced* factors and potentially alpha. Covariance between stocks arises through covariation with common underlying factors, whether priced or unpriced. We provide novel multivariate analysis of the characteristic and covariance relationship in both unconditional and conditional settings, and we show empirically how these covariance functions relate to multivariate functions of average returns. In doing so, our paper contributes to the characteristics versus covariance literature by examining the role of characteristics through a new lens: the multivariate relationship between characteristics and the cross-section of covariances.

Ross's Arbitrage Pricing Theory (APT) (Ross, 1976) posits that latent factors generate stock returns. Within this paradigm a predictor of covariance predicts loadings on underlying factors.

These factors could be priced or unpriced, where a priced factor confers a risk premium. Importantly, as emphasized by Kozak et al. (2018), "risk premium" in this sense only distinguishes that a factor contributes to expected returns and is not a statement about whether this premium arises for rational or behavioral reasons or represents "risk" in a deeper economic sense. If pairwise covariances between stocks are significantly predicted by a characteristic, that characteristic is a proxy for loadings on some underlying factor(s). Importantly, this is true regardless of what the true underlying factors are. Our model determines which characteristics are significant predictors of pairwise covariances. From the model's predictions we infer which characteristics are predictors of loadings on factors underlying the return generating process.

We first show there is considerable diversity in the way characteristics relate to covariances and hence systematic risk. Many characteristics that are strong predictors of covariance in univariate settings are insignificant once we control for other characteristics. We therefore focus on the marginal predictive power of characteristics. Our analysis uses individual stocks' characteristics as opposed to sorted portfolios since each stock has a well-defined vector of characteristics. This allows us to easily control for many confounding characteristic effects where portfolio sorting is often infeasible. Our results suggest that on the margin some characteristics predict average returns and covariances, consistent with proxying for exposure to priced risk. Other characteristics predict only covariance and not average returns, consistent with proxying for exposure to unpriced risk, and still others predict only average returns and not covariances, consistent with a statistical near-arbitrage, an "anomaly." The results suggest the common methodology of building factors of high minus low portfolio sorts on one or a small number of characteristics may obscure much of the variety and interplay across characteristics which our approach highlights.

Distilling the massive breadth of cross-sectional predictors in an intelligible way presents a major challenge when trying to understand the cross-section of covariances. Our goal is not necessarily to maximally explain or predict covariances, but rather to capture relationships across important cross-sectional characteristics in a transparent and interpretable way. To this end, we examine a set of 13 characteristic groups organized by Jensen et al. (Forthcoming) who cluster 153 separate firm characteristics by theme in order to reduce the dimensionality of the characteristic space. Importantly, the themed groupings neatly classify the large number of existing characteristics into disparate groups, which largely span the space of characteristics from the literature. Using this classification allows us to digest a large number of related firm characteristics while maintaining interpretability. The thirteen groups (and the number of grouped characteristics) include: Accruals* (5), Debt Issuance* (6), Investment* (22), Leverage* (11), Low Risk (21), Momentum (8), Profitability (12), Profit Growth (13), Quality (17), Seasonality (14), Size* (5), Skewness* (6), and Value (13), where an asterisk designates that average returns are highest with low values of the named characteristics and decreasing with higher values.

We analyze quarterly pairwise covariances constructed using daily stock returns over the course of the quarter. This short time horizon allows us to capture covariances on the same scale as short horizon return predictability. Since daily return data can potentially bias measures of covariation due to frictions and non-synchronous trading, we take measures to mitigate these concerns. We remove all micro caps from the sample since their daily returns can be heavily influenced by illiquidity and microstructure noise. These stocks account for only 3% of the market equity of the CRSP universe (Fama and French, 2008). Additionally, we adjust the remaining sample for asynchronous trading as prescribed in Shanken (1987).

Each characteristic group is divided into 5 quintiles and each stock is assigned to a quintile for each of the 13 characteristics. We begin with univariate analysis of each characteristic group separately. For instance, for the momentum group we first categorize each firm by its composite score over eight momentum measures. Since average returns are increasing in the momentum variables, the highest momentum stocks are in quintile five and the lowest momentum stocks are in quintile one. Each quarterly covariance observation comes from a pair of two stocks and each has a momentum quintile. This yields fifteen momentum pairing groups, (1,1), (1,2)=(2,1), (2,2), (2,3)=(3,2), ..., (5,5), where pairings with quintiles transposed are combined as the ordering is arbitrary.

We then model covariances as a multivariate function that depends upon pairwise characteristic values of all 13 characteristic groups simultaneously with additional controls for industry overlap. Since indicator functions can be used to approximate any measurable function (Billingsley (2008)), we model the cross-section of covariances using a large number of characteristic-based indicator functions that approximate a surface measuring covariation over pairs of characteristic values. We use this multivariate model to examine predictions of standard linear factor models. For example, two stocks with large positive loadings on a common volatile factor will have positive covariance holding all else equal. The same is true of two stocks with large negative loadings on the same factor. If a particular characteristic proxies for loadings on a latent factor, our model will predict two stocks with large values of the characteristic to have positive covariance controlling for other characteristics in the model. Similar logic suggests that the covariance between a high characteristic and low characteristic stock should be small. Our analysis begins by examining whether these predictions are supported by the data.

In univariate models, we find substantial diversity in the patterns of covariance across the 13 characteristics. Some characteristics such as Accruals, Debt Issuance, Momentum, Profit Growth, and Seasonality present as U-shaped patterns with covariance maxima around low characteristic portfolio pairs (1,1), high characteristic pairs (5,5), and low-high pairs (1,5). Others like Investment, Leverage, Low Risk, Profitability, Quality, Size, and Value predict covariance only in high (5,5) or low (1,1) pairs but not both. While the univariate results are of interest as they capture relationships in the high-low univariate sorts typical in the literature, our main focus is on the marginal predictive power of each characteristic, controlling for all other characteristics. Practically, this is important because it allows us to see how tilting a diverse portfolio toward certain characteristics will change the risk profile. We show that the typical characteristic is confounded, as characteristic exposures overlap. For example, high investment firms, tend to be high accruals, high debt issuance, low leverage, high risk, growth stocks. Low momentum firms tend to be small, high risk firms with low profit growth. In contrast to sorts that only control for one or a small number of characteristics, our multivariate regressions are well-suited to tease out the underlying relationships.

When we examine the multivariate association of characteristics with covariances, the patterns across many of the groups dramatically shift. This suggests that much of the comovement we tend to attribute to characteristics in isolation is actually confounded exposure to other characteristics. Across characteristic portfolios there is considerable overlap of firms. Many firms fall in the extreme portfolios of more than one characteristic. This overlap confounds low dimensional characteristic sorts so that covariances of characteristic-sorted portfoios are not appropriate for estimating the marginal predictive power of a characteristic for covariances. Our results show that for those characteristics that have incremental predictive power, covariance is almost always concentrated in only one of the extreme legs. Surprisingly, this predictive power is not consistently concentrated in the leg with high average returns. This pattern is easily missed when analyzing only long-short, characteristic-sorted portfolios.

Next, we examine the extent to which our model predicted covariances and the risks they imply for long-short strategies are compensated with commensurate expected returns. The multivariate model allows us to isolate the marginal risk associated with each characteristic. Combining our multivariate covariance model with an analogous model for expected returns, we examine the model-implied risk-return trade-off associated with each characteristic. We find significant variation across predictors. Some behave like risk factors, they are strong predictors of expected returns and covariances. Some behave like unpriced factors, they predict covariances, but not expected returns, and some behave like anomalies, they predict expected returns but not covariances, producing high implied Sharpe ratios.

We then combine our characteristics into a proxy for expected returns. We show that the expected return proxy explains only a very small portion of the explainable covariance, less than 10% of the baseline model and 3% of the most expansive model. We interpret our results through the lens of the one factor rotation. Our results suggest that the SDF only explains a small portion of the explainable variation. This result is inconsistent with the CAPM, but suggests an even stronger claim. It suggests the underlying SDF cannot closely resemble the market factor or any factor that explains a large amount of the covariation across firms.

Finally, a natural question is how the characteristic factor risk we find is associated with "deeper" models of risk. Systematic factor risk can arise for both rational and behavioral reasons. For instance, priced characteristic associated covariance may capture firms differential exposure to the investment opportunity set generating state variable hedging demands (Merton et al., 1973), or conversely, it may capture sentiment-investor demand that is aligned with factors capturing the covariation in cash-flows across firms (Kozak et al., 2018). We look for evidence for and against these deeper models by examining conditional versions of our covariance model. We interact our characteristics with measures capturing business cycle risk and investor sentiment to see if characteristic-associated covariance is higher or lower during times of differential exposure to either market conditions.

The first state variable we use is a recession dummy, which is an intuitive proxy for investor appetite for bearing risk. The second state variable we explore is the lower bound on expected market excess returns as derived in Martin (2017), which the author argues is a good estimate of the time-varying equity premium. This measure is an intuitive proxy for macroeconomic shocks relevant to the marginal investor, but this continuous measure also proxies for the intensity of changes in investor appetite for bearing risk. Lastly, we also interact our coefficients with the market-wide sentiment measure of Baker and Wurgler (2006), which is meant to capture high levels of aggregate investor exuberance or pessimism.

We find that the systematic risk of high-risk stocks, momentum losers and high investment firms interact especially strongly with the recession dummy. Additionally, the covariances of these characteristics along with those of small stocks and value stocks are positively correlated with Martin's (2017) equity premium proxy. This suggests that a number of characteristics proxy for business cycle risk. By contrast, when we interact our indicator model with market-wide sentiment as measured by Baker and Wurgler (2006), only covariances of growth stocks and stocks in technology or internet related industries vary with sentiment. Time-varying sentiment explains far less of the characteristic-covariance relationship than the recession dummy or expected return proxy, but while confined, sentiment associated comovement appears in the high priced and tech sector stocks where it may be most expected.

Our work builds on several important papers. In an early attempt to explore firm by firm covariances, Chan et al. (1999) predict firmwise covariances over 60 months using both past covariances and factor models like Fama and French (1993). Moskowitz (2003) uses a multivariate GARCH model to characterize the time-varying covariance structure of returns of portfolios formed on various firm characteristics. Brandt et al. (2009) link firm characteristics to the investor's portfolio problem by modeling optimal portfolio weights as a function of firm characteristics. Gao (2011) models pairwise annual returns as a function of firm characteristics using a kernel density approach, and then aggregates these exposures to create alternative firm-level measures of a stock's covariation with the market, size, and value factors.

A number of recent papers have used covariances of daily returns as the dependent variable to study the effect of individual characteristics on covariances (or correlations). Anton and Polk (2014) and Greenwood and Thesmar (2011) show stocks connected through mutual funds comove. Lou and Polk (2021) use the comovement of momentum stocks as a proxy for arbitrage activity in those stocks and explore the implications for mispricing. Further, there is a burgeoning literature on "excess comovement" and its implications.[1] While this literature is growing rapidly, little emphasis has been put on developing a "baseline model" of comovement. What is excess comovement in excess of? Most commonly, some combination of Fama and French (2016) characteristics and industry factors are used as the relevant comparison, but relatively little has been done in the more expansive characteristic "zoo." Our results show comovement associated with a characteristic is often absorbed in a multivariate setting and covariance associated with a characteristic is often concentrated in only one extreme end of the characteristic's distibution. This suggests that interpretation of the excess comovement literature may depend crucially on particular regression specifications.

Our paper is also related to the growing literature that tries to tame the factor zoo using novel methods of factor analysis. We share a common framework that characteristics may proxy for systematic risk exposure. Kelly et al. (2019) and Kim et al. (2021) develop new forms of principal component analysis where intercepts and loadings on extracted latent factors are modelled as fixed, linear functions of characteristics. Lettau and Pelger (2020) develop a generalized form of principal component analysis where factors are chosen jointly to explain both the time-series variation in a panel of assets and the cross-sectional mean returns.

In contrast, our paper reintroduces regressions as an approach for taming the factor zoo, highlighting the complex relationship between characteristics and systematic risk. We consider our approach the covariance analogue of traditional characteristic-expected return, cross-sectional regressions in the style of Fama and MacBeth (1973). While the literature exploring characteristics and average returns has been more balanced in its use of both factor model approaches and cross-sectional regression approaches, the literature on covariances has been heavily and increasingly skewed toward the former. As Cochrane (2011) points out, the regression approach has well-known

[1] For example, Grieser et al. (2020) replicate and review studies of comovement driven by common firm headquarters (Pirinsky and Wang, 2006), similar share prices (Green and Hwang, 2009), common analyst coverage (Chung and Kang, 2016), and common prime brokers (Chung and Kang, 2016).

pitfalls, and care is required in choosing functional forms and focusing on interesting variation in the data, rather than on tiny firms or outliers. Despite these difficulties, by virtue of their flexibility and interpretability, panel regressions are a valuable alternative approach for taming the factor zoo.

2 Motivation

Seminal papers Fama and French (1993) and Daniel and Titman (1997) initiated a huge literature that aims to understand why certain stock characteristics are associated with higher returns. Asset pricing theory suggests that exposure to priced systematic risks must be compensated with positive expected returns. If a characteristic predicts expected returns, it also predicts covariances through loadings on common, volatile risk factors. Fama and French (1993) show certain characteristics like market capitalization and book-to-market may be associated with sensitivity to latent risk factors. Central to their reasoning is the fact that the returns firms with shared characteristics tend to comove.

Daniel and Titman (1997) respond that comovement is weak evidence that characteristics proxy for risk factors. They argue that mimicking portfolios are unlikely to remove factor risk exposure unrelated to the characteristics. They give the example of a latent oil factor. After a string of negative oil shocks, energy companies end up in the value portfolio. This common industry exposure generates the appearance of factor comovement. They show portfolios buying the comovement and hedging the characteristic are not priced, while portfolios buying the characteristic and hedging the comovement produce a spread in average returns (see also Daniel et al. (2020)).

The return generating process underlying our analysis closely follows the process underlying the analysis in Daniel et al. (2020). The only departure is that we allow for an intercept term. Including an intercept term allows for more generality and is common in the most general models of recent empirical papers like Kelly et al. (2019), Kim et al. (2021) and Lettau and Pelger (2020). When the intercept is zero, the model is also consistent with the APT, allowing for both priced and unpriced factors. Let stock i's realized excess returns $R_{i,t}$ be described by the following process:

$$R_{i,t} = \alpha_i + \sum_{h=1}^{N_p} \beta^p_{i,h,t} f_{h,t} + \sum_{h=1}^{N_u} \beta^u_{i,h,t} g_{h,t} + \varepsilon_{i,t}, \tag{1}$$

where f denotes the set of N_p priced factors, g denotes the set of N_u unpriced factors and ε denotes idiosyncratic risk. Finally, $\mathbb{C}ov(\varepsilon_{i,t}, \varepsilon_{j,t}) = 0$ for $i \neq j$. We allow for α_i in order to account for possible returns that are unrelated to risk. For example, α may represent violations of arbitrage pricing due to irrational mispricing or trading frictions. In matrix notation we can simply write

$$R_{i,t} = \alpha_i + \beta_{i,t}' F_t + \varepsilon_{i,t}, \tag{2}$$

where $\beta_{i,t}$ denotes a $N \equiv N_p + N_u$ vector of loadings on the priced and unpriced factors, $F_t = [f_t,\ g_t]'$.

Fama and French (1993) motivate the creation of their multifactor model by stating: "if assets are priced rationally, variables that are related to average returns, such as size and book-to-market equity, must proxy for sensitivity to common (shared and thus undiversifiable) risk factors in returns." Our paper examines implications of this hypothesis within the linear factor model framework described by Equations (1) and (2). If characteristics are proxies for loadings on factors, pairwise covariances will also be functions of characteristics. This is true regardless of the form of vector-valued function β mapping from the space of relevant characteristics $\mathcal{C}$, to loadings on each of the factors that generate realized returns. It is also true whether there are 5, 10 or any arbitrary number, K factors driving returns.

If β and α map characteristics to loadings, $\alpha, \beta : \mathcal{C} \rightarrow \mathbb{R}^N$, we can re-express Equation (2) as

$$R_{i,t} = \alpha(c_{i,t}) + \beta(c_{i,t})' F_t + \varepsilon_{i,t}, \tag{3}$$

If Equation (3) is the true return generating process, then the standard assumption that true errors are uncorrelated, $\mathbb{C}ov(\varepsilon_{i,t}, \varepsilon_{j,t}) = 0$ for $i \neq j$, implies the covariance between returns is given by

$$\mathbb{C}ov\Big(R_{i,t}, R_{j,t}\Big) = \beta_t(c_{i,t})' \Sigma_F \beta_t(c_{j,t}), \tag{4}$$

where Σ_F denotes the covariance matrix of factor returns.

It is important to note that the assumption of uncorrelated errors is not restrictive. Correlation among errors would suggest some common factor is missing from the specified factor model. However, Equation (3) specifies the true underlying factor structure including all priced and unpriced factors. This differs from the scenario commonly found in the literature where the econometrician tests a specific linear model which is (potentially) mispecified or only represents a subset of the underlying factor structure. In that case, time-varying missing factors are attributed to the *model's* error term. This results in correlated model errors even though residuals of the true underlying process are uncorrelated.

Equation (4) shows that when loadings on the true factors are functions of only a characteristic vector, $c_{i,t}$, cross-sectional variation in pairwise covariances depends only on variations in pairs of characteristic vectors $(c_{i,t}, c_{j,t})$. This means that one can model cross-sectional variation in covariances of (R_i, R_j) pairs as a function $\mathcal{H}$, of characteristic vector pairs $(c_{i,t}, c_{j,t})$, such that

$\mathcal{H} : \mathcal{C} \times \mathcal{C} \to \mathbb{R}$. If β is a function of additional predictors one can incorporate these into the function $\mathcal{H}$ as well. Directly estimating $\mathcal{H}$ accommodates an arbitrary number of factors K, arbitrary types of factors and arbitrary functional forms β. No matter which linear process generates returns, observed covariances are generated through loadings on the set of true underlying factors, $F = [f, g]'$. The advantage of directly estimating $\mathcal{H}$ is that it circumvents the need to rely explicitly on factors observed and chosen by the econometrician.

Without loss of generality one can assume the N factors making up F are orthogonal to each other meaning that Σ_F is a diagonal matrix with σ_n^2 on the nth diagonal, $n \in \{1, ..., N\}$.[2] In this case, Equation (4) can be written as

$$\mathbb{C}ov\Big(R_{i,t}, R_{j,t}\Big) = \sum_{n=1}^{N} \sigma_n^2 \beta_t^n(c_{i,t}) \beta_t^n(c_{j,t}), \tag{5}$$

a notational convenience which is often invoked in the literature. Rotating the factors to make them mutually orthogonal does not affect the covariance of returns and therefore does not affect the function $\mathcal{H}$. We therefore refer to the form in Equation (5) throughout the paper.

We note that it is important for our analysis that we examine covariances as opposed to correlations. The objective of the current study is to understand how characteristics are related to the factor structure underlying stock returns. Given our motivation, correlations are an inappropriate measure of comovement because they require dividing by each firm's volatility. Since the idiosyncratic part of volatility likely varies across firms this would confound our object of interest. Pairwise firm correlations may be appropriate in pursuit of other research questions like those found in Chordia and Shivakumar (2002), Ang and Chen (2002), and Lou and Polk (2021), for example.

As an alternative method for taming the growing "factor zoo" a number of recent papers propose new dimension reduction techniques using characteristics and characteristic-sorted portfolios in an attempt to uncover factors generating returns, for example Kelly et al. (2019), Kim et al. (2021), Lettau and Pelger (2020) and Kozak et al. (2020). Many of these papers explicitly impose structure on how characteristics relate to the covariance structure of returns, and then use modified forms of principle components analysis (PCA) to extract factors. They also implicitly impose structure by combining stocks into characteristic-managed portfolios before statistical analysis is performed. This "asset repackaging" can alter the prominence of factors and characteristics' factor exposures. The resultant estimated factor structure may not coincide with the factor structure governing the underlying assets (Shanken, 1982; Gilles and LeRoy, 1991; Bray, 1994).

These PCA and modified PCA approaches reduce dimension by extracting a small number of factors. Individual characteristics often appear in multiple factors and factor loadings, obscuring

[2] One can always find a rotation of factors such that this is true.

the totality of the characteristic's relationship with factor risk. Furthermore, the factors are often difficult to assign economic meaning to. Our approach complements the PCA approach. We reduce dimension at the characteristic level. This allows us to identify the marginal contribution of each characteristic across all latent factors underlying the true model.

Unlike the modified PCA approaches, we do not identify a set of latent factors. While this is a limitation of our approach, it is not generally possible to identify the true economic sources of risk underlying factors of the APT model. Nawalkha (1997) and Lewellen et al. (2010) show that a set of factors that is correlated with the true economic factors and uncorrelated with idiosyncratic risk will price assets just as well as the true factors. A strength of our approach is that we can identify the combined contribution of a characteristic to factor risk across a very general set of linear models. The latent factors could be traded or non-traded, rational or behavioral, macroeconomic shocks (Chen et al., 1986) or characteristic sorted portfolios (Fama and French, 1993). Our flexible model can accommodate a rich variety of mappings between characteristics and factor loadings. Furthermore, we do not need to specify a certain number of factors that must be assumed as in the modified PCA approach. Without knowledge of the true number of factors, the PCA approach may miss important factors for covariance between stocks. Our method reflects the covariance associated with all underlying factors implicitly.

By generating new evidence using the cross-section of covariances that is robust to the many alternative types of latent factors and factor loadings, our approach complements papers that innovate on factor extraction. We estimate exposure to a combination of latent factors as predicted by each individual characteristic. In this sense, our paper is analogous to the large literature that directly models the cross-section of expected returns as functions of firm characteristics. While these studies do not directly identify the SDF, they summarize exposure potentially stemming from unidentified, latent factors. They complement and inform studies that aim to explicitly construct factor models to price the cross-section of average returns. Similarly, we complement the factor extraction approaches by describing the marginal contribution of individual characteristics to factor risk, independent of any narrow assumptions about the underlying factor model.

3 Data

To grapple with the many firm characteristics in the cross-section, we focus on 153 firm characteristics collected and categorized into 13 groups by Jensen et al. (Forthcoming): accruals, debt issuance, investment, leverage, low risk, momentum, profit growth, profitability, quality, seasonality, size, skewness, and value. Jensen et al. (Forthcoming) form the groupings using a hierarchical clustering algorithm over univariate portfolio returns, so that characteristic groups are formed by combining characteristics with a tendency to covary. These groupings allow us to study a comprehensive and diverse set of firm characteristics while maintaining interpretability of the results.

Using the 153 characteristics, we first create composite group variables. Different characteristics may have dramatically different cross-sectional distributions in the data. To make them comparable, we follow the approach developed in Asness et al. (2019). We first transform each characteristic to a vector of ranks. We then standardize each characteristic $z_i = \frac{rank(x_i)-\mu}{\sigma}$ subtracting the cross-sectional mean and dividing by the cross-sectional standard deviation each month. Additionally, we re-sign each variable so that the measures are increasing in their prediction of average returns. Lastly, we combine the characteristics by summing the standardized characteristics in each group to form a composite ranking. If a characteristic is missing, it is not included in the composite ranking. If all the characteristics in a group are missing, the composite ranking is considered missing.

Next, we compute firmwise quarterly covariances for stock pairs. Since we are computing stock-by-stock covariances, we remove the plentiful, but tiny "micro cap" stocks from our sample. These stocks, measured as stocks below the 20th percentile of NYSE market equity, make up only three percent of aggregate market equity. We expect problems arising from non-synchronous trading, stale prices and illiquidity to be most severe in this group. Additionally, to be included in the sample, stocks are required to have non-missing values for each of our 13 composite groupings. We restrict the stock sample to those stocks for which at least 70 percent of the daily returns over the past year are non-missing.

Once we have the set of candidate firms, we divide them into two groups, large and small, by the Fama and French (2008) definitions, where large stocks are larger than the NYSE median stock's market equity and small stocks are between the median and twentieth percentile. We then randomly select 250 firms from each of these two groups at the end of June each year. This collection of randomly selected stocks is then held fixed from July through June of the following year. Each quarter within that period, we compute pairwise covariances of all 500 stocks using daily data within the quarter. This gives us a sample size of roughly 24 million pairwise correlations and covariances. Sampling covariances at the quarterly horizon assures our covariances are aligned with each characteristic's return predictability, even if that return predictability is short-lived.

Our tests use covariances that have been adjusted to robustly account for nonsynchrous trading as described by Shanken (1987), which is similar in spirit to the Dimson (1979) and Cohen et al. (1983) adjustments for betas. The adjustment incorporates the possibility of slow price adjustment for each stock's daily price. Slow adjustment of stock i's price can result in covariance of stock i's returns with lagged returns of stock j provided stock j's price incorporates information more efficiently. The adjustment accounts for the possibility that stock i and stock j may have prices that adjust at different rates. We calculate quarterly adjusted covariances using the following specification for arbitrary stocks i and j:

$$\mathbb{C}ov^{adj}(r_i, r_j) = \mathbb{C}ov(r_i, r_j) + \sum_{m=1}^{m=4} \mathbb{C}ov(r_{i,t}, r_{j,t-m}) + \sum_{m=1}^{m=4} \mathbb{C}ov(r_{i,t-m}, r_{j,t}), \tag{6}$$

where $\mathbb{C}ov(r_i, r_j)$ denotes quarterly covariance of stock i and j calculated using daily returns and $\mathbb{C}ov(r_{i,t}, r_{j,t-m})$ denotes the covariance between returns of stock i and m-day lagged returns of stock j using daily returns within the given quarter. Thus, our covariance estimates contain 5 different temporal pairings: lags of 0,1,2,3 and 4 trading days for each stock in a given i, j pair.

Our stock return data comes from CRSP with corresponding firm characteristics taken from COMPUSTAT. Our sample runs from July 1964 to December 2018. The sample includes only common equity securities (share codes 10 and 11) for firms traded on NYSE, NASDAQ, or AMEX. Additionally, we exclude financial firms (Standard Industry Classification codes of 6000 to 6999). We reproduce the details of the Jensen et al. (Forthcoming) characteristic groups and underlying definitions in the Internet Data Appendix.[3]

4 Univariate Sorts

We start by forming univariate portfolio sorts on each of the thirteen composite characteristics. For each grouping, we present two figures. The first is a sort into quintiles of each characteristic that captures its relationship with average returns. The second figure takes the same quintiles and forms a five by five sort that captures the relationship of the characteristic with firm-level covariances.

Figures 1 and 2 show the average returns and covariance sorts of the first two characteristic groups Accruals and Debt Issuance. As expected, the quintile sorts generate a monotonically increasing sort on average returns. The covariance sorts generate maxima at the extreme portfolios (1,1), (5,5) and (1,5). This pattern at the extremes (1,1) and (5,5) is natural if characteristics create loadings $\beta_i \beta_j$ that are extremely positive or negative in the extreme portfolios. High accrual stocks covary with each other and low accrual stocks covary with each other. But the (1,5) portfolios covary in similar amounts to the (1,1) and (5,5) portfolios. High and low accrual stocks covary with each other as much as they covary with stocks with similar accrual characteristics. The same pattern is seen with Debt Issuance.

We note here that these are univariate relationships. The results may be driven by similarity in the other characteristics of accrual and debt issuance stocks. Stocks in portfolios one and five may share similar exposures to market beta, firm size, or exposure to other factors that drive their

[3]Bryan Kelly provides the WRDS code and additional details of the data construction at `https://www.bryankellyacademic.org/`. We thank the authors for this tremendous service to the profession.

univariate covariances. In the next section, we show that a theme of our results is that univariate and multivariate relationships are very different.

Figure 3 shows the results for the Investment group. While the average return sorts are still monotonic, though strongest in the extreme quintiles, the pattern of the covariances looks very different than the previous two figures. The high investment, low average return stocks in portfolio (1,1) covary a lot, while the low investment, high average returns stocks in portfolio (5,5) covary only slightly more than stocks with neutral loadings in portfolio (3,3). The pattern is consistent across near portfolios (1,2) and (2,1). High investment stocks comove, but low investment stocks do not. This differs from the usual evidence of comovement presented when creating long-short portfolios as factors. Typically it is stocks within the high average return portfolio that are shown to comove most strongly. For example, Fama and French (1993) show that small stocks and value stocks comove more with SMB and HML respectively than large stocks and growth stocks.

Figure 4 shows the results for Leverage. Like all the characteristic groupings, the Leverage characteristic is signed by Jensen et al. (Forthcoming) so that stocks in portfolio five should have high average returns (in this case companies with little leverage), but the Leverage characteristic does not create a spread in average returns, if anything decreasing leverage is associated with slightly lower returns. The characteristic sort does create a prominent spread in covariances. The low leverage (5,5) stocks covary the most and the high leverage (1,1) stocks covary the least. Perhaps surprisingly, given that all else equal more leverage should raise a stock's market beta, stocks with less leverage covary more. The pattern from the low leverage extreme to the high levarage extreme is monotonically decreasing in covariance.

Figure 5 shows that sorts on Low Risk create an average return differential. Since the Low Risk anomaly refers to "low risk" stocks having anomalously higher average returns than "high risk" stocks, the "low risk" stocks are categorized as being in the Low Risk (5) quintile and the "high risk" stocks are in the Low Risk (1) quintile. Consistent with the Low Risk anomaly, Figure 5 shows low returns concentrated in the extreme "high risk" portfolio (Low Risk (1)) with no evidence of monotonically different returns in the other four portfolios. The high risk portfolio also manifests as a high covariance portfolio. Stock-by-stock covariances are highest in portfolio (1,1) and lowest in portfolio (5,5), creating a monotonic sort declining in covariance as the characteristic decreases.

Figure 6 shows the Momentum sorted portfolios. These average return differences are most obvious in the extreme quintiles. The covariances return to a pattern of local maxima at the extremes (1,1), (5,5) and (1,5). But the low momentum stocks in portfolio (1,1) covary more than the high momentum stocks in portfolio (5,5).

Figures 7 and 8 show the results for Profitability and Profit Growth sorts. Profitability looks to be a more robust predictor of average returns as its sort is monotonic and creates a slightly larger spread from the low quintile to the high quintile. However, despite the conceptual similarity of the two categories, the resulting patterns in covariances are very different. The low profit companies in portfolio (1,1) covary the most, while the high profit companies in portfolio (5,5) covary the least. On the other hand, low profit growth companies covary and high profit growth companies covary, and we again see the common pattern of maxima at the (1,1), (1,5) and (5,5) extremes. In Section 5.4 we show that despite the conceptual similarities, high Profitablity and Profit Growth stocks differ substantially in how they tend to overlap with other characteristics used in our paper.

Figures 9 and 10 show sorts on Quality and Seasonality. Again we see close to monotonic quintile sorts on average returns, but the two have very different patterns in covariances. Low quality stocks in portfolio (1,1) covary the most, whereas portfolio pairs of stocks above the second quintile do not covary much. The covariance is concentrated in the low quality stocks. Seasonality shows a more U-shaped sort where stocks in portfolio (1,1), (5,5) and (1,5) covary more than the neutral stocks in portfolio (3,3).

Figure 11 shows the sorts on Size. The difference in average returns across size quintiles is about 50 basis points quarterly. The small spread is partially due to the fact that we've removed microcaps from the sample, about 30% of the typical spread across quintile portfolios, and partially due to the decline in small stock returns in the time period since Banz (1981). The rise in covariances from large stocks in portfolio (1,1) to small stocks in portfolio (5,5) is quite large. Small stocks covary more than large stocks, and this pattern is consistent across size portfolios. Figure 12 shows that Skewness sorts are rather slight in both average returns and covariances. Average returns are slightly increasing from the high skewness (Skewness (1)) to the low skewness (Skewness (5)) and the stocks in extreme skew portfolios covary more than neutral stocks.

Figure 13 shows Value sorted portfolios. Growth to value creates a strong monotonically increasing pattern in average returns. While value stocks have the higher average returns, the growth stocks in portfolio (1,1) covary the most and the pattern is mostly decreasing as stocks move toward value, though a slight uptick suggests that two value stocks tend to covary more than two neutral stocks.

5 Multivariate Analysis

In this section, we move from univariate sorts to a multivariate regression model. In order to capture the diversity of patterns we find in Figures 1 through 13, we employ a flexible, minimally parameterized, regression model. Our goal is to understand how exposure to each characteristic contributes to systematic risk. To this end, we make choices that ensure the transparency and interpretability of our model, while still maintaining the flexibility to explore the patterns revealed in the previous section.

5.1 Multivariate Baseline

We operationalize Figures 1 through 13 by creating indicator functions for each of the 13 characteristics. In essence, our model transforms the covariance figures into a set of indicator functions for each characteristic. Every pair of stocks therefore has 13 indicator functions, each of which describes the pair's quintile assignments for a particular characteristic. For example, if a pair of stocks, i and j is composed of one of the lowest and one of the highest momentum firms, the momentum indicator $\mathbb{1}_{1,5}(Mom_{i,j}) = \mathbb{1}(Mom_i = 1, Mom_j = 5) = 1$. Since the designation i or j is arbitrary, for each characteristic $\mathbb{C}$, we set $\mathbb{1}_{a,b}(\mathbb{C}_{ij}) = \mathbb{1}_{a,b}(\mathbb{C}_{ji})$ so that all 13 indicator functions are symmetric in each of the stocks' characteristic quintiles. More formally we define:

$$\mathbb{1}_{a,b}(\mathbb{C}_{ij}) \equiv \mathbb{1}((\mathbb{C}_i = a, \mathbb{C}_j = b) \cup (\mathbb{C}_i = b, \mathbb{C}_j = a)) \qquad a, b \in \{1, 2, 3, 4, 5\}$$

so that for each characteristic, we have fifteen unique bins describing the quintile pairs. In light of the Daniel and Titman (1997) critique, we control for industry risk exposure by including indicators for each of the 49 Fama French industry classifications. This controls for the possibility that industry exposure underlies covariation among firms with similar characteristics. For each of the industries, the indicator is equal to 1 if both firms whose covariance is observed are classified in the same industry.

$$\mathbb{1}^I_k(Ind_{ij}) = \mathbb{1}(Ind_i = k, Ind_j = k), \qquad k = 1, 2, 3, ...49,$$

where Ind_i and Ind_j denote the industry classification of firms i and j.

A common simplification in modeling multivariate functions is to assume the function is additive in predictors. Similar to Freyberger et al. (2020) who model expected returns as non-parametric functions of characteristics, we assume our function of interest is additive in characteristics. This assumption reduces model complexity and allows for easier interpretation of results. The full baseline model consists of quintile sorts on accruals, debt issuance, investment, leverage, low risk, momentum, profitability, profit growth, quality, seasonality, size, skewness, and value:

$$
\begin{aligned}
\mathbb{C}ov_{ij}^{adj} = \mu + \sum_{a=1}^{5} \sum_{b \geq a,\ a,b \neq 3,3}^{5} \Big[&\phi_{1,ab} \mathbb{1}_{a,b}(Acc_{ij}) + \phi_{2,ab} \mathbb{1}_{a,b}(DIss_{ij}) + \phi_{3,ab} \mathbb{1}_{a,b}(Inv_{ij}) \\
&+ \phi_{4,ab} \mathbb{1}_{a,b}(Lev_{ij}) + \phi_{5,ab} \mathbb{1}_{a,b}(LowRisk_{ij}) + \phi_{6,ab} \mathbb{1}_{a,b}(Mom_{ij}) + \phi_{7,ab} \mathbb{1}_{a,b}(Prof_{ij}) \\
&+ \phi_{8,ab} \mathbb{1}_{a,b}(ProfGr_{ij}) + \phi_{9,ab} \mathbb{1}_{a,b}(Qual_{ij}) + \phi_{10,ab} \mathbb{1}_{a,b}(Seas_{ij}) + \phi_{11,ab} \mathbb{1}_{a,b}(Size_{ij}) \\
&+ \phi_{12,ab} \mathbb{1}_{a,b}(Skew_{ij}) + \phi_{13,ab} \mathbb{1}_{a,b}(Val_{ij}) \Big] + \sum_{k=1}^{49} \phi_k^I \mathbb{1}_k^I (Ind_{ij}) + e_{ij}.
\end{aligned} \tag{7}
$$

While each characteristic is represented as fifteen unique indicator variables, we can only identify fourteen of the fifteen indicators in the multivariate model. For every characteristic, we therefore omit $(a, b) = (3, 3)$ as these will be absorbed into a constant term μ. Since firms in portfolio three have roughly neutral loadings on the characteristic, the constant term represents the pairwise covariance of two hypothetical firms neutral on all characteristics. To ease interpretation of the coefficients, we standardize the lefthand side variable across the full sample, so that the coefficients are rescaled and can be interpreted as measuring standard deviations above (positive) or below (negative) the covariance of a hypothetical pair of firms with all neutral characteristic values (3,3), belonging to different industries.

We include only the 13 characteristic groups from Jensen et al. and same-industry indicators. Same-industry indicators are important to include because covariation of a certain characteristic may actually be due to certain industries having similar characteristic values. This point is emphasized in Daniel and Titman (1997). There are a number of additional predictors that would increase in-sample performance of the model which we do not include. For example, we could include each of the 153 characteristics used by Jensen et al. to create the 13 characteristic groupings. However, the analysis in Jensen et al. shows that the full set of 153 characteristics are largely repetitions of these 13 themes. If we were to include all 153 characteristics, our model would perform better in-sample but the estimates would be difficult to interpret since there would be so much confounding between characteristics within each of the 13 groups.

Similarly, we could include lagged covariances as these are known to predict future covariances. Lagged covariances would improve the predictive power of the model, but since characteristics in our sample are often persistent, the characteristic information we are interested in is embedded in the lagged covariances. The lagged covariances would confound the relationship between characteristics and covariances. Our goal is to understand the relationship between characteristics and systematic risk. Including lagged characteristics would necessarily obscure this relationship.

The dependent variable in our analysis, $\mathbb{C}ov_{ij}^{adj}$ is computed using daily returns. Using adjusted covariance measurements following Shanken (1987) reduces some aspects of noise associated with measured covariances. However, they are still likely to be very imprecise measurements of pairwise covariances. The presence of noise associated with estimated dependent variables causes larger standard errors of estimated regression coefficients but does not bias the point estimates. Larger standard errors are less of a concern for us because the number of quarterly pairwise covariances from our random sample of 500 stocks exceeds 120,000. Our main empirical specifications all have over 20 million observations. Even though we estimate a large number of regression coefficients and we cluster standard errors, we have sufficient power to have informative tests.

Table 1 shows the multivariate coefficients across the thirteen characteristics. Broadly the relationships across the coefficients are consistent moving from the center toward the extremes, so that the coefficients on the extreme groupings (1,1), (5,5) and (1,5) capture the relationships across quintiles. We will use this fact to condense later tables. For the characteristics that we observe significantly predicting firmwise covariances, the predictability is always concentrated on one leg. This stands in stark contrast to the univariate relationships. For example, Figure 1 shows that both low (5,5) and high (1,1) accrual stocks tend to covary, but the coefficients in Table 1 show that the covariance of stocks in accruals (1,1) is absorbed by the tendency for high accrual stocks to have exposure to other characteristics.

Debt issuance, investment, low risk, momentum, profit growth, seasonality, and size are all strong and significant positive predictors of covariances, but all of the predictive power is concentrated on one leg. The leg with strong predictive power is neither consistently the high nor low expected return leg. For instance, firms with low debt issuance, high profit growth, and small size in the (5,5) portfolios covary in the higher than average return legs. While firms with high risk, low momentum, and low seasonality in the (1,1) portfolios covary in the lower average return legs.

If we compare to a setting where factor betas are independent, and identically distributed around zero, we might expect the coefficients to follow a very different pattern:

$$\mathbb{C}ov(R_i, R_j) = \sum_{l=1}^{K} \beta_i \beta_j \sigma_l^2$$

If factors are related to characteristics and factor betas are distributed around zero, we might expect two value stocks to load positively on HML and two growth stocks to load negatively on HML. Thus, value stocks covary with other value stocks and growth stocks covary with other growth stocks. By the same logic, one might expect that growth and value stocks covary less than two average stocks.

Fama and French (2020) codify this natural relationship by introducing a factor model where the factor loadings are set equal to standardized characteristic values and the factors are formed as cross-sectional regression coefficients. In their set up, if a stock has a characteristic one standard deviation above or below the mean, it gets a factor loading of one or negative one, respectively. This pattern does not appear in the coefficients of the multivariate regression. At times, this is very natural. For example, if we think of the Low Risk characteristic, or even more specifically, market beta, it is natural to think of betas as varying across a completely positive range. Frazzini and Pedersen (2014) form a strong sort on ex post betas from ex ante information that generates a span of betas across the low and high deciles from 0.67 to 1.85. It may well be the case than there are no negative market beta stocks. But a common rationale that underlies the high versus low sorts so common in the literature is that these portfolios capture opposing loadings on underlying factors after removing the common exposure all stocks have to the market.

Looking across characteristic groups, some extreme legs actually covary significantly less than two neutral stocks after controlling for exposure to other characteristics. For instance, high skewness stocks in (1,1) which have low average returns covary significantly less than two neutral stocks. Low investment stocks in portfolio (5,5) with higher average returns tend to covary less than average at a 10% level of significance. In the framework above, we would expect stocks at the opposite extremes of characteristics, those in portfolio (1,5) to have betas of opposite sign, suggesting they covary less than two average stocks. We do see this pattern in leverage, momentum, skewness, value, and at 10% levels of significance with low risk, but often there is no clear relationship, and in the case of accruals and debt issuance the relationship is even positive.

Not surprisingly, the largest point estimate in the entire model is the coefficient for the Low-Risk (1,1) portfolio. Stocks in the first quintile of the Low-Risk anomaly are high risk stocks which include high market beta stocks and high volatility stocks. High beta, high volatility stocks tend to covary with each other. Even though this result is not surprising, it is important that we include the low-risk anomaly in our analysis. To determine the marginal contribution of a certain characteristic for the covariance of two firms' stock, we need to control for other characteristics. Clearly, it is important to control for two stocks being in the high-risk portfolio when evaluating whether the contribution of any other characteristics significantly impacts covariances.

5.2 Multivariate versus Univariate

Table 2 compares the multivariate coefficients in Table 1 to univariate coefficients from regressions of our firmwise covariances separately for each characteristic group. This table captures the change in the relationships between characteristics and covariances that results from the impact of controlling for the presence of other characteristics. Since the patterns are quite consistent moving from the center toward the extreme portfolios, we condense the table by showing only the extreme portfolios (1,1), (1,5), and (5,5).

The coefficients are typically smaller as we move from univariate to multivariate characteristics suggesting that much of the univariate relationships between stocks is driven by exposure to other characteristics. This outcome is sometimes extreme. For instance, the univariate regression suggests that the growth stocks in Value (1,1) comove 0.25 standard deviations more than neutral stocks with a t-statistic of 4.50. This relationship is completely absorbed by other characteristics as the multivariate coefficient on value (1,1) is 0.02 with a t-statistic of 0.61. The strong univariate association between low profitability stocks in (1,1) is completely absorbed in the multivariate setting. In a univariate sense, growth stocks and low profitability stocks comove, but they do not comove because they are growth stocks or low profit stocks. They comove because they tend to share similar characteristics that are associated with comovement.

High investment (1,1) remains a strong and significant predictor of covariances, though the coefficient falls by half relative to the univariate regression. While in the univariate case, both high and low profit growth stocks covary at the margins of conventional levels of significance, in the multivariate case all of the association is in the high profit growth leg. These characteristics are of particular interest as they have risen in prominence in new factor models, such as Hou et al. (2015) and Fama and French (2015).

Hou et al. (2015) motivate the characteristics of investment and profitability (and later profit growth Hou et al. (2021)) from the neoclassical q-theory of investment. Fama and French (2015) motivate their model as a version of the dividend discount model. While both approaches motivate the use of characteristics as predictors of average returns, neither approach gives much theoretical insight to how firms should covary. In both approaches the authors move from their motivation to a reduced form factor pricing model with two or three characteristic portfolio sorts. This move requires that the characteristics are associated with expected returns because they are associated with loadings on priced factors. Our approach is able to test this proposition.

We find that profitability is not a marginal predictor of covariance. When firms comove with profitability, it is likely due to common associations they have with other latent factors. We find large differences in univariate and multivariate regressions. Our results suggest the common practice of forming reduced form models from one-diminsional or multi-dimension sorts on only a small number characteristics in the style of Fama and French (1993) may obscure important patterns in the data. The resulting covariation may be largely due to omitted latent factors with loadings correlated with these characteristics.

One of the largest predictors of covariances in the multivariate model is momentum "losers" in portfolio (1,1). Fama and French (2020) call momentum, "a hard sell for a world of rational pricing," but this must not be because they represent an anomaly of arbitrage pricing as outlined by Ross (1976). Momentum is a strong multivariate predictor of covariances. The coefficient on momentum "losers" is larger than the coefficient on small stocks in Size (5,5), and while neither "winners" nor large stocks significantly comove controlling for other characteristics, high and low momentum stocks tend to comove less with each other than neutral stocks.

If a characteristic proxies for factor risk, we would expect firms in characteristic pair (1,5) to covary less than two neutral firms yielding a negative coefficient. We see this pattern arise in the multivariate regression with leverage, momentum, skewness, and value. For each of these four groups, the (1,5) pair is negative and significant in the multivariate, but not the univariate specification. This suggests that confounding relationships with other characteristics tend to obscure the relationships between the characteristics in the extreme groups.

The multivariate and univariate results are so different because firm characteristics share considerable overlap. Table 3 captures the raw overlap across the 13 characteristics groups. For the first and fifth quintiles of each characteristic, we report the percentage of stocks in that quintile that are also in an extreme quintile of another characteristic. Since stocks may overlap with either the low or high end of another characteristic, for example the high leverage stocks in Leverage (1) are more likely to be in the value leg, Value (5), we report the maximum overlap of either the first or fifth quintile of the other characteristic. If the first quintile is matched with the fifth quintile, we bold the number reported. Panel A reports the overlap of the first quintile of each characteristic with the other 12 characteristics, while Panel B reports their overlap with the fifth quintile.

If characteristics were completely independent of one another, we would expect all of the off-diagonal elements to be equal to 20%. Therefore, an overlap of 30% indicates approximately 50% higher overlap than we would expect under a null hypothesis of independence in characteristics. Looking broadly, approximately a quarter of the values across the two panels are above 30%, indicating high overlap with an extreme quintile of another characteristic group. The largest overlaps jump out immediately. Growth stocks in Value (1) have 58% overlap with low leverage (5) stocks, while value stocks in (5) have a 46% overlap with high leverage stocks in (1). Low quality stocks in (1) have 53% overlap with low profit stocks in (1). While these extreme examples capture conceptually closely related characteristics, substantial overlap is quite common. For example, high investment firms (in (1)) also tend to be high accruals (40%), high debt issuance (49%), low leverage (34%), high risk (41%), and growth stocks (47%). Low investment companies (in (5)) overlap with low accruals (39%), low debt issuance (43%), and value (35%), but have weaker overlap with leverage (23%) and low risk (25%) and a stronger overlap with high profitability (31%).

The overlaps are not always as symmetric. The high beta, high volatility stocks in Low Risk (1) have substantial overlap with high investment (41%), low leverage (43%), low momentum (35%), low profit (42%), small size (31%), and growth stocks (50%). On the other hand the low beta and volatility stocks in Low Risk (5) only overlap more than 30% with high leverage (32%) and value stocks (34%). Some groupings are more independent of other characteristics. Seasonality has no overlap with another quintile above 30%. Skewness has its greatest overlap when the short leg (1) has a 30% overlap with the high beta, high volatility stocks in Low Risk (1). Low Momentum (1) stocks have a 35% overlap with Low Profit Growth (1) and 35% with the high risk stocks in Low Risk (1). Winners in Momentum (5) tend to overlap with the large stocks in Size (1).

Stocks share correlated characteristics, and this overlap drives the mapping of characteristics onto covariances. Regressions are a natural tool to tease out the relationships between characteristics and covariances, but while the literature is large and well-developed when the target is expected returns, it is comparatively sparse when the target is expected covariances. Regressions excel at disentangling the complicated relationships between characteristics and covariances.

Our regression results showing that multivariate relationships between characteristics and covariances are much different than univariate relationships suggest that factors underlying the cross-section are not as neatly aligned with characteristics as suggested by the small dimensional portfolio sorts frequently used in the literature. This result seems to favor recent attempts to expand traditional principal component techniques to capture factors common across stocks such as Kelly et al. (2019), Kim et al. (2021), and Lettau and Pelger (2020). The flexibility in these models to find factors across portfolios may better allow them to model the relationships we find. The additional flexibility often occurs with a loss of transparency. It can be difficult to understand the resulting statistical factors. Transparency is a strength of our regression-centric approach. We can disentangle the complicated linkages across characteristics. Additionally, our results suggest that empiricists should be wary of imposing excessive symmetry in relationships across characteristics.

5.3 Isolating Marginal Factor Risk in Each Characteristic

Our approach isolates the marginal predictive power a characteristic has for the covariance between two firms. In this section, we combine that with estimates for the marginal predictive power each characteristic has for expected returns. Combining these two measures allows us to estimate the Sharpe ratio of buying the isolated marginal factor risk associated with each characteristic.

Our model allows us to approximate the volatility of long-short positions that can be attributed only to systematic factor exposure of characteristic h, while controlling for all other characteristics in the model. A position that is long a high characteristic h stock and short a low characteristic h stock but with neutral exposure to other characteristics has variance equal to:

$$\mathbb{V}ar(P_5^h - P_1^h) = \mathbb{V}ar(P_5^h) + \mathbb{V}ar(P_1^h) - 2\mathbb{C}ov(P_5^h, P_1^h), \tag{8}$$

where P_k^h denotes returns of a stock with characteristic h in the kth quintile while having neutral exposures to all other characteristics.

Our model estimates the contribution to covariance between any two stocks that can be isolated as coming from characteristic exposure to systematic factors. Therefore, a stock's isolated variance from a single characteristic's factor exposure can be described by $\mathbb{V}ar(P_5^h) = \mathbb{C}ov(P_5^h, P_5^h)$ and $\mathbb{V}ar(P_1^h) = \mathbb{C}ov(P_1^h, P_1^h)$. Equation (8) can therefore be written as:

$$\mathbb{V}ar(P_5^h - P_1^h) = \mathbb{C}ov(P_5^h, P_5^h) + \mathbb{C}ov(P_1^h, P_1^h) - 2\mathbb{C}ov(P_5^h, P_1^h). \tag{9}$$

The point estimates from our model allow us to estimate the values in Equation 9. The estimated variance of the long short position described in Equation 9 is given by:

$$\widehat{\sigma_h^2} = \widehat{\phi}_{h,55} + \widehat{\phi}_{h,11} - 2\widehat{\phi}_{h,15}$$

We combine these estimates with estimates of the marginal influence of each characteristic on expected returns. Each characteristic's influence on expected returns is estimated with the following regression model,

$$\begin{aligned} Ret_{i,t+1} = \gamma_0 + \sum_{a \in \{1,2,4,5\}} \Big[& \gamma_{1,a}\mathbb{1}_a(Acc_i) + \gamma_{2,a}\mathbb{1}_a(DIss_i) + \gamma_{3,a}\mathbb{1}_a(Inv_i) \\ & + \gamma_{4,a}\mathbb{1}_a(Lev_i) + \gamma_{5,a}\mathbb{1}_a(LowRisk_i) + \gamma_{6,a}\mathbb{1}_a(Mom_i) + \gamma_{7,a}\mathbb{1}_a(Prof_i) \\ & + \gamma_{8,a}\mathbb{1}_a(ProfGr_i) + \gamma_{9,a}\mathbb{1}_a(Qual_i) + \gamma_{10,a}\mathbb{1}_a(Seas_i) + \gamma_{11,a}\mathbb{1}_a(Size_i) \\ & + \gamma_{12,a}\mathbb{1}_a(Skew_i) + \gamma_{13,a}\mathbb{1}_a(Val_i) \Big] + e_i. \end{aligned} \tag{10}$$

where excess returns are regressed on indicators for each characteristic grouping in a panel regression with robust standard errors clustered by firm and time period. The regression captures the marginal contribution of each characteristic in the presence of other characteristics.

Table 4 shows the results of this regression. It shows monotonic sorts for accruals, debt issuance, momentum, quality, size, skewness, and value, which are all monotonically increasing from portfolio one to five as suggested by previous empirical work and the univariate results. In each of these characteristics, at least one extreme quintile's point estimate is significant. Investment is close to monotonic and significant at the 10% level on the short leg. Profitability and seasonality have monotonic patterns in the coefficients, but they are not statistically different than zero. There is not a consistent monotonic pattern in leverage, low risk, or profit growth, suggesting these are not robust marginal predictors of expected returns.

The coefficients in Tables 1 and 2 show the marginal contribution of each characteristic to covariances across firms. The coefficients in Table 4 yield the marginal contribution each characteristic makes to expected returns. Table 5 combines these two estimates. The first column in Table 5 shows the characteristic groups sorted by magnitude of the volatility of implied factor exposure as described by Equation 9. Low risk, momentum, and value are the most associated with factor volatility, followed by size, leverage, and investment, and then followed by seasonality and profit growth. All of these have volatilities significantly different than zero at the 1% level. Profitability and skewness follow at the margins of statistical significance. The characteristics least associated with volatility are debt issuance, quality, and accruals. Debt issuance's estimated covariance-implied variance is slightly negative, but since that is not possible the estimate is displayed at its lower bound of 0.

The third column shows the average return estimates for each characteristic. Starting at the bottom with the low volatility factors is particularly interesting. Even though debt issuance is not associated with factor exposure, it does predict average returns. Debt issuance behaves like a statistical arbitrage, an APT anomaly. The estimated Sharpe ratio is effectively infinite (though our estimate is subject to estimation error). Next, quality, with its expected return of 1.62% and its modest volatility of 0.93% has a Sharpe ratio of 1.74. This is followed by accruals with an average return of 1.15% and volatility of 1.38% yielding an implied Sharpe ratio of 0.83. The arbitrage pricing theory is built around the idea that Sharpe ratios cannot get too high. We do not attempt to define a strict threshold determining arbitrage or near-arbitrage opportunities, but we regard these Sharpe ratios as suspiciously high. These characteristics are least suitable for risk-based explanations.

At the top of the table, we see that the Low Risk anomalies, which are good covariance predictors, are not good marginal predictors of average returns in our sample. This result is consistent with Novy-Marx (2014), which shows that strategies on beta and idiosyncratic volatility are absorbed by exposure to size, value and profitability. In our framework, the Low Risk stocks behave like an unpriced factor. It is associated with a lot of factor volatility, but also a 0.00 Sharpe ratio.

The next three characteristics all have similar Sharpe ratios between 0.27 and 0.35. Again, momentum is a surprisingly a good candidate for a risk explanation. It has the second highest implied factor volatility and a Sharpe ratio of 0.27. Both value and size have slightly lower average returns and slightly lower volatilities leading to similar overall Sharpe ratios. Leverage is associated with a lot of volatility, but low and insignificant average returns, again symptomatic of an unpriced factor. Investment and seasonality are statistically insignificant predictors of average returns, leading to smaller Sharpe ratios of 0.20 and 0.13, at most on the margin of being associated with priced factors. Profitability has a marginally significant volatility and an insignificant average return. Lastly, skewness has a significant average return of 0.66 and a small and marginally significant volatility of 1.70%. The resulting Sharpe ratio of 0.39 suggests the reward to the small volatility risk is not as outsized as the debt issuance, quality, and accruals factors, leaving it as a modest near-arbitrage or a small risk, small reward factor.

Figure 14 shows this result graphically. On the y-axis we graph each isolated factor's average return, while on the x-axis we graph each isolated factor's covariance. When buying a characteristic is a statistical arbitrage opportunity, the characteristic will hug the left-hand side of the graph, near the y-axis having high average returns, but only implying a little factor volatility. If a characteristic behaves like an unpriced factor, it will be near the x-axis, since it will imply a large amount of volatility, but only a small amount of average return. Finally, the priced factors will balance the risk-return trade-off. For reference, we fit a dashed line from the origin through our most obvious candidates for priced factors, the momentum, value, and size characteristics. The slope of the line suggests an implied Sharpe ratio of 0.29. Buying characteristics along this line involves a risk-return trade-off. Despite different volatilities, several of the factors bunch at around a similar implied Sharpe ratio.

5.4 Isolating Expected Returns

Next, we explore the implications of our regressions for the stochastic discount factor representation of the APT. In order to do so, we first introduce the notion of a single characteristic encompassing expected returns for a given stock. Without loss of generality, we can rotate the factors in Equation 1, so that only one factor has a non-zero risk premium and all other factors are unpriced (Roll (1977), Hansen and Richard (1987)).

$$r_{i,t} = \alpha_i + \beta_{i,t}^{sdf} f_t^{sdf} + \sum_{h=1}^{N-1} \beta_{i,h,t}^{u} \widetilde{g}_{h,t} + \varepsilon_{i,t}, \tag{11}$$

The single factor is a mean-variance efficient portfolio of all investable assets. In the one factor representation, only a stock's β^{sdf} and the mean-variance efficient portfolio's premium, λ_{mve} determine expected returns:

$$E[r_{i,t}] = E_{i,t} = \lambda_{mve} \beta_{i,t}^{sdf}. \tag{12}$$

Therefore, at time t, if we could observe the β^{sdf} of all firms, we could consider it an additional characteristic that completely determines the relative expected returns of stocks. Including it in our model of covariances would allow us to estimate the amount of covariation that comes from exposure to the stochastic discount factor. The remaining covariation of stocks is due to unpriced risk.

In this section we ask, of all the covariation across the cross-section of stocks, how much is due to exposure to the stochastic discount factor compared to the other unpriced factors? In the paradigm since Sharpe (1963) extended Markowitz (1952) into a one factor model of covariances as a precursor to development of the CAPM, there has been a conception that the most important determinants of the covariance of stocks are the most important determinants of equilibrium asset pricing. This intuition was an important motivator of the reimagination of the CAPM in Ross (1976).[4] The market factor describes much of the comovement across stocks. If the stochastic discount factor is strongly correlated with the market factor, it should also describe a lot of the comovement across stocks.

This intuition suggests a new test for how much of the overall explainable comovement across stocks is explained by exposure to the stochastic discount factor. From Equation 12, we can see that exposure to the SDF is proportional to expected returns. If we observed expected returns, we would first sort into quintiles on expected returns and then include expected return sorted portfolios in our cross-sectional covariance regressions. Since expected returns are unobservable, we proxy for expected returns using the characteristic regression model described in Equation 19 and shown in Table 4.

We use the predicted expected return, $\widehat{Ret}$, from these regressions as a proxy for the unobserved true expected excesss returns. These are full sample estimates, so they are not tradeable, but since they use the entire sample are more efficient than expanding window regressions and therefore are the better proxies for expected returns. We group the proxies as we would any characteristic dividing each stock into one of five quintiles and grouping covariance pairs into fifteen groups. We first perform the univariate regression of covariances on the $\widehat{Ret}$ quintile pairs alone.

Table 6 shows the univariate regression results. The upper left corner of Table 6 shows that low expected return stocks in ERet (1,1) covary the most, and this is monotonically decreasing as we move toward ERet (1,5) and rapidly decreasing as we move toward ERet (5,5). Naively, these results seem to suggest that low expected return stocks have large positive or negative exposure to the stochastic discount factor. Large positive exposures are not consistent with these stocks having relatively low expected returns in the cross-section. Large negative exposures would suggest that they hedge risk exposure and should earn a return below the risk-free rate, again, inconsistent with the data. A more likely explanation is that the issue of confounding is quite severe for expected return sorted portfolios. If the loadings on the mean-variance efficient stochastic discount factor in equation 11 are correlated with the loadings of the other factors, then the loadings on the other

[4]Speaking about the CAPM in an AFA interview with Richard Roll, Stephen Ross said, "The words and the music didn't fit, the math didn't fit the words, the intuitions were much better than the mathematics that described it, so the APT came from my attempt to understand what was really going on, and sort of make the math consistent with what I thought the good intuitions were in the field."

factors may dominate the overall patterns of covariances.

Next, the indicators from Equation (7) are added to the univariate indicators for $\widehat{Ret}$ quintile pairs. Table 7 reports the results for the multivariate regression. The pattern of covariances across expected return sorted portfolios reverses, suggesting that the odd pattern in the univariate results was in fact due to confounding exposure to other factors. High expected return stocks comove with a coefficient of 0.04 and marginally significant t-statistic of 1.73, but this evidence is buttressed by the covariance of the (4,5) and (4,4) portfolios (not shown) that have coefficients of 0.03 and significant t-statistics of 2.02 and 2.38, respectively. When other characteristics are included, controlling for the presence of unpriced factors, the high expected return stocks do comove as predicted by theory. The covariances monotonically decrease as they approach the low expected return stocks in ERet (1,1), which has a coefficient of -0.03 that is not significantly different than zero.

Table 6 and Table 7 together demonstrate a fact about the importance of the SDF for explaining covariances. In Table 6 which only considers the expected return proxy, the R-squared is 0.02%, whereas the multivariate regression including the extended set of characteristics is 2.69%. Of the explainable variance, comparatively little is explained by exposure to the mean-variance efficient factor. It follows that comparatively little is explained by exposure to the stochastic discount factor.

At the bottom of Table 7 we push further along this dimension. Explanatory variables are added incrementally to maximize explanatory power. This captures how much of the total covariance explainable by our characteristics is attributable to exposure to the SDF. In the second row, to capture that firms in similar industries may also covary, we create a variable that captures across industry variation. This variable uses the Fama and French (1997) definitions for ten industries and creates indicators from the ten by ten matrix of across industry interactions. If firms in two different industries tend to covary, these indicators will capture that variation. The R-squared from adding across industry interactions rises to 3.17%.

Lastly, we add lagged covariance over the last year. We do not include lagged covariances in the main specification because they are functions of lagged values of the characteristics and thus confound the relationships between characteristics and covariances that we seek to uncover. In a purely predictive exercise, lagged covariances capture past values of our modeled characteristics and factors. In addition, they capture characteristics, factors, and exposures that may yet be unknown. This specification gives us insight into the maximal amount of explainable variance in the cross-section of covariance. However, it requires that we relax the main goal of our analysis, that results must be transparent and interpretable. Lagged covariances increase the explainable R-squareds to 6.03%. There is considerable additional predictable variation in lagged covariances, and it is almost as large as the variation we understand through a broad cross-section of characteristics. This suggests there may be many more characteristics associated with covariances that are not well captured by our thirteen characteristic groups.

Provided our expected return proxies are reasonable approximations of the underlying expected returns, our analysis suggests that of all the predictable variation in future covariances, only a small portion comes from variation in the SDF. These results do not rule out rational pricing or factor pricing. For example, the consumption CAPM could hold with expected returns across stocks proportional to consumption betas, while beta on consumption growth explained only a small portion of the explainable covariance across stocks. This also does not rule out multi-factor models that include the market. However, it does imply that the market accounts for only a small part of the variation in the SDF. This finding does seem to rule out SDFs that are close to the original hypothesis of Sharpe (1964) in explaining a large portion of the covariance across stocks. Further, the evidence suggests that the SDF is neither closely correlated with the market nor any portfolio that explains a large portion of the covariance across individual stocks, which is a much stronger statement than the CAPM does not hold.

Lopez-Lira and Roussanov (2020) construct zero-cost portfolios designed to maximally hedge ex ante systematic risk. They show that their zero-cost, factor hedged portfolios do not have zero returns. They argue this is devastating for the APT. Our results suggest an alternative explanation. If our estimates of expected return are close approximations of the true expected returns, then one can hedge almost all factor risk, without hedging any exposure to the stochastic discount factor. Expected returns represent only a small portion of explainable variation in the cross-section of covariances. As the single factor rotation in Equation 9 makes clear, while expected returns can be reduced to exposure to a single factor, there is no requirement that this single factor explain a large portion of the covariation across stocks. Our results provide direct evidence that in fact, it does not explain much of variation in covariances. Our best predictors of expected returns, combined into a single characteristic, are only weak predictors of the cross-section of covariances.

This finding also relates to the growing literature on weak factors for which the dispersion of risk exposures is small in the cross-section.[5] Since only a small fraction of the spread in covariances across individual stocks is related to their expected returns, our results suggest that even the ex ante mean-variance efficient combination of all latent factors is potentially a weak factor in individual stock returns.

[5]See for example Kan and Zhang (1999), Kleibergen (2009), Bryzgalova (2015), Burnside (2016), Gospodinov et al. (2017), Anatolyev and Mikusheva (2021), Giglio and Xiu (2021), and Giglio et al. (2021).

6 Time Variation in Characteristics-Covariance relation

Our framework so far has been agnostic about the source of risk underlying the latent factor structure of returns. Our APT motivation makes no distinction about the economic processes driving the factors. A natural question is whether these systematic factors line up with standard notions of economic risk. In this section, we examine how the relationship between covariances and characteristics varies in response to the macroeconomic environment.

Any characteristic that contributes more to a portfolio's systematic risk in times when risk premia are high should be associated with higher risk premiums. We therefore use economic state variables commonly known to covary with equity risk premiums as conditioning variables. We interact our indicator model with these state variables. This approach yields a conditional model, allowing us to estimate how risk exposure associated with a characteristic varies across economic states. The conditional relationship offers further insights into how the cross-section of characteristics contribute to cross-sectional risk premiums.

We use three state variables: an indicator for NBER's recession, Martin's (2017) proxy for the time-varying equity premium, and Baker and Wurgler's (2006) measure of market-wide sentiment. Recessions are a natural measure of macroeconomic distress. The equity premium proxy is a continuous measure that captures a host of risks relevant to the marginal investor. And market-wide investor sentiment encapsulates behavioral or psychological phenomena associated with investor behavior. We ask whether the systematic risks that are reflected in characteristics vary dynamically with these measures.

To explore the link between characteristics, pairwise stock-level covariances and economic state variables, we first revisit the stochastic discount factor implied by a linear pricing model:

$$M_t = a + \sum_{n \in N_p} b_n f_{n,t}, \tag{13}$$

where N_p denotes the subset of factors from Equation 2 which are priced. For each factor, b_n represents the associated price of risk. The well-known pricing relation then follows:

$$
\begin{aligned}
\mathbb{E}_t(R_{i,t+1}) - R_f &= -R_f \mathbb{C}ov_t(M_{t+1}, R_{i,t+1}) \\
&= -R_f \sum_{n \in N_p} b_n \mathbb{C}ov_t(f_{n,t+1}, R_{i,t+1}) \\
&= -R_f \sum_{n \in N_p} b_n \beta_{i,n,t} \sigma^2_{n,t} \\
&= -R_f \sum_{n \in N_p} b_n \beta^n_t(c_{i,t}) \sigma^2_{n,t}
\end{aligned} \tag{14}
$$

where $\beta_{i,n,t}$ denotes the time t loading of stock i on factor n and $\sigma^2_{n,t}$ denotes the time t variance of factor n. In the final line of Equation 14 we write $\beta_{i,n,t}$ as $\beta^n_t(c_{i,t})$ where $c_{i,t}$ denotes a vector of firm i's characteristics in order to highlight the hypothesis we are analyzing.

Equivalently, we can use the single-factor representation of the SDF described in Equations 11 and 12:

$$
M_t = a + b^{sdf} f^{sdf}_t. \tag{15}
$$

The pricing relation then becomes:

$$
\begin{aligned}
\mathbb{E}_t(R_{i,t+1}) - R_f &= -R_f \mathbb{C}ov_t(M_{t+1}, R_{i,t+1}) \\
&= -R_f b^{sdf} \mathbb{C}ov_t(f^{sdf}_{t+1}, R_{i,t+1}) \\
&= -R_f b^{sdf} \beta^{sdf}_{i,t}(c_{i,t}) \sigma^2_{sdf,t}
\end{aligned} \tag{16}
$$

Equation 14 shows that the risk premium for a specific stock varies with: priced factor variance $\sigma^2_{n,t}$ or sensitivities to those variances, $\beta_{i,n,t}$. Similarly, for the one-factor rotation, Equation (16) shows that the risk premium for a stock varies with $\sigma^2_{sdf,t}$, or β^{sdf}. If a particular economic variable coincides with variation in a stock's conditional expected returns, it must also coincide with variation in $\sigma^2_{sdf,t}$ or $\beta^{sdf}_{i,t}$.

Revisiting Equation 5, we can break up the summation describing covariance into the sum over N_p priced factors and a separate summation over N_u unpriced factors:

$$
\begin{aligned}
Cov\Big(r_{i,t}(c_{i,t}), r_{j,t}(c_{j,t})\Big) &= \sum_{n=1}^{N} \sigma^2_n \beta^n_t(c_{i,t}) \beta^n_t(c_{j,t}), \\
&= \sum_{h \in N_p} \sigma^2_h \beta^h_t(c_{i,t}) \beta^h_t(c_{j,t}) + \sum_{m \in N_u} \sigma^2_m \beta^m_t(c_{i,t}) \beta^m_t(c_{j,t}).
\end{aligned} \tag{17}
$$

For any $n \in (N_p \cup N_u)$, variation in $\sigma^2_{n,t}$, $\beta_{i,n,t}$ or $\beta_{j,n,t}$ results in variation of the conditional covariance between stocks i and j. Combining this with Equation 14, shows that time-varying risk premiums for either stock i or stock j coincide with time variation of the summation over the N_p priced factors in Equation 17. Holding all else equal, time variation in either stocks' equity premium results in time variation in their covariance. We therefore choose state variables known to covary with factor risk premiums. Interacting the indicators from Equation (7) with macroeconomic state variables allows us to see how the product of loadings and corresponding factor variance, as functions of characteristics, change with the economy. Insofar as the state variables generate cross-sectional variation in risk premiums, the interactions show us which characteristics tend to contribute differently to systematic risk when risk premiums are high.

As is clear from Equation 5, increases in time-varying covariance can either be due to two stocks exposed to the macroeconomic risk or that hedge the macroeconomic risk. That is, because $\sigma^2_n \beta^n_t(c_{i,t})\beta^n_t(c_{j,t}) = \sigma^2_n(-\beta^n_t(c_{i,t}))(-\beta^n_t(c_{j,t}))$, when we observe an increase in covariances, we cannot distinguish whether $\sigma^2_n \beta^n_t(c_{i,t})$ is becoming more positive or more negative. Either can be consistent with time-varying covariances in response to SDF shocks. If firm characteristics proxy for latent factors exposed to time-varying economic state variables, exposure could appear on either the high (5) or the low (1) expected return leg.

In order to examine how characteristics interact with exposures to economic state variables $\mathbb{ES}_t$, we interact state variables with the indicators from Equation (7):

$$\begin{aligned}\mathbb{C}ov^{adj}_{ij} = \mu_0 + \phi^{ES}\mathbb{ES}_t + \sum_{z=1}^{13}\left[\sum_{a=1}^{5}\sum_{b\geq a,\ a,b\neq 3,3}^{5} \phi_{z,ab}\mathbb{1}_{a,b}(c^z_{ij}) + \phi^{ES}_{z,ab}\mathbb{ES}_t * \mathbb{1}_{a,b}(c^z_{ij})\right] \\ + \sum_{k=1}^{49}\theta_k \mathbb{1}^I_k(Ind_{ij}) + \sum_{k=1}^{49}\theta^{ES}_k\mathbb{ES}_t * \mathbb{1}^I_k(Ind_{ij}) + e_{i,j},\end{aligned} \tag{18}$$

where c^z_{ij} denotes the characteristic z quintile pair of firms i and j. Importantly, because the cross-section of covariances we analyze is so large (nearly 24 million observations), we are able to estimate the conditional model with relatively high statistical power.

A number of studies use recessions as proxies for deterioration of the SDF and study stock returns over the business cycle.[6] Additionally, several papers have looked at characteristic sorted

[6]Some examples include, for expected stock returns Fama and French (1989), Ferson and Harvey (1991), Campbell and Diebold (2009); expected Sharpe ratios Brandt and Kang (2004), Ludvigson and Ng (2007), Lustig and Verdelhan (2012).

portfolios over the business cycle to look for evidence of priced exposure to macroeconomic risk.[7] Most closely related to our setting, Moskowitz (2003) develops a multivariate GARCH to study time-varying covariances of size, book-to-market, and momentum portfolios and finds only size robustly responds to the businesss cycle. As shown earlier, these univariate characteristic sorted portfolios are exposed simultaneously to many characteristics. Our approach isolates the marginal predictive power of each characteristic.

We use a quarterly NBER recesssion indicator as our first measure of macroeconomic distress. As in Equation 18, we add the variable linearly to our baseline model and interact it with every characteristic pair as well as the same-industry indicators. The interaction between same-industry indicators and recession indicators controls for cyclical industry variation. The coefficient on the recession indicator estimates the average effect of recessions on covariances. The coefficients on interactions between the recession indicator and characteristic pairs capture the changes to characteristic-covariance relationships during recessions.

Table 8 reports results for the extreme quintiles. In order to conserve space, we do not report estimates for the industry controls. The recession indicator alone is positive but statistically insignificant, indicating that the covariance of a stock pair with neutral loadings on the characteristics does not significantly increase in recessions. Interestingly, this finding is different than the average effect of recessions on the covariance across stocks. In results not shown, we add only a recession indicator without the characteristic interactions. The coefficient on the indicator is 0.44 with a t-statistic of 2.70, suggesting that the average increase in covariances is quite large and significant. When we interact the recession indicator with the characteristics, the coefficient on the indicator term falls to 0.10 and an insignificant t-statistic of 1.44. This indicates that stocks covary more in recession, but this covariation is mediated through the characteristics. Stocks with neutral or irrelevant characteristic exposure do not covary more during recessions, but the typical stock has characteristic exposures associated with more covariation in recessions. These characteristics may load on factors that become more volatile in recessions or the characteristics may load more heavily on the underlying factors during recessions.

Many of the extreme characteristic pairs have very large and significant coefficient estimates when interacted with the recession indicator. Similar to the asymmetry of covariances among extreme characteristic pairs documented in Table 2, we see asymmetry in recessions. That is, most of the characteristics that show large, positive coefficients on the recession interaction term, only have large coefficients for one of the extreme quintile pairs. We see that high risk firms (Low-

[7] See for example studies of size (Perez-Quiros and Timmermann, 2000), book-to-market (Lakonishok et al., 1994; Liew and Vassalou, 2000), momentum (Chordia and Shivakumar, 2002) and a large collection of characteristics (Dittmar and Lundblad, 2017).

risk(1,1)), high investment firms (Investment(1,1)), low past return firms (Momentum(1,1)), high quality firms (Quality(5,5)) all covary more in recessions. Additionally, value stocks (Value(5,5)) and low accrual stocks (Accruals(5,5)) covary more in recessions at the 10% level of significance with very large point estimates. In all of these characteristics, the increase in covariation during recessions is concentrated in one side of the characteristic. While the Value, Accruals and Quality characteristics show increased covariances in the legs typically associated with higher average returns, Low-Risk, Momentum and Investment show higher covariation in the legs associated with lower average returns. We emphasize here that given our dependent variable is pairwise covariance, we cannot distinguish between characteristics that hedge the increased risk during recessions and those that provide exposure to the risk. However, holding all else equal, assets that hedge exposure to risk should be associated with lower risk premiums and assets that provide exposure should be associated with higher risk premiums. This suggests that if the average returns of high investment, high-risk, and low momentum firms are due to risk, then these types of stocks are likely to hedge recession risk. Similarly, high-quality, low accruals and value stocks are likely exposed to recession risk.

Recall that in Table 1 momentum was one of the strongest predictors of covariances with the predictive power concentrated on the low expected return leg (Momentum (1,1)). Table 8 shows that all of that predictive power is concentrated in recessions. In non-recessionary times, high expected return, high momentum stocks in Momentum (5,5) covary more, while low momentum stocks do not covary significantly more than two stocks with neutral loadings. The pattern reverses in recessions, low momentum stocks covary more, while the covariance of high momentum stocks falls in recession.

The size of the effects are large. Low momentum stocks covariation increases from 0.03 to 0.41, which is about the size of the unconditional coefficient on high risk stocks, the largest predictor of covariances in Table 1. Consistent with our finding, Liu and Zhang (2008) ties momentum exposure to industrial production, and Li and Zhang (2017) ties momentum exposure to consumption risk in univariate momentum portfolios. Our approach finds strong business cycle effects in the marginal exposure of momentum to factor risk.

Table 8 also shows a large effect in high investment stocks (Investment (1,1)). The multivariate effect from Table 1 that high investment stocks covary more is entirely found in recessions in Table 8. The high risk stocks in Low Risk (1,1) covary more in recessions with a marginal increase of 0.40 doubling their unconditional coefficient of 0.40. This result extends the conditional CAPM literature (e.g. Jagannathan and Wang (1996), Cederburg and O'Doherty (2016)) into our multifactor setting with potentially many latent factors.

Value has a large but marginally significant increase in the long leg (Value (5,5)), while the growth stocks in the short leg covary more outside of recessions. Low accrual stocks and high quality stocks do not covary significantly in the unconditional estimation, but both covary considerably more in recessions. Investors exposed to their long legs will see large and significant increases in their portfolio volatilities during recessions.

While average returns are notoriously difficult to pin down and require large samples, average covariances converge to their true value much faster. Using covariances alone, in Table 8, we provide strong evidence that covariances of some low-expected return stocks, high-risk, low momentum, and high investment stocks, increase during recessions. These strong effects in the short legs suggest these characteristics must hedge macroeconomic risk in order to rationalize their lower unconditional average returns. Since recessions occur in only 14% of quarters in our data, it is difficult to provide strong statistical evidence for marginal average returns of such stocks increasing in recessions and thus hedging business cycle risk. While we cannot make strong statistical claims about the conditional average returns, Table 9 provides some evidence for this interpretation.

Table 9 shows the results of a regression of returns on characteristics, where the returns are also interacted with the recession state variable.

$$Ret_{i,t+1} = \gamma_0 + \sum_{z=1}^{13} \left[\sum_{a=\in\{1,2,4,5\}} \gamma_{z,a} \mathbb{1}_a(c_i^z) + \sum_{a=\in\{1,2,4,5\}} \gamma_{z,a}^R \mathbb{1}(Rec) * \mathbb{1}_a(c_i^z) \right] + e_i. \tag{19}$$

The table shows that low momentum stocks, which on average have low returns, have high returns in recessions, while high momentum stocks have low average returns. This result is only suggestive of the plausibility that momentum might be hedging exposure to a macroeconomic state variable, because while the effect is large in economic magnitude, 2.62% per quarter, it is not statistically significant.[8] Table 9 shows the same pattern across the investment characteristic. High investment stocks in quintile one have had returns in recessions that were 1.20% per quarter higher than low investment stocks in quintile 5.

Table 9 is merely suggestive as the test has low statistical power. None of the five minus one interacted coefficients are statistically significant, despite sometimes large magnitudes. However, this suggested evidence paired with the low unconditional average returns and the large and statistically significant covariances of these stocks in recessions suggests a risk-based explanation for the average returns. Namely, the stocks hedge business cycle risk.

Next, we interact the model with the estimated equity premium extracted from option prices.

[8] The p-value of the difference between the interacted coefficients of the low and high momentum is 0.26. The p-value of the difference between the high and low coefficients in normal times relative to recessionary times is 0.06, at the boarder of conventional significance thresholds.

We estimate the lower bound on the market risk premium using risk-neutral market variance multiplied by the risk-free rate. Martin (2017) shows that this is a lower bound for the conditional market risk premium and argues that it can be used as a proxy for the premium since the bound is tight. We use the estimated market risk premium, as a single systematic variable that largely captures time varying risk premia of individual stocks which appear on the left side of Equation 14.

Risk-neutral variance is typically calculated using Option Metrics data which goes back only as far as 1996. To extend the data set, we use the NVIX index of Manela and Moreira (2017) which uses natural language processing to extrapolate the VIX index as far back as 1926. We square the index and multiply it by the risk-free rate to get a longer time series of forward looking lower bounds on the market expected excess returns. The data runs through 2016 so that it covers most of our sample.

As before, we interact the equity premium lower bound and the same-industry indicator with all characteristic pair indicators from Equation (7) as well as adding both linearly in the model. We standardize the equity premium proxy in order to make interpretation easier. The results are reported for extreme quintile pairs in Table 10. Again, we do not report estimates for the same industry dummies or their interactions with the equity premium proxy. The equity premium proxy is a large and significant predictor of covariances. Average covariances rise 0.16 standard deviations with a one standard deviation increase in the proxy. In results not shown, when we add the equity premium proxy without interacting it with the characteristics, the coefficient is 0.29 with a t-statistic of 4.99. Just over 40% of the average effect is mediated through the characteristics.

The results are largely consistent with those in Table 8. High-risk firms (Low Risk (1,1)) and past loser firms (Momentum (1,1)) all have large positive and significant coefficients on their interactions with the equity premium proxy. Comovement in value stocks (Value (5,5)), which was at the margins of statistical significance in Table 8, is significant at the 5% level in Table 10. Low accrual stocks (Accruals (5,5)) remains at the margins of statistical significance and is joined by quality which is also significant at the 10% level. The lone exception is small firms (Size (5,5)), which interact significantly and positively with the equity premium proxy, whereas they did not significantly interact with the recession indicator.

In summary, our results suggest that many characteristics proxy for covariances and this covariation is time-varying in conjunction with the business cycle. During times of macroeconomic stress such as recessions or spikes in the estimated equity premium, low accruals, high investment, high risk, low momentum, small, and value stocks are all associated with increases in covariation. These increases are often large enough to drive the unconditional patterns in the cross-section of covariances.

While our results suggest that covariances vary over time in accordance with the business cycle, it is natural to ask whether covariation also arises in response to non-fundamental drivers such as market-wide sentiment. Baker and Wurgler (2006) show that sentiment has cross-sectional implications for stock returns, even when it has been orthogonalized to fundamental economic shocks. They find that returns of characteristic-sorted long-short portfolios that are sensitive to speculative demand are predicted by beginning of period sentiment. Distressed stocks, growth stocks, small stocks, high volatility stocks, young stocks and unprofitable stocks are found to be particularly sensitive to market-wide sentiment.

Under mild assumptions, Kozak et al. (2018) show that even in a world where all deviations from the CAPM are driven by irrational belief distortions of sentiment traders, cross-sectional differences in expected returns must align with the common factor structure of returns. Their model crucially does not require a strong factor structure in biased beliefs, rather it implies that arbitrageurs' activities ensure that only belief distortions aligned with loadings on major common factors can have price effects. In other words Equations (14) and (16) still hold in such a model. Since Baker and Wurgler (2006) show that expected returns of certain stocks vary with market-wide sentiment, Equations (14) and (16) imply that $\sigma^2_{sdf,t}$ or $\beta^{sdf}_{i,t}$ must vary with sentiment. As discussed previously, for stocks i and j, if $\sigma^2_{sdf,t}$, $\beta^{sdf}_{i,t}$ or $\beta^{sdf}_{j,t}$ vary with sentiment, then the covariances between them vary with sentiment. To investigate the relation between sentiment and covariances, we interact the orthogonalized sentiment measure developed by Baker and Wurgler (2006) with our characteristic indicators as in Equation 18.

Table 11 shows the results of interacting sentiment with the indicator model. The largest and arguably most interesting result is the large and significant interaction between market-wide sentiment and the growth portfolio, Value (1,1). The coefficient suggests that the covariation of two growth stocks increases by 0.11 (with a t-statistic of 2.73) when investor sentiment is one standard deviation above average, while the coefficient on the linear Value (1,1) term is insignificant and close to zero. This suggests that covariation among growth stocks is driven by high sentiment market environments.

We suppress the industry controls to save space, but we do note large and significant interactions with sentiment in four industries: Electrical Equipment (0.50, t-statistic of 2.68), Computer Software (0.53, t-statistic of 2.50), Computer Hardware (0.42, t-statistic of 2.10), and Shipping Containers (0.05, , t-statistic of 2.32). While industries naturally have a fundamental source of covariance, these results suggest certain industries have an additional source of sentiment related covariance, perhaps due to their ability to capture the imagination of sentiment investors.

While we see large effects from sentiment in growth firms and in specific high-tech industries, it is arguably as interesting that we do not see effects across many other characteristic groups.

While covariation across firms in many characteristic groups is sensitive to the business cycle and specifically spiked in bad times among firms with low accruals, high investment, high risk, low momentum and low prices (value), the influence of sentiment is isolated to growth firms and industries. Rather than competing to explain similar sensitivities to time-varying macroeconomic factors, sentiment and business cycle effects combine to explain complementary phenomena in the cross-section of covariances.

In summary, our results suggest that many characteristics proxy for covariances and this covariation varies over time in conjunction with both the business cycle and investor sentiment. However, there are many more instances of characteristics whose covariances align with the business cycle than of those that vary with sentiment. Our approach allows us to isolate these time-varying relationships by controlling for the confounding presence of other related characteristics as well as industry effects. While conditional regressions over the cross-section of stock returns have low statistical power and require many more years of data to bound reasonable estimates of time-varying risk premia, our conditional regressions over the cross-section of covariances find large and significant patterns across characteristics that are robust across different measures of macroeconomic activity. Our analysis suggests that the cross-section of covariances offers a new perspective on the time-varying relationship between risk and return across stocks.

7 Conclusion

We develop a simple and transparent, multivariate regression approach to modeling the cross-section of pairwise covariances as a function of firm characteristics. This approach allows us to isolate each characteristic's marginal contribution to systematic factor risk after controlling for exposure to other characteristics. We provide new evidence and insights into the "characteristics versus covariances" debate by leaning on the strength of regressions to tease out patterns across related firm characteristics. While this approach is widely used in research on the cross-section of expected returns, it is almost completely absent from the literature exploring the covariation across stock returns.

We show by comparing univariate regression results to multivariate regression results that approaches that fail to take into account the rich multidimensional patterns across firm characteristics are likely to be confounded by interactions between characteristics. Our approach also identifies asymmetries in the relationships between characteristics and covariances. Almost all characteristics are only associated with higher than average covariance in one extreme leg of their portfolio sorts.

We organize our thirteen composite characteristic groups into priced factors, unpriced factors, and statistical arbitrage anomalies. Our analysis shows that marginal exposure to momentum, size, and value that controls for confounding characteristic exposure generates a spread in average returns and covariance, which is consistent with priced factor exposure. Marginal exposure to low risk, leverage, investment, seasonality, profit growth, and profitability generates robust spreads in covariances but not average returns, consistent with unpriced factor exposure. Marginal exposures to accruals, quality, and debt issuance generate spreads in average returns but not covariances, consistent with a statistical near-arbitrage anomaly.

We show characteristic exposure drives differences in covariances across the business cycle and in response to macroeconomic stress. While high momentum firms covary more than low momentum firms outside of recessions, low mometum firms covary substantially more in recessions. The same pattern holds in response to volatility spikes that proxy for changes in the equity premium. High investment firms covary more than low investment firms and this is entirely due to covariance during recessions. We find that investor sentiment also impacts the cross-section of covariances, but the effects are confined to growth stocks and stocks in technology related industries.

Our analysis advances understanding of the relationships between firm characteristics and systematic risk. The simple regression approach allows us to parse the relationship between characteristics and covariance in a way that controls for confounding characteristic exposure. By reducing dimension at the characteristic level and summarizing the marginal contribution of each characteristic across all latent factors, we complement the growing literature which uses factor analysis to reduce factor model dimension. Our approach offers a new lens through which the relationship between characteristics and asset prices can be examined.

References

Anatolyev, S. and Mikusheva, A. (2021). Factor models with many assets: strong factors, weak factors, and the two-pass procedure. *Journal of Econometrics*.

Ang, A. and Chen, J. (2002). Asymmetric correlations of equity portfolios. *Journal of Financial Economics*, 63(3):443–494.

Anton, M. and Polk, C. (2014). Connected stocks. *Journal of Finance*, 69(3):1099–1127.

Asness, C. S., Frazzini, A., and Pedersen, L. H. (2019). Quality minus junk. *Review of Accounting Studies*, 24(1):34–112.

Baker, M. and Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. *The journal of Finance*, 61(4):1645–1680.

Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics*, 9(1):3–18.

Billingsley, P. (2008). *Probability and measure*. John Wiley & Sons.

Brandt, M. W. and Kang, Q. (2004). On the relationship between the conditional mean and volatility of stock returns: A latent var approach. *Journal of Financial Economics*, 72(2):217–257.

Brandt, M. W., Santa-Clara, P., and Valkanov, R. (2009). Parametric portfolio policies: Exploiting characteristics in the cross-section of equity returns. *Review of Financial Studies*, 22(9):3411–3447.

Bray, M. (1994). The arbitrage pricing theory is not robust 2: Factor structures and factor pricing. *LSE Financial Markets Group Discussion Paper Series*.

Bryzgalova, S. (2015). Spurious factors in linear asset pricing models. *LSE manuscript*, 1(3).

Burnside, C. (2016). Identification and inference in linear stochastic discount factor models with excess returns. *Journal of Financial Econometrics*, 14(2):295–330.

Campbell, S. D. and Diebold, F. X. (2009). Stock returns and expected business conditions: Half a century of direct evidence. *Journal of Business & Economic Statistics*, 27(2):266–278.

Cederburg, S. and O'Doherty, M. S. (2016). Does it pay to bet against beta? on the conditional performance of the beta anomaly. *Journal of Finance*, 71(2):737–774.

Chan, L. K., Karceski, J., and Lakonishok, J. (1999). On portfolio optimization: Forecasting covariances and choosing the risk model. *Review of Financial Studies*, 12(5):937–974.

Chen, N.-F., Roll, R., and Ross, S. A. (1986). Economic forces and the stock market. *Journal of Business*, pages 383–403.

Chordia, T. and Shivakumar, L. (2002). Momentum, business cycle, and time-varying expected returns. *Journal of Finance*, 57(2):985–1019.

Chung, J.-W. and Kang, B. U. (2016). Prime broker-level comovement in hedge fund returns: information or contagion? *Review of Financial Studies*, 29(12):3321–3353.

Cochrane, J. H. (2011). Presidential address: Discount rates. *Journal of Finance*, 66(4):1047–1108.

Cohen, K. J., Hawawini, G. A., Maier, S. F., Schwartz, R. A., and Whitcomb, D. K. (1983). Friction in the trading process and the estimation of systematic risk. *Journal of Financial Economics*, 12(2):263–278.

Daniel, K., Mota, L., Rottke, S., and Santos, T. (2020). The cross-section of risk and returns. *The Review of Financial Studies*, 33(5):1927–1979.

Daniel, K. and Titman, S. (1997). Evidence on the characteristics of cross sectional variation in stock returns. *Journal of Finance*, 52(1):1–33.

Dimson, E. (1979). Risk measurement when shares are subject to infrequent trading. *Journal of Financial Economics*, 7(2):197–226.

Dittmar, R. F. and Lundblad, C. T. (2017). Firm characteristics, consumption risk, and firm-level risk exposures. *Journal of Financial Economics*, 125(2):326–343.

Fama, E. F. and French, K. R. (1989). Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics*, 25(1):23–49.

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

Fama, E. F. and French, K. R. (1997). Industry costs of equity. *Journal of Financial Economics*, 43(2):153–193.

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

Fama, E. F. and French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*.

Fama, E. F. and French, K. R. (2016). Dissecting anomalies with a five-factor model. *Review of Financial Studies*, 29(1):69–103.

Fama, E. F. and French, K. R. (2020). Comparing cross-section and time-series factor models. *Review of Financial Studies*, 33(5):1891–1926.

Fama, E. F. and MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, 81(3):607–636.

Ferson, W. E. and Harvey, C. R. (1991). The variation of economic risk premiums. *Journal of Political Economy*, 99(2):385–415.

Frazzini, A. and Pedersen, L. H. (2014). Betting against beta. *Journal of Financial Economics*, 111(1):1–25.

Freyberger, J., Neuhierl, A., and Weber, M. (2020). Dissecting characteristics nonparametrically. *Review of Financial Studies*, 33(5):2326–2377.

Gao, G. (2011). Characteristic-based covariances and cross-sectional expected returns. *Available at SSRN 1786911*.

Giglio, S. and Xiu, D. (2021). Asset pricing with omitted factors. *Journal of Political Economy*, 129(7):1947–1990.

Giglio, S., Xiu, D., and Zhang, D. (2021). Test assets and weak factors. Technical report, National Bureau of Economic Research.

Gilles, C. and LeRoy, S. F. (1991). On the arbitrage pricing theory. *Economic theory*, 1(3):213–229.

Gospodinov, N., Kan, R., and Robotti, C. (2017). Spurious inference in reduced-rank asset-pricing models. *Econometrica*, 85(5):1613–1628.

Green, T. C. and Hwang, B.-H. (2009). Price-based return comovement. *Journal of Financial Economics*, 93(1):37–50.

Greenwood, R. and Thesmar, D. (2011). Stock price fragility. *Journal of Financial Economics*, 102(3):471–490.

Grieser, W., Lee, J. H., and Zekhnini, M. (2020). Ubiquitous comovement. *Available at SSRN 3385841*.

Hansen, L. P. and Richard, S. F. (1987). The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models. *Econometrica*, pages 587–613.

Harvey, C. R., Liu, Y., and Zhu, H. (2016). . . . and the cross-section of expected returns. *Review of Financial Studies*, 29(1):5–68.

Hou, K., Mo, H., Xue, C., and Zhang, L. (2021). An augmented q-factor model with expected growth. *Review of Finance*, 25(1):1–41.

Hou, K., Xue, C., and Zhang, L. (2015). Digesting anomalies: An investment approach. *Review of Financial Studies*, 28(3):650–705.

Jagannathan, R. and Wang, Z. (1996). The conditional capm and the cross-section of expected returns. *Journal of Finance*, 51(1):3–53.

Jensen, T. I., Kelly, B. T., and Pedersen, L. H. (forthcoming). Is there a replication crisis in finance? *Journal of Finance*.

Kan, R. and Zhang, C. (1999). Two-pass tests of asset pricing models with useless factors. *Journal of Finance*, 54(1):203–235.

Kelly, B. T., Pruitt, S., and Su, Y. (2019). Characteristics are covariances: A unified model of risk and return. *Journal of Financial Economics*, 134(3):501–524.

Kim, S., Korajczyk, R. A., and Neuhierl, A. (2021). Arbitrage portfolios. *Review of Financial Studies*, 34(6):2813–2856.

Kleibergen, F. (2009). Tests of risk premia in linear factor models. *Journal of Econometrics*, 149(2):149–173.

Kozak, S., Nagel, S., and Santosh, S. (2018). Interpreting factor models. *The Journal of Finance*, 73(3):1183–1223.

Kozak, S., Nagel, S., and Santosh, S. (2020). Shrinking the cross-section. *Journal of Financial Economics*, 135(2):271–292.

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

Lettau, M. and Pelger, M. (2020). Factors that fit the time series and cross-section of stock returns. *Review of Financial Studies*, 33(5):2274–2325.

Lewellen, J., Nagel, S., and Shanken, J. (2010). A skeptical appraisal of asset pricing tests. *Journal of Financial economics*, 96(2):175–194.

Li, J. and Zhang, H. H. (2017). Short-run and long-run consumption risks, dividend processes, and asset returns. *Review of Financial Studies*, 30(2):588–630.

Liew, J. and Vassalou, M. (2000). Can book-to-market, size and momentum be risk factors that predict economic growth? *Journal of Financial Economics*, 57(2):221–245.

Liu, L. X. and Zhang, L. (2008). Momentum profits, factor pricing, and macroeconomic risk. *Review of Financial Studies*, 21(6):2417–2448.

Lopez-Lira, A. and Roussanov, N. L. (2020). Do common factors really explain the cross-section of stock returns? *Available at SSRN 3628120*.

Lou, D. and Polk, C. (2021). Comomentum: Inferring Arbitrage Activity from Return Correlations. *Review of Financial Studies*.

Ludvigson, S. C. and Ng, S. (2007). The empirical risk–return relation: A factor analysis approach. *Journal of Financial Economics*, 83(1):171–222.

Lustig, H. and Verdelhan, A. (2012). Business cycle variation in the risk-return trade-off. *Journal of Monetary Economics*, 59:S35–S49.

Manela, A. and Moreira, A. (2017). News implied volatility and disaster concerns. *Journal of Financial Economics*, 123(1):137–162.

Markowitz, H. (1952). Portfolio selection. *Journal of Finance*, 7(1):77–91.

Martin, I. (2017). What is the expected return on the market? *Quarterly Journal of Economics*, 132(1):367–433.

Merton, R. C. et al. (1973). An intertemporal capital asset pricing model. *Econometrica*, 41(5):867–887.

Moskowitz, T. J. (2003). An analysis of covariance risk and pricing anomalies. *Review of Financial Studies*, 16(2):417–457.

Nawalkha, S. K. (1997). A multibeta representation theorem for linear asset pricing theories. *Journal of Financial Economics*, 46(3):357–381.

Novy-Marx, R. (2014). Understanding defensive equity. Technical report, National Bureau of Economic Research.

Perez-Quiros, G. and Timmermann, A. (2000). Firm size and cyclical variations in stock returns. *Journal of Finance*, 55(3):1229–1262.

Pirinsky, C. and Wang, Q. (2006). Does corporate headquarters location matter for stock returns? *Journal of Finance*, 61(4):1991–2015.

Roll, R. (1977). A critique of the asset pricing theory's tests part i: On past and potential testability of the theory. *Journal of Financial Economics*, 4(2):129–176.

Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, 13(3):341–360.

Shanken, J. (1982). The arbitrage pricing theory: is it testable? *Journal of Finance*, 37(5):1129–1140.

Shanken, J. (1987). Nonsynchronous data and the covariance-factor structure of returns. *Journal of Finance*, 42(2):221–231.

Sharpe, W. F. (1963). A simplified model for portfolio analysis. *Management science*, 9(2):277–293.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, 19(3):425–442.

Table 1: Multivariate Regression of Covariances on Characteristic Indicators

This table shows a panel regression of pairwise, firm covariances on characteristic indicators from our 13 Characteristic groups. The regression equation is described in Equation 7. The pairwise, covariances are calculated with the Shanken (1987) adjustments and standardized over the full sample so the coefficients are interpretable in units of standard deviation. Standard errors are clustered by pair and quarter.

Accruals	1	2	3	4	5
1	0.01	-0.00	0.00	0.01	0.02
	(0.42)	(-0.06)	(0.16)	(1.27)	(2.07)
2		-0.00	-0.00	0.01	0.01
		(-0.25)	(-0.31)	(1.11)	(1.25)
3				0.01	0.01
				(1.63)	(1.51)
4				0.02	0.02
				(1.51)	(1.55)
5					0.03
					(1.63)

Debt Issuance	1	2	3	4	5
1	0.01	0.01	0.00	0.02	0.02
	(0.81)	(0.76)	(0.50)	(1.83)	(2.00)
2		0.01	0.00	0.02	0.02
		(0.69)	(0.65)	(2.25)	(2.53)
3				0.01	0.01
				(3.15)	(2.68)
4				0.03	0.03
				(3.01)	(3.37)
5					0.02
					(2.73)

Investment	1	2	3	4	5
1	0.07	0.05	0.03	0.02	-0.00
	(2.40)	(2.22)	(2.15)	(0.96)	(-0.22)
2		0.04	0.02	0.01	-0.01
		(2.00)	(1.86)	(0.48)	(-1.29)
3				-0.01	-0.02
				(-1.40)	(-2.54)
4				-0.02	-0.03
				(-1.42)	(-2.17)
5					-0.03
					(-1.85)

Leverage	1	2	3	4	5
1	-0.05	-0.04	-0.04	-0.04	-0.05
	(-2.10)	(-2.38)	(-2.99)	(-2.91)	(-3.64)
2		-0.01	-0.00	-0.00	-0.02
		(-0.85)	(-0.69)	(-0.14)	(-1.89)
3				0.00	-0.02
				(0.30)	(-1.85)
4				0.02	0.01
				(1.32)	(0.53)
5					0.02
					(0.88)

Low Risk	1	2	3	4	5
1	0.40	0.23	0.13	0.05	-0.02
	(10.45)	(9.80)	(8.98)	(3.40)	(-1.85)
2		0.12	0.05	-0.00	-0.07
		(7.67)	(8.05)	(-0.46)	(-6.74)
3				-0.04	-0.09
				(-6.17)	(-10.00)
4				-0.07	-0.11
				(-5.81)	(-8.94)
5					-0.14
					(-9.89)

Momentum	1	2	3	4	5
1	0.09	0.05	0.01	-0.01	-0.04
	(3.15)	(2.36)	(0.96)	(-1.03)	(-2.72)
2		0.03	0.01	-0.00	-0.01
		(2.16)	(1.69)	(-0.15)	(-1.22)
3				-0.00	-0.01
				(-0.78)	(-0.96)
4				-0.00	0.00
				(-0.09)	(0.28)
5					0.03
					(1.41)

Profit Growth	1	2	3	4	5
1	-0.01	-0.01	-0.01	-0.01	-0.00
	(-0.32)	(-0.93)	(-1.19)	(-1.06)	(-0.29)
2		-0.01	-0.00	-0.00	0.00
		(-1.11)	(-1.19)	(-0.57)	(0.34)
3				0.00	0.01
				(0.31)	(1.84)
4				0.00	0.01
				(0.34)	(1.46)
5					0.03
					(2.63)

Profitability	1	2	3	4	5
1	0.02	0.01	0.00	0.00	0.01
	(0.95)	(0.47)	(0.14)	(0.20)	(0.87)
2		0.00	0.00	0.00	0.01
		(0.27)	(0.05)	(0.31)	(0.84)
3				0.00	0.01
				(0.40)	(0.97)
4				0.01	0.01
				(0.51)	(0.72)
5					0.01
					(0.90)

Quality	1	2	3	4	5
1	0.02	0.00	0.00	0.01	0.00
	(0.62)	(0.04)	(0.28)	(0.84)	(0.34)
2		-0.01	-0.00	-0.00	-0.01
		(-0.56)	(-0.87)	(-0.03)	(-1.11)
3				0.00	-0.00
				(0.33)	(-0.65)
4				0.01	-0.00
				(0.57)	(-0.10)
5					-0.01
					(-0.55)

Seasonality	1	2	3	4	5
1	0.04	0.02	0.02	0.02	0.00
	(3.13)	(2.70)	(2.70)	(2.10)	(0.37)
2		0.01	0.01	0.01	0.00
		(1.26)	(1.38)	(1.22)	(0.03)
3				0.01	-0.00
				(1.25)	(-0.44)
4				0.01	0.01
				(1.06)	(0.44)
5					-0.00
					(-0.23)

Size	1	2	3	4	5
1	-0.02	-0.02	-0.02	-0.03	-0.03
	(-1.40)	(-1.65)	(-2.46)	(-2.22)	(-1.56)
2		-0.01	-0.01	-0.01	0.00
		(-0.96)	(-1.30)	(-1.14)	(0.04)
3				0.00	0.01
				(0.13)	(1.28)
4				0.00	0.02
				(0.33)	(1.42)
5					0.06
					(2.27)

Skewness	1	2	3	4	5
1	-0.06	-0.04	-0.03	-0.03	-0.02
	(-4.21)	(-3.17)	(-4.51)	(-3.61)	(-2.13)
2		-0.01	-0.01	-0.00	0.00
		(-1.09)	(-1.13)	(-0.51)	(0.16)
3				0.00	0.01
				(0.95)	(1.47)
4				0.01	0.01
				(0.90)	(1.51)
5					0.02
					(1.61)

Value	1	2	3	4	5
1	0.02	-0.01	-0.03	-0.03	-0.04
	(0.60)	(-0.73)	(-1.98)	(-2.47)	(-2.35)
2		-0.01	-0.01	-0.01	-0.01
		(-0.85)	(-1.34)	(-0.82)	(-0.54)
3				0.01	0.01
				(0.91)	(0.72)
4				0.01	0.02
				(1.21)	(1.21)
5					0.05
					(1.62)

Cons	-0.06
	(-1.78)
Obs	23,874,657
R^2	0.03

Coal	1.22	**Petroleum & Gas**	0.28	**Business Services**	0.13	**Aircraft**	0.09	**Candy & Soda**	0.06	**Apparel**	0.05	**Non-Metallic and Industrial Metal**	0.03	**Beer & Liquor**	-0.02	**Food Products**	-0.06
	(2.76)		(5.14)		(1.97)		(2.50)		(1.44)		(1.58)		(0.51)		(-0.27)		(-2.38)
Electronic Equip	0.48	**Computer Software**	0.26	**Steel Works Etc**	0.12	**Restaurants & Hotels**	0.09	**Healthcare**	0.06	**Machinery**	0.05	**Recreation**	0.02	**Printing & Publicshing**	-0.02	**Utilities**	-0.07
	(4.22)		(3.08)		(4.41)		(2.70)		(1.42)		(2.73)		(0.27)		(-0.90)		(-3.45)
Precious Metals	0.45	**Construction**	0.18	**Communication**	0.12	**Business Services**	0.08	**Chemicals**	0.06	**Transportation**	0.05	**Consumer Goods**	-0.00	**Rubber and Plastic Products**	-0.02	**Textiles**	-0.08
	(5.52)		(4.35)		(2.29)		(2.19)		(2.20)		(1.84)		(-0.19)		(-0.55)		(-1.48)
Electrical Equip	0.36	**Shipping & Rail**	0.18	**Computer Hardware**	0.11	**Tobacco Products**	0.07	**Fabricated Products**	0.06	**Automobiles & Trucks**	0.04	**Pharmaceutical Products**	0.00	**Shipping Containers**	-0.03	**Almost Nothing**	-0.13
	(2.99)		(1.99)		(2.30)		(1.76)		(0.41)		(1.76)		(0.01)		(-0.97)		(-1.17)
Entertainment	0.28	**Personal Services**	0.13	**Transportation**	0.11	**Measuring & Control Equip**	0.07	**Retail**	0.06	**Construction Materials**	0.03	**Defense**	0.00	**Medical Equip**	-0.04	**Agriculture**	-0.24
	(1.72)		(2.24)		(4.84)		(2.31)		(3.34)		(1.00)		(0.02)		(-1.53)		(-3.06)

Table 2: Univariate vs Multivariate Coefficients

This table compares univariate vs multivariate regressions of pairwise, firm covariances on characteristic indicators from our 13 Characteristic groups. The univariate regressions regress pairwise, firm covariances on 15 indicators representing only one characteristic group without a constant. The multivariate regression includes all thirteen characteristic groups and is described in Equation 7. The pairwise, covariances are calculated with the Shanken (1987) adjustments and standardized over the full sample so the coefficients are interpretable in units of standard deviation. Standard errors are clustered by pair and quarter.

Cov	Uni	Multi	Cov	Uni	Multi	Cov	Uni	Multi
Accruals 11	0.10	0.01	Momentum 11	0.19	0.09	Size 11	-0.09	-0.02
	(2.24)	(0.42)		(3.19)	(3.15)		(-3.72)	(-1.40)
Accruals 15	0.05	0.02	Momentum 15	0.00	-0.04	Size 15	-0.03	-0.03
	(1.30)	(2.07)		(0.01)	(-2.72)		(-0.94)	(-1.56)
Accruals 55	0.05	0.03	Momentum 55	0.04	0.03	Size 55	0.18	0.06
	(1.06)	(1.63)		(1.25)	(1.41)		(3.23)	(2.27)
Debt Iss 11	0.06	0.01	Prof Gr 11	0.07	-0.01	Skewness 11	0.02	-0.06
	(1.30)	(0.81)		(1.97)	(-0.22)		(0.54)	(-4.21)
Debt Iss 15	0.04	0.02	Prof Gr 15	0.04	-0.00	Skewness 15	0.01	-0.02
	(0.94)	(2.00)		(1.07)	(-0.29)		(0.19)	(-2.13)
Debt Iss 55	0.04	0.02	Prof Gr 55	0.07	0.03	Skewness 55	0.02	0.02
	(1.12)	(2.73)		(1.86)	(2.63)		(0.53)	(1.61)
Investment 11	0.20	0.07	Profitability 11	0.17	0.02	Value 11	0.25	0.02
	(3.53)	(2.40)		(3.40)	(0.95)		(4.50)	(0.60)
Investment 15	0.03	-0.00	Profitability 15	0.02	0.01	Value 15	-0.04	-0.04
	(0.78)	(-0.22)		(0.65)	(0.87)		(-1.19)	(-2.35)
Investment 55	-0.03	-0.03	Profitability 55	-0.03	0.01	Value 55	-0.06	0.05
	(-1.12)	(-1.85)		(-1.11)	(0.90)		(-1.40)	(1.62)
Leverage 11	-0.10	-0.05	Quality 11	0.07	0.02			
	(-4.69)	(-2.10)		(1.51)	(0.62)			
Leverage 15	-0.05	-0.05	Quality 15	0.02	0.00			
	(-1.67)	(-3.64)		(0.47)	(0.34)			
Leverage 55	0.21	0.02	Quality 55	0.01	-0.01			
	(4.10)	(0.88)		(0.38)	(-0.55)			
Low Risk 11	0.40	0.40	Seasonality 11	0.07	0.04			
	(6.36)	(10.45)		(1.79)	(3.13)			
Low Risk 15	-0.09	-0.02	Seasonality 15	0.01	0.00			
	(-2.98)	(-1.85)		(0.39)	(0.37)			
Low Risk 55	-0.20	-0.14	Seasonality 55	0.00	-0.00			
	(-11.41)	(-9.89)		(0.10)	(-0.23)			

Table 3: Characteristic Overlap by Quintile

The table shows the overlap of extreme quintiles for each of the thirteen characteristics. The row of each matrix represents the characteristic whose first (fifth) quintile's overlap is being examined in Panel A (B). The i, j entry in Panels A and B are given by the following:

$$PanelA(i,j) = \frac{MAX(\#firms \in Q_{1,i} \cap Q_{1,j}, \#firms \in Q_{1,i} \cap Q_{5,j})}{\#firms \in Q_{1,i}}$$

$$PanelB(i,j) = \frac{MAX(\#firms \in Q_{5,i} \cap Q_{1,j}, \#firms \in Q_{5,i} \cap Q_{5,j})}{\#firms \in Q_{5,i}}$$

where $Q_{N,j}$ denotes quintile N of characteristic j. We report the maximum overlap of a given characteristic's extreme quintiles in order to save space and because the purpose is to show that an extreme quintile of characteristic i may overlap with at least one of the extreme quintiles of characteristic j. Bold numbers denote entries where the maximum overlap is between a fist quintile and fifth quintile pair.

Panel A: First Quintile

	Accr	DbtIss	Inv	Lev	Lowrisk	Mom	ProfitGr	Prof	Qual	Season	Size	Skew	Val
Accr	100	23	40	**27**	29	25	21	18	22	25	**21**	**21**	31
DbtIss	23	100	49	27	25	25	29	24	25	23	20	20	22
Inv	40	49	100	**34**	41	27	**25**	**22**	**29**	25	**19**	21	47
Lev	18	27	**23**	100	**32**	**18**	21	21	46	20	28	**19**	**46**
Lowrisk	29	25	41	**43**	100	35	28	42	29	27	**31**	30	50
Mom	25	25	27	**26**	35	100	35	29	26	22	**35**	22	**24**
ProfitGr	**23**	29	**22**	21	28	35	100	35	31	21	**29**	21	**23**
Prof	**31**	24	**31**	**26**	42	29	35	100	53	**23**	**44**	23	34
Qual	22	25	**28**	46	29	26	31	53	100	27	**29**	20	**29**
Season	25	23	25	**29**	27	22	21	22	27	100	22	**20**	27
Size	17	20	18	28	**28**	**24**	**22**	**41**	**24**	22	100	18	**21**
Skew	**21**	**21**	21	**21**	30	22	21	23	20	**22**	**21**	100	23
Val	31	**25**	47	**58**	50	**29**	**33**	34	**30**	27	**23**	23	100

Panel B: Fifth Quintile

	Accr	DbtIss	Inv	Lev	Lowrisk	Mom	ProfitGr	Prof	Qual	Season	Size	Skew	Val
Accr	100	25	39	19	**23**	23	**23**	**31**	**19**	27	24	**21**	23
DbtIss	25	100	43	33	**24**	23	28	24	27	25	26	**21**	**25**
Inv	39	43	100	**23**	25	22	**22**	**31**	**28**	24	26	**21**	35
Lev	**27**	33	**34**	100	**43**	**26**	32	30	40	**29**	30	**22**	**59**
Lowrisk	16	18	25	**32**	100	18	18	25	21	19	**28**	23	34
Mom	23	23	22	21	**21**	100	31	22	25	22	**24**	**20**	**29**
ProfitGr	21	28	**25**	32	**25**	31	100	31	34	**21**	**22**	20	**33**
Prof	**18**	24	**22**	30	25	22	31	100	57	**20**	**41**	19	**22**
Qual	**21**	27	**29**	40	21	25	34	57	100	21	**24**	21	**30**
Season	27	25	24	**19**	**20**	22	20	**24**	21	100	25	**22**	**20**
Size	24	26	26	30	**31**	**35**	**29**	**44**	**29**	25	100	22	**23**
Skew	**21**	20	**21**	21	23	**21**	20	**20**	21	20	22	100	21
Val	23	19	35	**46**	34	**24**	**23**	**15**	**29**	19	21	21	100

Table 4: Multivariate Regression of Returns on Characteristic Indicators

This table shows the results of regressions of quarterly stock returns regressed on indicator functions representing each quintile of our 13 characteristic groups. The indicator equals one if the stock has a characteristic value in the given quintile. To identify the coefficients we have no indicator for stocks in quintile three, which is a neutral loading on the characteristic. Standard errors are clustered by stock and quarter.

$$Ret_{i,t+1} = \gamma_0 + \sum_{z=1}^{13} \Big[\gamma_{z,1}\mathbb{1}_1(z_i) + \gamma_{z,2}\mathbb{1}_2(z_i) + \gamma_{z,4}\mathbb{1}_4(z_i) + \gamma_{z,5}\mathbb{1}_5(z_i)\Big] + e_i.$$

Accruals 1	-0.63	Leverage 1	0.15	Prof Gr1	-0.24	Seasonality 1	-0.21	Value 1	-0.53
	(-2.75)		(0.53)		(-1.10)		(-1.12)		(-1.29)
Accruals 2	-0.15	Leverage 2	0.11	Prof Gr2	-0.03	Seasonality 2	0.03	Value 2	-0.26
	(-1.12)		(0.72)		(-0.20)		(0.21)		(-1.39)
Accruals 4	0.24	Leverage 4	-0.24	Prof Gr4	-0.15	Seasonality 4	0.08	Value 4	0.46
	(1.76)		(-0.99)		(-1.19)		(0.73)		(3.24)
Accruals 5	0.51	Leverage 5	0.27	Prof Gr5	-0.16	Seasonality 5	0.16	Value 5	1.15
	(2.87)		(0.59)		(-0.63)		(1.03)		(4.18)
Debt Iss 1	-0.58	Low Risk 1	-0.40	Profitability 1	-0.33	Size 1	-0.75	Cons	2.66
	(-2.16)		(-0.65)		(-0.88)		(-2.62)		(3.67)
Debt Iss 2	-0.05	Low Risk 2	0.06	Profitability 2	-0.06	Size 2	-0.45		
	(-0.43)		(0.25)		(-0.43)		(-3.30)		
Debt Iss 4	0.03	Low Risk 4	-0.21	Profitability 4	0.24	Size 4	0.35		
	(0.24)		(-1.38)		(1.91)		(2.05)		
Debt Iss 5	0.20	Low Risk 5	-0.44	Profitability 5	0.36	Size 5	0.86		
	(1.30)		(-1.54)		(1.48)		(2.76)		
Investment 1	-0.63	Momentum 1	-0.96	Quality 1	-0.90	Skewness 1	-0.42		
	(-1.70)		(-2.65)		(-2.91)		(-2.04)		
Investment 2	-0.11	Momentum 2	-0.09	Quality 2	0.03	Skewness 2	-0.10		
	(-0.71)		(-0.61)		(0.17)		(-0.87)		
Investment 4	-0.14	Momentum 4	0.11	Quality 4	0.31	Skewness 4	0.02		
	(-0.96)		(0.82)		(2.09)		(0.12)		
Investment 5	0.07	Momentum 5	0.89	Quality 5	0.72	Skewness 5	0.24		
	(0.23)		(2.98)		(2.62)		(1.36)		

Table 5: Model Implied Factor Volatilities, Average Returns, and Sharpe Ratios

The table captures the model implied factor volatilies, average returns, and Sharpe ratios of purchasing the isolated factor exposure after controlling for exposures to the other characteristics. The average returns are given by the coefficients of the multivariate regression in Table 4,

$$\widehat{\mu_k} = \widehat{\gamma}_{k,5} - \widehat{\gamma}_{k,1}.$$

The implied factor volatilites are derived from the coefficient estimates in Table 1,

$$\widehat{\sigma_k^2} = \widehat{\phi}_{k,55} + \widehat{\phi}_{k,11} - 2\widehat{\phi}_{k,15}.$$

	Vol	p-val	AvgRet	p-val	Sharpe
Low Risk	8.87	0.00	0.04	0.96	0.00
Momentum	6.89	0.00	1.85	0.00	0.27
Value	5.88	0.00	1.68	0.00	0.29
Size	4.59	0.00	1.61	0.00	0.35
Leverage	4.40	0.00	0.12	0.84	0.03
Investment	3.45	0.00	0.70	0.23	0.20
Seasonality	2.89	0.00	0.38	0.20	0.13
Profit Growth	2.81	0.00	0.08	0.84	0.03
Profitability	2.30	0.05	0.69	0.21	0.30
Skewness	1.70	0.12	0.66	0.04	0.39
Accruals	1.38	0.30	1.15	0.00	0.83
Quality	0.93	0.73	1.62	0.00	1.74
Debt Issuance	0.00	0.75	0.78	0.03	-

Table 6: Univariate Regression of Covariances on Indicators for Expected Return Proxy

This table presents the results of a regression of quarterly, pairwise stock covariances on indicators formed from quartile sorts on a proxy for expected returns. We proxy expected returns with the predicted values of the average returns on characteristics regressions described in Table 4, $\widehat{Ret}$. We sort each stock into quintiles. Using the quintile for each firm we create 15 indicators representing every possible two firm combination. Standard errors are clustered by time and firm pair.

	ERet 1	ERet 2	ERet 3	ERet 4	ERet 5
ERet 1	0.19	0.11	0.07	0.06	0.05
	(6.08)	(5.58)	(5.70)	(4.71)	(4.28)
ERet 2		0.06	0.03	0.02	0.02
		(4.41)	(4.45)	(3.02)	(2.04)
ERet 3			-0.04	0.00	-0.00
			(-1.40)	(0.06)	(-0.07)
ERet 4				0.00	0.00
				(0.23)	(0.38)
ERet 5					0.01
					(0.81)
Observations	23874657				
R-squared	0.00				

Table 7: Expected Returns, Characteristics and Covariances

The table shows results of the model when including estimated expected returns as an additional characteristic. Expected returns are proxied by the predicted returns from the regression equation in Table 4. Panel A shows the baseline model with the addition of the expected return proxies. Panel B shows the increase in percent variation explained, R^2, as more extensive industry controls and then lagged covariances are added to the model. Standard errors are clustered by pair and quarter.

Accruals 11	0.02	Momentum 11	0.11	Size 11	-0.01
	(0.89)		(3.45)		(-0.41)
Accruals 15	0.02	Momentum 15	-0.04	Size 15	-0.03
	(2.18)		(-2.78)		(-1.55)
Accruals 55	0.03	Momentum 55	0.01	Size 55	0.05
	(1.32)		(0.57)		(1.82)
Debt Iss 11	0.03	Profit Gr 11	0.00	Skewness 11	-0.05
	(1.40)		(0.02)		(-3.22)
Debt Iss 15	0.02	Profit Gr 15	0.00	Skewness 15	-0.02
	(2.29)		(0.09)		(-2.00)
Debt Iss 55	0.02	Profit Gr 55	0.04	Skewness 55	0.02
	(2.41)		(2.80)		(1.33)
Investment 11	0.09	Profitability 11	0.03	Value 11	0.03
	(2.73)		(1.21)		(0.94)
Investment 15	0.00	Profitability 15	0.01	Value 15	-0.04
	(0.17)		(0.93)		(-2.56)
Investment 55	-0.03	Profitability 55	0.01	Value 55	0.02
	(-1.80)		(0.53)		(0.83)
Leverage 11	-0.06	Quality 11	0.03	ERet 11	-0.03
	(-2.51)		(0.94)		(-1.41)
Leverage 15	-0.06	Quality 15	0.00	ERet 15	0.01
	(-3.89)		(0.18)		(0.98)
Leverage 55	0.03	Quality 55	-0.02	ERet 55	0.04
	(0.97)		(-1.27)		(1.73)
Low Risk 11	0.41	Seasonality 11	0.05	Constant	-0.07
	(10.35)		(3.47)		(-2.17)
Low Risk 15	-0.02	Seasonality 15	0.00		
	(-1.29)		(0.34)		
Low Risk 55	-0.13	Seasonality 55	-0.01	Observations	23,874,657
	(-9.82)		(-0.43)	R^2	0.03

	R^2 (%)
Eret Only	0.20
+Baseline	2.69
+Across Industries	3.19
+LagCov	6.03

Table 8: Multivariate Regression of Covariances Interacted with Recession Indicator

This table shows the results of a regression of quarterly pairwise covariances regressed on our characteristic indicators representing thirteen characteristic groups where each indicator is additionally interacted with a recession indicator variable. The recession indicator is one if the economy is in a recession in the current quarter according to NBER and zero otherwise. Standard errors are clustered by pair and quarter.

		*Rec			*Rec			*Rec
RecQ	0.10							
	(1.44)							
Accruals11	0.01	-0.12	Mom11	0.03	0.41	Size11	-0.04	0.06
	(0.72)	(-1.39)		(1.39)	(2.96)		(-2.04)	(1.06)
Accruals15	0.00	0.06	Mom15	-0.02	-0.11	Size15	-0.04	0.02
	(0.62)	(2.16)		(-1.71)	(-1.98)		(-3.01)	(0.37)
Accruals55	0.00	0.23	Mom55	0.06	-0.20	Size55	0.03	0.00
	(0.40)	(1.91)		(3.57)	(-2.71)		(1.44)	(0.00)
DebtIss11	0.01	0.05	ProfitGr11	0.00	-0.05	Skew11	-0.05	-0.03
	(0.88)	(0.57)		(0.19)	(-0.63)		(-4.27)	(-0.46)
DebtIss15	0.02	-0.01	ProfitGr15	0.00	0.00	Skew15	-0.02	0.01
	(2.06)	(-0.33)		(-0.23)	(0.1)		(-1.8)	(0.35)
DebtIss55	0.03	-0.07	ProfitGr55	0.03	0.02	Skew55	0.02	0.05
	(3.14)	(-2.08)		(2.54)	(0.41)		(1.37)	(0.83)
Invest11	0.02	0.32	Profit11	0.01	0.14	Value11	0.04	-0.13
	(1.21)	(2.17)		(0.63)	(1.12)		(2.13)	(-0.84)
Invest15	-0.01	0.07	Profit15	0.00	0.06	Value15	-0.04	0.01
	(-1.2)	(1.16)		(0.33)	(1.5)		(-2.50)	(0.27)
Invest55	-0.01	-0.11	Profit55	0.02	-0.05	Value55	0.00	0.27
	(-0.67)	(-1.47)		(1.52)	(-0.65)		(-0.11)	(1.91)
Lev11	-0.02	-0.20	Qual11	0.03	-0.14	cons	-0.07	
	(-1.03)	(-1.76)		(0.93)	(-1.32)		(-1.82)	
Lev15	-0.03	-0.06	Qual15	0.00	0.03			
	(-2.6)	(-1.25)		(0.04)	(0.56)	R^2(%)	9.73	
Lev55	0.04	0.07	Qual55	-0.02	0.13	Obs	23,687,384	
	(1.66)	(0.76)		(-1.31)	(2.13)			
LowRisk11	0.34	0.40	Season11	0.04	-0.01			
	(10.26)	(2.49)		(2.56)	(-0.11)			
LowRisk15	-0.04	0.09	Season15	0.00	0.03			
	(-3.22)	(1.38)		(-0.29)	(0.68)			
LowRisk55	-0.13	-0.02	Season55	-0.02	0.12			
	(-9.08)	(-0.29)		(-1.87)	(1.42)			

Table 9: Multivariate Regression of Returns Interacted with Recession Indicator

This table shows the results of a regression of quarterly returns regressed on our characteristic indicators representing thirteen characteristic groups where each indicator is additionally interacted with a recession indicator variable. The recession indicator is one if the economy is in a recession in the current quarter according to NBER and zero otherwise. Standard errors are clustered by pair and quarter.

		*Rec			*Rec			*Rec
RecQ	-2.72							
	(-0.92)							
Accruals1	-0.5	-1.23	Mom1	-1.26	2.25	Size1	-0.55	-1.78
	(-2.22)	(-1.24)		(-4.11)	(1.31)		(-1.89)	(-1.57)
Accruals5	0.45	0.44	Mom5	0.99	-0.36	Size5	0.59	2.13
	(2.6)	(0.55)		(3.24)	(-0.32)		(2.00)	(1.75)
DebtIss1	-0.39	-1.22	ProfitGr1	-0.36	0.66	Skew1	-0.27	-0.97
	(-1.76)	(-0.92)		(-1.63)	(0.8)		(-1.57)	(-0.94)
DebtIss5	0.21	-0.25	ProfitGr5	-0.22	0.67	Skew5	0.12	1.04
	(1.22)	(-0.55)		(-0.78)	(0.97)		(0.73)	(1.17)
Invest1	-0.82	0.63	Profit1	-0.43	0.06	Value1	-0.26	-1.68
	(-2.49)	(0.41)		(-1.05)	(0.05)		(-0.65)	(-1.15)
Invest5	0.20	-0.57	Profit5	0.35	0.29	Value5	0.99	1.10
	(0.66)	(-0.5)		(1.37)	(0.30)		(4.53)	(0.87)
Lev1	0.23	-0.45	Qual1	-0.83	-0.92			
	(0.82)	(-0.47)		(-2.49)	(-0.78)			
Lev5	0.06	1.43	Qual5	0.7	0.29			
	(0.13)	(0.79)		(2.51)	(0.29)			
LowRisk1	-0.46	0.15	Season1	-0.19	-0.09	Const	2.97	
	(-0.8)	(0.06)		(-0.97)	(-0.12)		(4.07)	
LowRisk5	-0.42	-0.03	Season5	0.19	0.12	$R^2(\%)$	1.00	
	(-1.44)	(-0.03)		(1.26)	(0.18)	Obs	309,535	

Table 10: Multivariate Regression of Covariances Interacted with Equity Premium Proxy

This table shows the results of a regression of quarterly pairwise covariances regressed on our characteristic indicators representing thirteen characteristic groups where each indicator is additionally interacted with a time-varying estimator of the equity premium. Standard errors are clustered by pair and quarter.

		*EquityPrem			*EquityPrem			*EquityPrem
EquityPrem	0.16							
	(3.91)							
Accruals11	0.02	0.00	Mom11	-0.25	0.16	Size11	0.05	-0.04
	(0.64)	(0.32)		(-2.58)	(3.16)		(2.08)	(-3.17)
Accruals15	-0.07	0.04	Mom15	0.03	-0.03	Size15	-0.04	0.00
	(-1.67)	(1.94)		(0.56)	(-0.98)		(-1.31)	(0.08)
Accruals55	-0.16	0.08	Mom55	0.16	-0.05	Size55	-0.15	0.09
	(-1.87)	(1.94)		(3.94)	(-2.80)		(-1.95)	(2.24)
DebtIss11	-0.08	0.04	ProfitGr11	0.07	-0.03	Skew11	0.00	-0.03
	(-1.87)	(1.81)		(1.36)	(-1.32)		(0.00)	(-1.88)
DebtIss15	0.01	0.00	ProfitGr15	0.03	-0.02	Skew15	-0.04	0.00
	(0.53)	(0.35)		(1.30)	(-1.27)		(-1.54)	(0.30)
DebtIss55	0.08	-0.02	ProfitGr55	0.03	0.00	Skew55	-0.07	0.04
	(4.10)	(-2.29)		(1.34)	(0.09)		(-1.09)	(1.29)
Invest11	-0.07	0.04	Profit11	-0.09	0.04	Value11	0.23	-0.09
	(-1.02)	(1.34)		(-1.38)	(1.28)		(4.89)	(-3.69)
Invest15	-0.04	0.01	Profit15	-0.03	0.02	Value15	-0.03	0.00
	(-1.05)	(0.67)		(-1.98)	(2.46)		(-1.55)	(0.08)
Invest55	0.04	-0.02	Profit55	0.04	0.00	Value55	-0.33	0.17
	(0.96)	(-1.00)		(0.78)	(-0.06)		(-2.92)	(2.94)
Lev11	0.13	-0.07	Qual11	-0.08	0.07	const	-0.39	
	(3.12)	(-3.34)		(-0.81)	(1.29)		(-5.20)	
Lev15	0.11	-0.08	Qual15	-0.08	0.04			
	(2.73)	(-3.65)		(-3.06)	(2.86)	R^2(%)	13.69	
Lev55	0.20	-0.10	Qual55	-0.11	0.04	obs	22,511,873	
	(2.59)	(-2.67)		(-2.09)	(1.73)			
LowRisk11	-0.26	0.33	Season11	0.01	0.02			
	(-1.92)	(4.75)		(0.34)	(1.47)			
LowRisk15	-0.08	0.02	Season15	-0.05	0.02			
	(-2.05)	(1.10)		(-1.64)	(1.55)			
LowRisk55	0.03	-0.09	Season55	-0.1	0.04			
	(0.63)	(-4.23)		(-2.03)	(1.59)			

Table 11: Multivariate Regression of Covariances Interacted with Sentiment

This table shows the results of a regression of quarterly pairwise covariances regressed on our characteristic indicators representing thirteen characteristic groups where each indicator is additionally interacted with a time-varying Sentiment. Standard errors are clustered by pair and quarter.

		*Sent			*Sent			*Sent
Sentiment	0.18							
	(1.57)							
Accruals11	0.00	0.01	Mom11	0.08	0.01	Size11	-0.02	-0.03
	(0.18)	(0.41)		(2.97)	(0.33)		(-0.97)	(-1.76)
Accruals15	0.02	-0.01	Mom15	-0.04	-0.01	Size15	-0.02	-0.03
	(1.89)	(-0.43)		(-2.78)	(-0.87)		(-1.61)	(-1.56)
Accruals55	0.04	-0.03	Mom55	0.03	-0.02	Size55	0.06	-0.04
	(1.55)	(-1.03)		(1.64)	(-0.8)		(2.24)	(-1.36)
DebtIss11	0.03	-0.07	ProfitGr11	0.00	-0.03	Skew11	-0.05	0.00
	(1.55)	(-2.55)		(0.19)	(-1.19)		(-3.63)	(-0.20)
DebtIss15	0.02	-0.02	ProfitGr15	0.00	-0.03	Skew15	-0.02	-0.01
	(2.22)	(-1.76)		(-0.13)	(-1.47)		(-1.72)	(-0.62)
DebtIss55	0.02	0.00	ProfitGr55	0.03	-0.02	Skew55	0.02	-0.01
	(1.88)	(-0.01)		(2.21)	(-0.73)		(1.58)	(-0.57)
Invest11	0.07	0.02	Profit11	0.03	0.02	Value11	0.01	0.12
	(2.08)	(0.34)		(1.3)	(0.62)		(0.30)	(2.73)
Invest15	0.00	-0.04	Profit15	0.01	-0.02	Value15	-0.03	0.01
	(0.41)	(-2.69)		(1.21)	(-1.62)		(-2.09)	(0.60)
Invest55	-0.02	-0.05	Profit55	0.01	-0.05	Value55	0.04	-0.02
	(-1.33)	(-2.37)		(0.86)	(-1.86)		(1.60)	(-0.87)
Lev11	-0.05	0.00	Qual11	0.00	0.00	const	-0.07	
	(-2.27)	(-0.08)		(-0.04)	(-0.04)		(-1.75)	
Lev15	-0.05	-0.02	Qual15	0.00	-0.01			
	(-3.58)	(-1.48)		(-0.38)	(-0.43)	R^2(%)	4.03	
Lev55	0.01	0.07	Qual55	0.00	-0.03	obs	23,641,328	
	(0.58)	(1.81)		(0.08)	(-1.56)			
LowRisk11	0.39	0.01	Season11	0.04	0.03			
	(10.98)	(0.32)		(2.82)	(1.79)			
LowRisk 15	-0.02	-0.04	Season15	0.00	0.00			
	(-1.81)	(-2.47)		(0.33)	(-0.05)			
LowRisk 55	-0.13	0.01	Season55	0.00	-0.02			
	(-9.18)	(0.42)		(0.12)	(-1.00)			

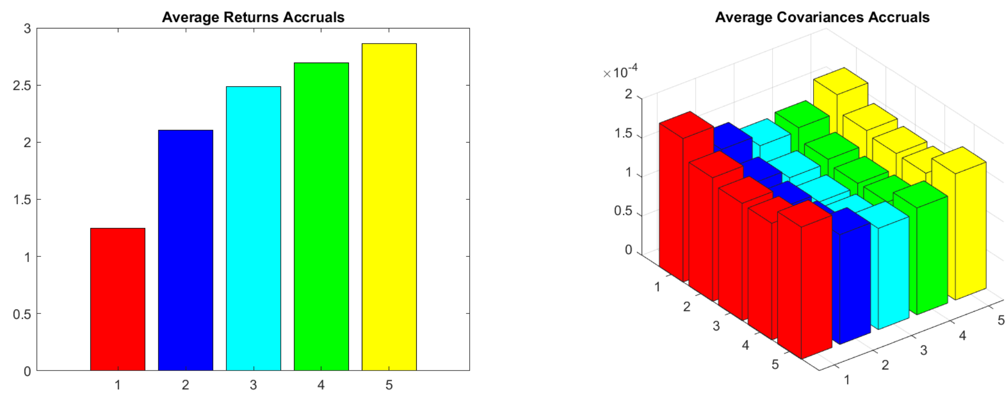

Figure 1: Average Returns and Covariances of Firms Grouped by Accruals

This figure shows average returns (left) and firm, pairwise covariances (right) of portfolios formed on accruals. Each quarter we sort each firm in the sample into quintiles by a composite of six accrual measures based on breakpoints using all eligible firms, such that quintile 1 is high accrual stocks and quintile 5 is low accrual stocks. The left panel shows quarterly average returns for the five sorted portfolios. The right panel shows the pairwise covariance between every two firms in our sample calculated using daily data over the quarter and then averaged within the appropriate bin.

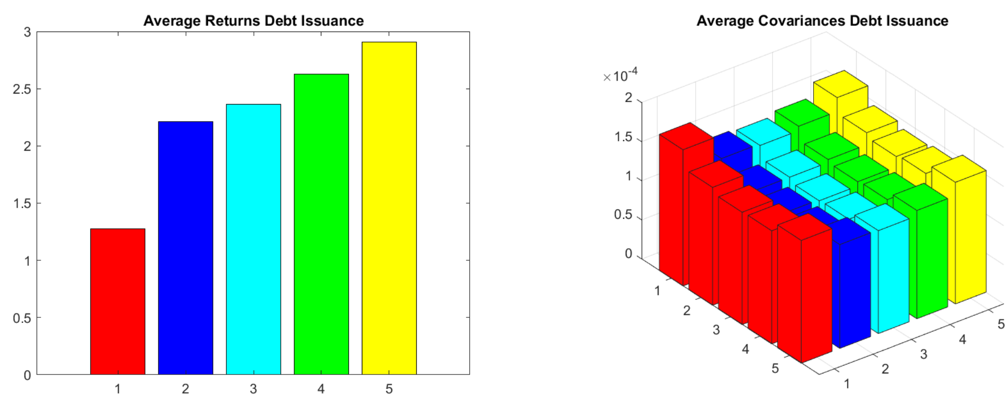

Figure 2: Average Returns and Covariances of Firms Grouped by Debt Issuance

This figure shows average returns (left) and firm, pairwise covariances (right) of portfolios formed on debt issuance. Each quarter we sort each firm in the sample into quintiles by a composite of seven debt issuance measures based on breakpoints using all eligible firms, such that quintile 1 is high debt issuance stocks and quintile 5 is low debt issuance stocks. The left panel shows quarterly average returns for the five sorted portfolios. The right panel shows the pairwise covariance between every two firms in our sample calculated using daily data over the quarter and then averaged within the appropriate bin.

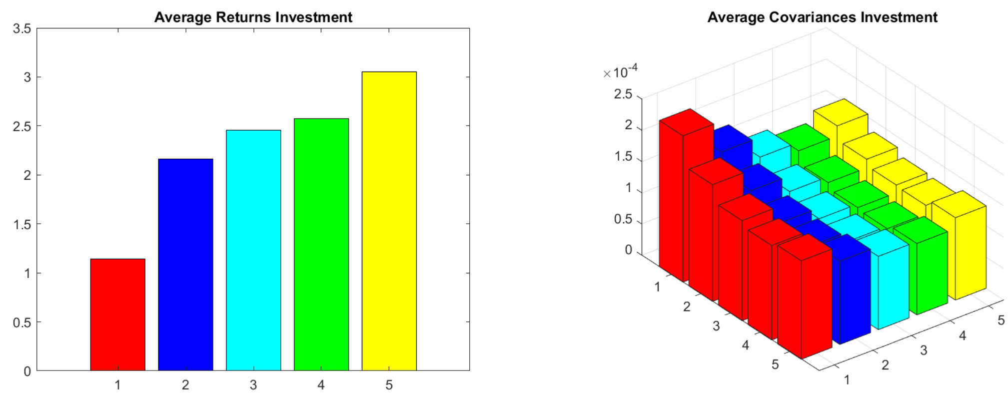

Figure 3: Average Returns and Covariances of Firms Grouped by Investment

This figure shows average returns (left) and firm, pairwise covariances (right) of portfolios formed on investment. Each quarter we sort each firm in the sample into quintiles by a composite of 22 investment measures based on breakpoints using all eligible firms, such that quintile 1 is low investment stocks and quintile 5 is low investment stocks. The left panel shows quarterly average returns for the five sorted portfolios. The right panel shows the pairwise covariance between every two firms in our sample calculated using daily data over the quarter and then averaged within the appropriate bin.

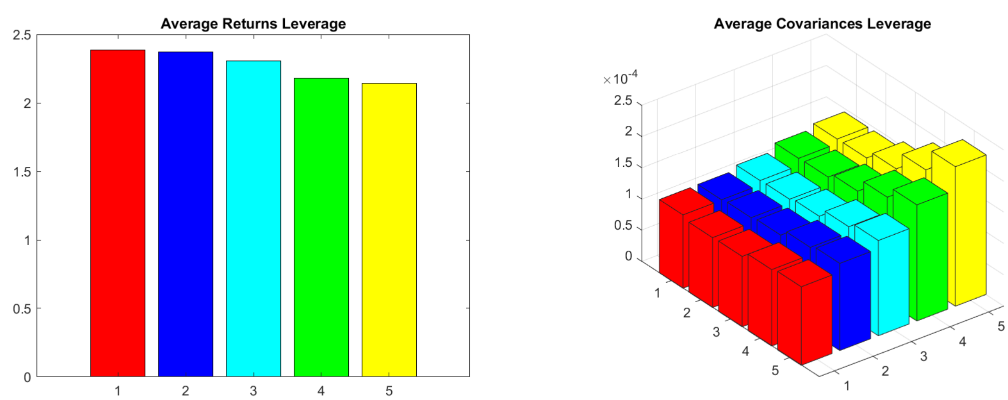

Figure 4: Average Returns and Covariances of Firms Grouped by Leverage

This figure shows average returns (left) and firm, pairwise covariances (right) of portfolios formed on leverage. Each quarter we sort each firm in the sample into quintiles by a composite of 11 leverage measures based on breakpoints using all eligible firms, such that quintile 1 is high leverage stocks and quintile 5 is low leverage stocks. The left panel shows quarterly average returns for the five sorted portfolios. The right panel shows the pairwise covariance between every two firms in our sample calculated using daily data over the quarter and then averaged within the appropriate bin.

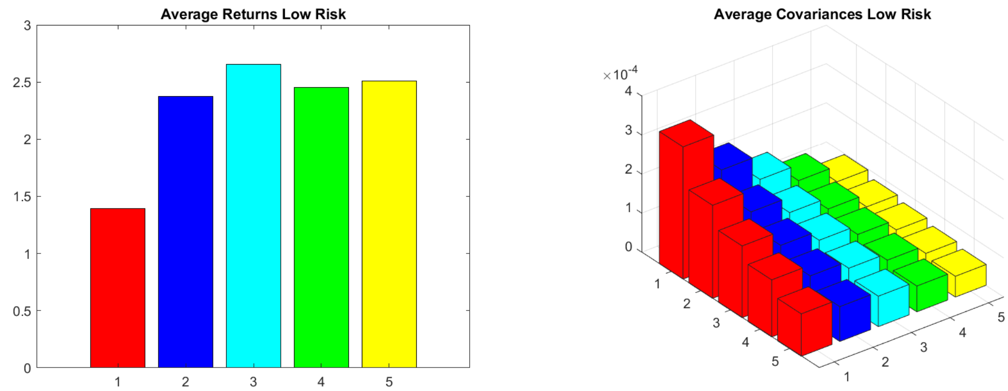

Figure 5: Average Returns and Covariances of Firms Grouped by Low Risk

This figure shows average returns (left) and firm, pairwise covariances (right) of portfolios formed on low risk. Each quarter we sort each firm in the sample into quintiles by a composite of 18 low risk measures based on breakpoints using all eligible firms, such that quintile 1 is high risk stocks and quintile 5 is low risk stocks. The left panel shows quarterly average returns for the five sorted portfolios. The right panel shows the pairwise covariance between every two firms in our sample calculated using daily data over the quarter and then averaged within the appropriate bin.

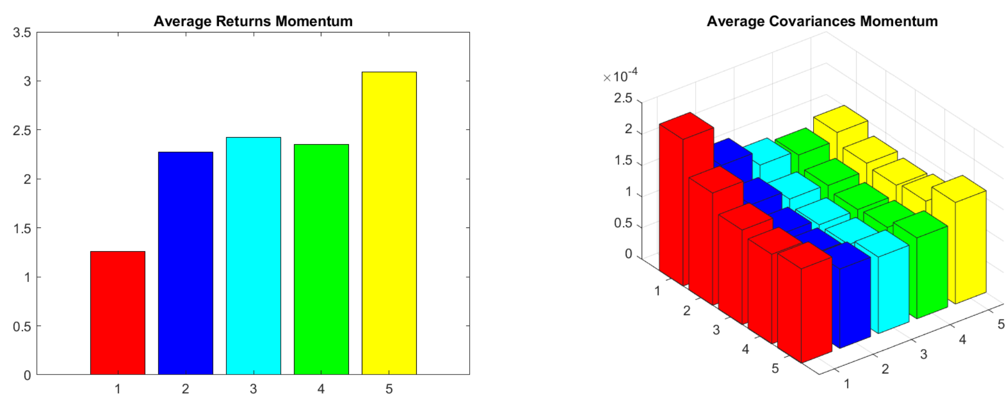

Figure 6: Average Returns and Covariances of Firms Grouped by Momentum

This figure shows average returns (left) and firm, pairwise covariances (right) of portfolios formed on momentum. Each quarter we sort each firm in the sample into quintiles by a composite of eight momentum measures based on breakpoints using all eligible firms, such that quintile 1 is low momentum stocks and quintile 5 is high momentum stocks. The left panel shows quarterly average returns for the five sorted portfolios. The right panel shows the pairwise covariance between every two firms in our sample calculated using daily data over the quarter and then averaged within the appropriate bin.

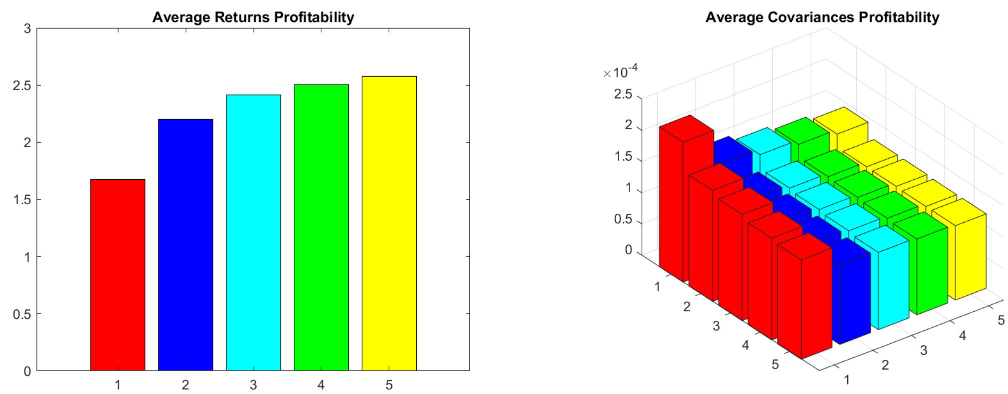

Figure 7: Average Returns and Covariances of Firms Grouped by Profitability

This figure shows average returns (left) and firm, pairwise covariances (right) of portfolios formed on profitability. Each quarter we sort each firm in the sample into quintiles by a composite of 12 profitability measures based on breakpoints using all eligible firms, such that quintile 1 is low profitability stocks and quintile 5 is profitability stocks. The left panel shows quarterly average returns for the five sorted portfolios. The right panel shows the pairwise covariance between every two firms in our sample calculated using daily data over the quarter and then averaged within the appropriate bin.

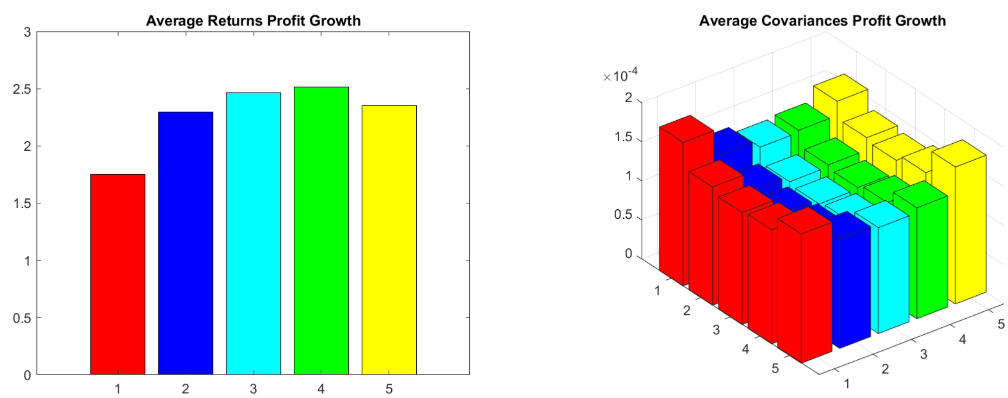

Figure 8: Average Returns and Covariances of Firms Grouped by Profit Growth

This figure shows average returns (left) and firm, pairwise covariances (right) of portfolios formed on profit growth. Each quarter we sort each firm in the sample into quintiles by a composite of 12 profit growth measures based on breakpoints using all eligible firms, such that quintile 1 is low profit growth stocks and quintile 5 is profit growth stocks. The left panel shows quarterly average returns for the five sorted portfolios. The right panel shows the pairwise covariance between every two firms in our sample calculated using daily data over the quarter and then averaged within the appropriate bin.

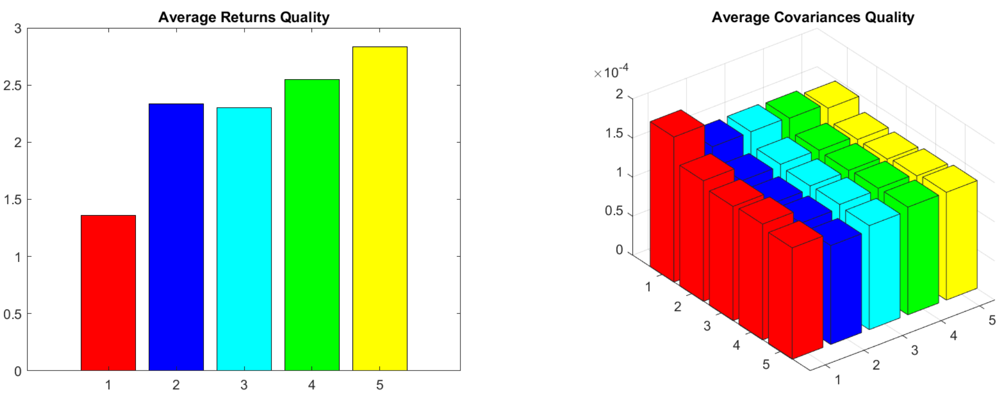

Figure 9: Average Returns and Covariances of Firms Grouped by Quality

This figure shows average returns (left) and firm, pairwise covariances (right) of portfolios formed on quality. Each quarter we sort each firm in the sample into quintiles by a composite of 17 quality measures based on breakpoints using all eligible firms, such that quintile 1 is low quality stocks and quintile 5 is high quality stocks. The left panel shows quarterly average returns for the five sorted portfolios. The right panel shows the pairwise covariance between every two firms in our sample calculated using daily data over the quarter and then averaged within the appropriate bin.

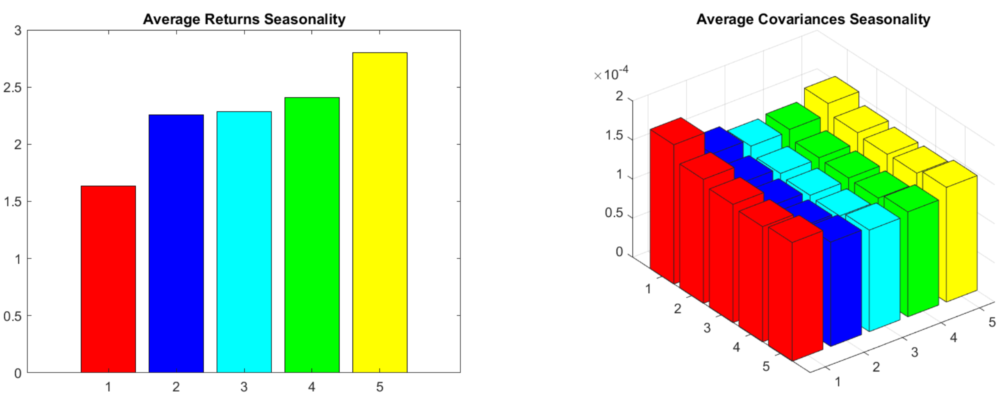

Figure 10: Average Returns and Covariances of Firms Grouped by Seasonality

This figure shows average returns (left) and firm, pairwise covariances (right) of portfolios formed on seasonality. Each quarter we sort each firm in the sample into quintiles by a composite of 17 seasonality measures based on breakpoints using all eligible firms, such that quintile 1 is low seasonality stocks and quintile 5 is high seasonality stocks. The left panel shows quarterly average returns for the five sorted portfolios. The right panel shows the pairwise covariance between every two firms in our sample calculated using daily data over the quarter and then averaged within the appropriate bin.

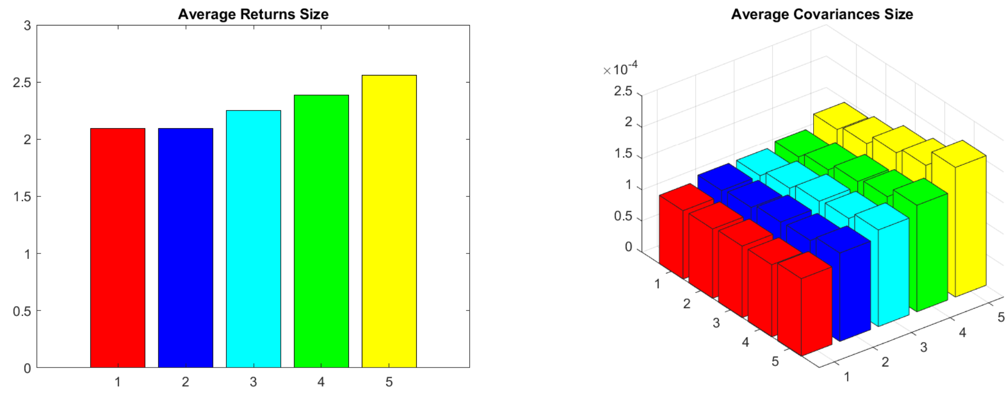

Figure 11: Average Returns and Covariances of Firms Grouped by Size

This figure shows average returns (left) and firm, pairwise covariances (right) of portfolios formed on size. Each quarter we sort each firm in the sample into quintiles by a composite of five size measures based on breakpoints using all eligible firms, such that quintile 1 is large stocks and quintile 5 is small stocks. The left panel shows quarterly average returns for the five sorted portfolios. The right panel shows the pairwise covariance between every two firms in our sample calculated using daily data over the quarter and then averaged within the appropriate bin.

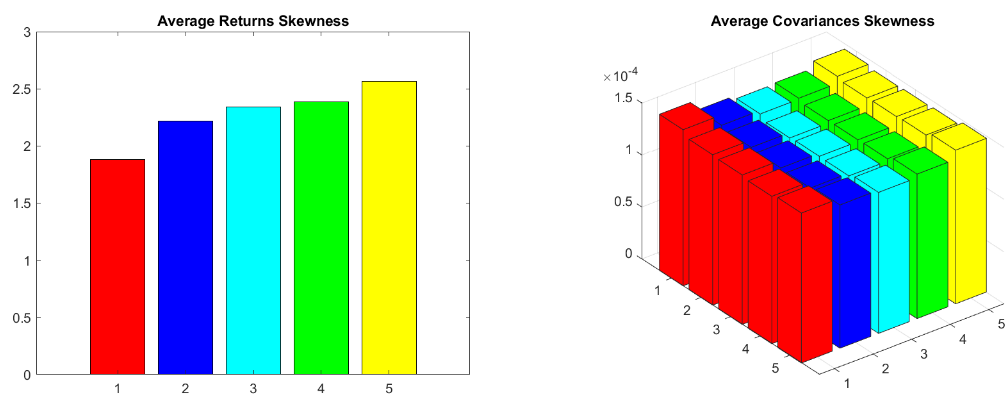

Figure 12: Average Returns and Covariances of Firms Grouped by Skewness

This figure shows average returns (left) and firm, pairwise covariances (right) of portfolios formed on skewness. Each quarter we sort each firm in the sample into quintiles by a composite of six skewness measures based on breakpoints using all eligible firms, such that quintile 1 is high skewness stocks and quintile 5 is low skewness stocks. The left panel shows quarterly average returns for the five sorted portfolios. The right panel shows the pairwise covariance between every two firms in our sample calculated using daily data over the quarter and then averaged within the appropriate bin.

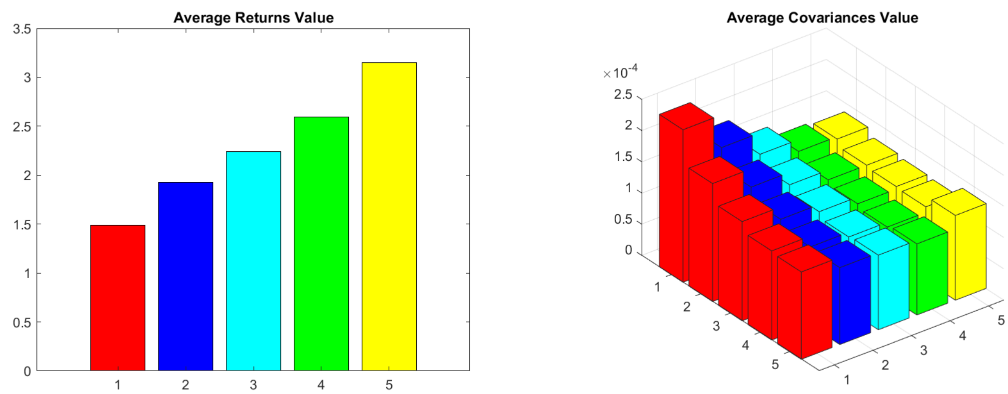

Figure 13: Average Returns and Covariances of Firms Grouped by Value

This figure shows average returns (left) and firm, pairwise covariances (right) of portfolios formed on value. Each quarter we sort each firm in the sample into quintiles by a composite of 18 value measures based on breakpoints using all eligible firms, such that quintile 1 is growth stocks and quintile 5 is value stocks. The left panel shows quarterly average returns for the five sorted portfolios. The right panel shows the pairwise covariance between every two firms in our sample calculated using daily data over the quarter and then averaged within the appropriate bin.

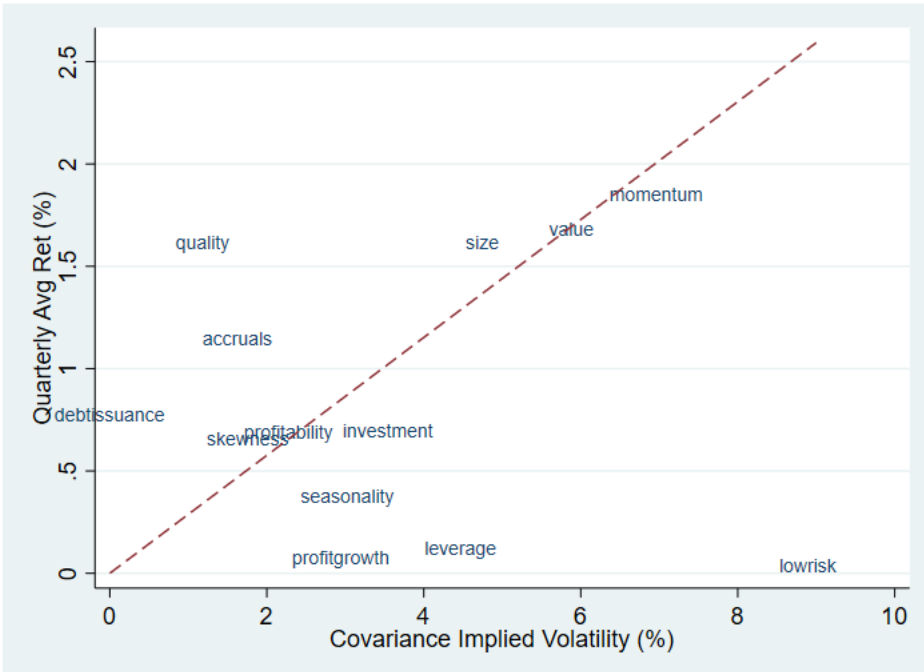

Figure 14: Average Returns vs Implied Factor Variance

This figure graphs quarterly average returns relative to implied factor variances. Both measures are based on estimates implied by regressions on the characteristic group.

The horizontal axis is formed using the coefficient estimates extracted from the multivariate regressions of firm, pairwise covariances on indicator function of the 13 characteristic groups. The implied factor variance is given by

$$\widehat{\sigma_k^2} = \widehat{\phi}_{k,55} + \widehat{\phi}_{k,11} - 2\widehat{\phi}_{k,15}.$$

The vertical axis captures the quarterly average returns of the factors implied by the multivariate panel regression of returns on characteristic indicators of the 13 groups.

$$\widehat{E[f_k]} = \widehat{\phi}_{k,5} - \widehat{\phi}_{k,1}$$

For reference, the dashed line is fit from the origin through the momentum, value, and size characteristics, an implied Sharpe ratio of 0.29.

A Internet Data Appendix

Table 12: **Characteristic Groups**

This table gives the classification of characteristics as described in Jensen et al. (ming).

Description	Citation	Sign
Accruals		
Change in current operating working capital	Richardson, Sloan, Soliman, and Tuna (2005)	-1
Operating accruals	Sloan (1996)	-1
Percent operating accruals	Hafzalla, Lundholm, and Matthew Van Winkle (2011)	-1
Years 16-20 lagged returns, nonannual	Heston and Sadka (2008)	-1
Total accruals	Richardson et al. (2005)	-1
Percent total accruals	Hafzalla et al. (2011)	-1
Debt Issuance		
Abnormal corporate investment	Titman, Wei, and Xie (2004)	-1
Growth in book debt (3 years)	Lyandres, Sun, and Zhang (2008)	-1
Change in financial liabilities	Richardson et al. (2005)	-1
Change in noncurrent operating liabilities	Richardson et al. (2005)	-1
Change in net financial assets	Richardson et al. (2005)	1
Earnings persistence	Francis et al. (2004)	1
Net operating assets	Hirshleifer, Hou, Teoh, and Zhang (2004)	-1
Investment		
Liquidity of book assets	Ortiz-Molina and Phillips (2014)	-1
Asset Growth	Cooper, Gulen, and Schill (2008)	-1
Change in common equity	Richardson et al. (2005)	-1
CAPEX growth (1 year)	Xie (2001)	-1
CAPEX growth (2 years)	Anderson and Garcia-Feijoo (2006)	-1
CAPEX growth (3 years)	Anderson and Garcia-Feijoo (2006)	-1
Change in current operating assets	Richardson et al. (2005)	-1
Change in current operating liabilities	Richardson et al. (2005)	-1
Hiring rate	Belo, Lin, and Bazdresch (2014)	-1
Inventory growth	Belo and Lin (2012)	-1
Inventory change	J. K. Thomas and Zhang (2002)	-1
Change in long-term net operating assets	Fairfield, Whisenant, and Yohn (2003)	-1
Mispricing factor: Management	Stambaugh and Yuan (2017)	1
Change in noncurrent operating assets	Richardson et al. (2005)	-1
Change in net noncurrent operating assets	Richardson et al. (2005)	-1
Change in net operating assets	Hirshleifer et al. (2004)	-1
Change PPE and Inventory	Lyandres et al. (2008)	-1
Long-term reversal	De Bondt and Thaler (1985)	-1
Sales Growth (1 year)	Lakonishok, Shleifer, and Vishny (1994)	-1
Sales Growth (3 years)	Lakonishok et al. (1994)	-1
Sales growth (1 quarter)	Hou, Xue, and Zhang (2018)	-1
Years 2-5 lagged returns, nonannual	Heston and Sadka (2008)	-1
Low Leverage		
Firm age	Jiang, Lee, and Zhang (2005)	-1
Liquidity of market assets	Ortiz-Molina and Phillips (2014)	-1
Book leverage	Fama and French (1992)	-1
The high-low bid-ask spread	Corwin and Schultz (2012)	1
Cash-to-assets	Palazzo (2012)	1
Net debt-to-price	Penman, Richardson, and Tuna (2007)	-1
Earnings volatility	Francis, LaFond, Olsson, and Schipper (2004)	1
R&D-to-sales	Chan, Lakonishok, and Sougiannis (2001)	1
R&D capital-to-book assets	Li (2011)	1
Asset tangibility	Hahn and Lee (2009)	1

Continued on next page

Table 12 – *Continued from previous page*

Description	Citation	Sign
Altman Z-score	Dichev (1998)	1
	Low Risk	
Market Beta	Fama and MacBeth (1973)	-1
Dimson beta	Dimson (1979)	-1
Frazzini-Pedersen market beta	Frazzini and Pedersen (2014)	-1
Downside beta	Ang, Chen, and Xing (2006)	-1
Earnings variability	Francis et al. (2004)	-1
Idiosyncratic volatility from the CAPM (21 days)	Hou, Xue, and Zhang (2018)	-1
Idiosyncratic volatility from the CAPM (252 days)	Ali, Hwang, and Trombley (2003)	-1
Idiosyncratic volatility from the Fama-French 3-factor model	Ang, Hodrick, Xing, and Zhang (2006)	-1
Idiosyncratic volatility from the qfactor model	Hou, Xue, and Zhang (2018)	-1
Cash flow volatility	Huang (2009)	-1
Maximum daily return	Bali, Cakici, and Whitelaw (2011)	-1
Highest 5 days of return	Bali, Brown, and Tang (2017)	-1
Return volatility	Ang, Hodrick, et al. (2006)	-1
Years 6-10 lagged returns, nonannual	Heston and Sadka (2008)	-1
Share turnover	Datar, Naik, and Radcliffe (1998)	-1
Number of zero trades with turnover as tiebreaker (1 month)	Liu (2006)	1
Number of zero trades with turnover as tiebreaker (6 months)	Liu (2006)	1
Number of zero trades with turnover as tiebreaker (12 months)	Liu (2006)	1
	Momentum	
Current price to high price over last year	George and Hwang (2004)	1
Residual momentum t-12 to t-1	Blitz, Huij, and Martens (2011)	1
Residual momentum t-6 to t-1	Blitz et al. (2011)	1
Price momentum t-3 to t-1	Jegadeesh and Titman (1993)	1
Price momentum t-6 to t-1	Jegadeesh and Titman (1993)	1
Price momentum t-9 to t-1	Jegadeesh and Titman (1993)	1
Price momentum t-12 to t-1	Jegadeesh and Titman (1993)	1
Year 1-lagged return, nonannual	Heston and Sadka (2008)	1
	Profit Growth	
Change sales minus change Inventory	Abarbanell and Bushee (1998)	1
Change sales minus change receivables	Abarbanell and Bushee (1998)	-1
Change sales minus change SG&A	Abarbanell and Bushee (1998)	1
Change in quarterly return on assets	Hou, Xue, and Zhang (2018)	1
Change in quarterly return on equity	Hou, Xue, and Zhang (2018)	1
Standardized earnings surprise	Foster, Olsen, and Shevlin (1984)	1
Change in operating cash flow to assets	Bouchaud, Krueger, Landier, and Thesmar (2019)	1
Price momentum t-12 to t-7	Novy-Marx (2012)	1
Labor force efficiency	Abarbanell and Bushee (1998)	1
Standardized Revenue surprise	Jegadeesh and Livnat (2006)	1
Year 1-lagged return, annual	Heston and Sadka (2008)	1
Tax expense surprise	J. Thomas and Zhang (2011)	1
	Profitabability	
Coefficient of variation for dollar trading volume	Chordia, Subrahmanyam, and Anshuman (2001)	-1
Return on net operating assets	Soliman (2008)	1
Profit margin	Soliman (2008)	1
Pitroski F-score	Piotroski (2000)	1
Return on equity	Haugen and Baker (1996)	1
Quarterly return on equity	Hou, Xue, and Zhang (2015)	1
Ohlson O-score	Dichev (1998)	-1
Operating cash flow to assets	Bouchaud et al. (2019)	1
Operating profits-to-book equity	Fama and French (2015)	1
Operating profits-to-lagged book equity	Hou, Xue, and Zhang (2018)	1
Coefficient of variation for share turnover	Chordia et al. (2001)	-1

Continued on next page

Table 12 – *Continued from previous page*

Description	Citation	Sign
Quality		
Capital turnover	Haugen and Baker (1996)	1
Cash-based operating profits-tobook assets	Hou, Xue, and Zhang (2018)	1
Cash-based operating profits-tolagged book assets	Ball, Gerakos, Linnainmaa, and Nikolaev (2016)	1
Change gross margin minus change sales	Abarbanell and Bushee (1998)	1
Gross profits-to-assets	Novy-Marx (2013)	1
Gross profits-to-lagged assets	Hou, Xue, and Zhang (2018)	1
Mispricing factor: Performance	Stambaugh and Yuan (2017)	1
Number of consecutive quarters with earnings increases	Barth, Elliott, and Finn (1999)	1
Quarterly return on assets	Balakrishnan, Bartov, and Faurel (2010)	1
Operating profits-to-book assets	Fama and French (2015)	1
Operating profits-to-lagged book equity	Hou, Xue, and Zhang (2018)	1
Operating leverage	Novy-Marx (2011)	1
Quality minus Junk: Composite	C. S. Asness et al. (2019)	1
Quality minus Junk: Growth	C. S. Asness et al. (2019)	1
Quality minus Junk: Profitability	C. S. Asness et al. (2019)	1
Quality minus Junk: Safety	C. S. Asness, Frazzini, and Pedersen (2019)	1
Assets turnover	Soliman (2008)	1
Seasonality		
Market correlation	C. Asness, Frazzini, Gormsen, and Pedersen (2020)	-1
Coskewness	Harvey and Siddique (2000)	-1
Net debt issuance	Bradshaw et al. (2006)	-1
Kaplan-Zingales index	Lamont, Polk, and SaaŽa-Requejo (2001)	1
Change in long-term investments	Richardson et al. (2005)	-1
Taxable income-to-book income	Lev and Nissim (2004)	1
Years 2-5 lagged returns, annual	Heston and Sadka (2008)	1
Years 6-10 lagged returns, annual	Heston and Sadka (2008)	1
Years 11-15 lagged returns, annual	Heston and Sadka (2008)	1
Years 11-15 lagged returns, nonannual	Heston and Sadka (2008)	-1
Years 16-20 lagged returns, annual	Heston and Sadka (2008)	1
Change in short-term investments	Richardson et al. (2005)	1
Amihud Measure	Amihud (2002)	1
Dollar trading volume	Brennan, Chordia, and Subrahmanyam (1998)	-1
Market Equity	Banz (1981)	-1
Price per share	Miller and Scholes (1982)	-1
R&D-to-market	Chan et al. (2001)	1
Size		
Amihud Measure	Amihud (2002)	1
Dollar trading volume	Brennan, Chordia, and Subrahmanyam (1998)	-1
Market Equity	Banz (1981)	-1
Price per share	Miller and Scholes (1982)	-1
R&D-to-market	Chan et al. (2001)	1
Skewness		
Idiosyncratic skewness from the CAPM	Hou, Xue, and Zhang (2018)	-1
Idiosyncratic skewness from the Fama-French 3-factor model	Bali, Engle, and Murray (2016)	-1
Idiosyncratic skewness from the qfactor model	Hou, Xue, and Zhang (2018)	-1
Short-term reversal	Jegadeesh (1990)	-1
Highest 5 days of return scaled by volatility	C. Asness et al. (2020)	-1
Total skewness	Bali et al. (2016)	-1
Value		
Assets-to-market	Fama and French (1992)	1
Book-to-market equity	Rosenberg, Reid, and Lanstein (1985)	1
Book-to-market enterprise value	Penman et al. (2007)	1

Continued on next page

Table 12 – *Continued from previous page*

Description	**Citation**	**Sign**
Net stock issues	Pontiff and Woodgate (2008)	-1
Debt-to-market	Bhandari (1988)	1
Dividend yield	Litzenberger and Ramaswamy (1979)	1
Ebitda-to-market enterprise value	Loughran and Wellman (2011)	1
Equity duration	Dechow, Sloan, and Soliman (2004)	-1
Net equity issuance	Bradshaw, Richardson, and Sloan (2006)	-1
Equity net payout	Daniel and Titman (2006)	1
Net payout yield	Boudoukh, Michaely, Richardson, and Roberts (2007)	1
Payout yield	Boudoukh et al. (2007)	1
Free cash flow-to-price	Lakonishok et al. (1994)	1
Intrinsic value-to-market	Frankel and Lee (1998)	1
Net total issuance	Bradshaw et al. (2006)	-1
Earnings-to-price	Basu (1983)	1
Operating cash flow-to-market	Desai, Rajgopal, and Venkatachalam (2004)	1
Sales-to-market	Barbee Jr, Mukherji, and Raines (1996)	1