A Portfolio Perspective on the Multitude of Firm Characteristics*

Victor DeMiguel Alberto Martín-Utrera Francisco J. Nogales Raman Uppal

March 8, 2018

Abstract

We investigate how many characteristics matter jointly for an investor who cares not only about average returns but also about portfolio risk and transaction costs. Our main finding is that transaction costs significantly increase the dimension of the cross section of stock returns. While in the absence of transaction costs only a small number of characteristics—about six—are significant, in the presence of transaction costs this number *increases* to 15. The explanation is that, as we show analytically and empirically, combining characteristics helps to substantially reduce transaction costs because the trades in the underlying stocks required to rebalance different characteristics net out. Our work demonstrates that transaction costs provide an economic rationale to consider a larger number of characteristics than that considered in prominent asset-pricing models.

Keywords: transaction costs, cross section of stock returns, trading diversification.
JEL Classification: G11

---

*DeMiguel, London Business School, e-mail: avmiguel@london.edu; Martín-Utrera, Lancaster University Management School, e-mail: a.martinutrera@lancaster.ac.uk; Nogales, Universidad Carlos III de Madrid, e-mail: FcoJavier.Nogales@uc3m.es; Uppal, Edhec Business School and CEPR, e-mail: raman.uppal@edhec.edu. We gratefully acknowledge comments from Torben Andersen, Turan Bali, Pedro Barroso, Kerry Back, Malcolm Baker, Alexandre Belloni, Jonathan Berk, Harjoat Bhamra, John Birge, Michael Brandt, Svetlana Bryzgalova, Andrea Buraschi, Veronika Czellar, Serge Darolles, Alex Edmans, Wayne Ferson, Lorenzo Garlappi, René Garcia, Francisco Gomes, Amit Goyal, Nick Hirschey, Christian Julliard, Petri Jylha, Bige Kahraman, Nishad Kapadia, Ralph Koijen, Robert Kosowski, Apostolos Kourtis, Serhiy Kozak, Anton Lines, Abraham Lioui, Raphael Markellos, Lionel Martellini, Spyros Mesomeris, Maurizio Montone, Narayan Naik, Stefan Nagel, Andreas Neuhierl, Ľuboš Pástor, Barbara Ostdiek, Anna Pavlova, Julien Penasse, Ludovic Phalippou, Ilaria Piatti, Jeffrey Pontiff, Riccardo Rebonato, Scott Richardson, Thierry Roncalli, Shrihari Santosh, James Sefton, Georgios Skoulakis, Stephen Taylor, Nikolaos Tessaromatis, Grigory Vilkov, Christian Wagner, Michael Weber, Dacheng Xiu, Paolo Zaffaroni, Frank Zhang, Lu Zhang, and seminar participants at American Finance Association, Cass Business School, Citi Global Quant Research Conference, Conference on Computational and Financial Econometrics, Deutsche Bank Global Quantitative Conference, Edhec Business School, European Finance Association, French Finance Association, HEC Paris, Imperial College Business School, INET Econometric Seminar Series (Oxford), INFORMS Annual Meeting, INQUIRE Europe, INQUIRE UK, Invesco, KU Leuven, Lancaster University Management School, London Business School, London School of Economics, New Methods for the Empirical Analysis of Financial Markets Conference, Northern Finance Association, Norwich Business School, Rice University, Saïd Business School, Universitat Pompeu Fabra, Université Catolique de Louvain, Université de Nantes, Vienna University of Economics and Business, 5th Luxembourg Asset Management Summit, 10th International Conference on Computational and Financial Econometrics (Seville), 30th Annual Seminar of the London Quant Group (Oxford), and XXIV Finance Forum (Madrid). Nogales acknowledges support from the Spanish government for projects MTM2013-44902-P and MTM2017-88797-P, and the UC3M-BS Institute of Financial Big Data.

1 Introduction

Hundreds of variables have been proposed to predict the cross-section of stock returns; see, for instance, Harvey, Liu, and Zhu (2015), McLean and Pontiff (2016), and Hou, Xue, and Zhang (2017).[1] This abundance of cross-sectional predictors leads Cochrane (2011) to ask, "Which characteristics really provide independent information about average returns? Which are subsumed by others?" Likewise, Goyal (2012) states that "these days one has a multitude of variables that seem to explain the cross-sectional pattern of returns. The amount of independent information in these variables is unclear as no study to date [...] has conducted a comprehensive study to analyze the joint impact of these variables."

Cochrane and Goyal challenge researchers to characterize the *dimension* of the cross-section of stock returns by identifying a small set of characteristics that subsume the rest. To address this challenge, several papers use cross-sectional regressions to shrink the cross section of *expected stock returns*; see Green, Hand, and Zhang (2017), Freyberger, Neuhierl, and Weber (2016), and Messmer and Audrino (2017). These papers, however, study the dimension of the cross section in the absence of transaction costs, but transaction costs matter because the hundreds of predicting variables discovered in the literature pose a challenge to the efficient market hypothesis only if investors can exploit them to build a portfolio that delivers superior *risk-adjusted* returns *net of transaction costs*. To address this gap in the literature, our objective is to study how transaction costs affect the dimension of the cross section of stock returns.

To achieve our objective, we study how many firm-specific characteristics matter *jointly* from a *portfolio perspective*; that is, from the perspective of an investor who cares not only about average returns, but also about portfolio risk and transaction costs. A portfolio perspective is required in order to assess how many characteristics matter *jointly* because it is optimal to trade *combinations* of characteristics to reduce both portfolio risk and transaction costs.

[1] See also the following survey papers: Subrahmanyam (2010), Richardson, Tuna, and Wysocki (2010), and Nagel (2013), and the book Bali, Engle, and Murray (2016).

Our main finding is that transaction costs significantly increase the dimension of the cross section of stock returns. We consider a dataset with 51 characteristics and we find that while in the absence of transaction costs only a small number of characteristics—about six—are significant, in the presence of transaction costs this number *increases* to 15. The intuition behind this result is that combining characteristics is advantageous in the presence of transaction costs because the trades in the underlying stocks required to rebalance different characteristics often cancel each other out, and thus, combining a larger number of characteristics allows one to substantially reduce transaction costs. Essentially, combining characteristics allows one to *diversify trading*, just as combining them allows one to diversify risk.

To quantify the benefits from trading diversification we first compare analytically and empirically the average trading volume (turnover) required to exploit an equally weighted portfolio of characteristics in combination with that required to exploit them in isolation. Analytically, we show that the turnover required to rebalance an equally weighted portfolio of the K characteristics is about $1/\sqrt{K}$ of that required to rebalance the characteristics separately. Empirically, we find that while the average monthly turnover required to exploit a characteristic in isolation is 24.09%, the turnover required to exploit an equally weighted combination of characteristics is only 6.71%; that is, trading diversification delivers a 72.15% reduction in turnover. Note that a reduction in turnover will translate into a reduction in transaction costs regardless of the particular manner in which transaction costs are modeled.

To achieve our main research goal, we then extend the parametric portfolio framework in Brandt, Santa-Clara, and Valkanov (2009). Parametric portfolios are obtained by adding to a benchmark portfolio a linear combination of the long-short portfolios associated with each of the firm-specific characteristics considered. To determine which characteristics are jointly significant, we use a screen-and-clean method to test which characteristics have parametric portfolio weights that are significantly different from zero. Then, we address three research questions. First, how many characteristics are jointly significant from a portfolio perspective? Second, how does the answer to this question change with transaction costs? Third, can an investor improve out-of-sample performance net of transaction costs by exploiting a large set of characteristics instead of the small number considered in prominent asset-pricing models?

Our answers to these three questions are as follows. First, in the absence of transaction costs, only a small number of characteristics—about six—are significant. Five characteristics—unexpected quarterly earnings, return volatility, asset growth, 1-month momentum, and gross profitability—are significant because they increase the mean return and also help to reduce the risk of the portfolio of characteristics. A sixth characteristic—beta—is significant *only* because of its ability to reduce the risk of the other characteristics, in particular, the return-volatility characteristic.[2] We find that traditional characteristics such as 12-month momentum and book to market are not significant because, although they have high mean returns, they do not offer a sufficiently attractive tradeoff between portfolio mean return and risk.

Second, in contrast to what one would find if evaluating characteristics in isolation, we find that the presence of transaction costs *increases* the number of jointly significant characteristics from six to 15. This is because the benefits of trading diversification are large when combining characteristics *optimally*. Indeed, we find empirically that the marginal transaction cost of trading the stocks underlying a characteristic is reduced by around 65% on average when they are combined in an optimal parametric portfolio. Consequently, certain characteristics that would require a large amount of trading in the underlying stocks if exploited in isolation, such as the short-term-reversal characteristic (i.e., 1-month momentum), continue to be significant in the presence of transaction costs because of the trading diversification possible from combining characteristics.

Finally, we show that an investor can exploit a large set of characteristics in the presence of transaction costs to achieve an out-of-sample Sharpe ratio that is larger than that obtained by exploiting small sets of characteristics that are typically considered in prominent asset-pricing models. For instance, we find the investor achieves an out-of-sample Sharpe ratio of returns net of transaction costs around 100% larger than that from

[2]The returns of the beta and return-volatility characteristics are highly correlated over time, but while return volatility has a large (negative) mean return, beta has a negligible mean return. Thus, going long the beta characteristic allows the investor to hedge the risk of her short position in the return-volatility characteristic, without compromising its mean return.

exploiting the three traditional characteristics considered in Brandt et al. (2009) and 25% higher than that from exploiting a set of four characteristics that include investment and profitability characteristics, which are highlighted in Hou, Xue, and Zhang (2014) and Fama and French (2015). These out-of-sample results confirm that in the presence of transaction costs the cross section of stock returns is not fully explained by a small set of characteristics.[3]

Our main finding complements the result in Kozak, Nagel, and Santosh (2017) that principal-component-sparse asset-pricing models explain the cross section better than characteristic-sparse models. Our work demonstrates that transaction costs *increase* the number of characteristics that are significant for an investor, and thus, transaction costs provide an economic rationale for non-sparse characteristic-based asset-pricing models.

1.1 Characteristics and factors

We now discuss the relation between firm-specific characteristics and risk factors. Characteristics are variables that can be computed using individual-firm data, e.g., the historical stock-return volatility of a firm. Factors, on the other hand, are variables that proxy for a common source of risk, e.g., the market return. Firm-specific characteristics are related to factors because the return of a long-short portfolio based on a characteristic can be used as a proxy for an underlying unknown risk factor. For instance, Fama and French (1993) finds that returns on long-short portfolios based on size and book to market explain the cross-section of stock returns, and thus argues that these characteristics are proxies for common risk factors.

The relation between characteristics and risk factors, however, is not always clear. For instance, Daniel and Titman (1997) challenges the findings in Fama and French (1993) and claims that it is the size and book-to-market characteristics themselves rather than the covariance structure that explains the cross-section of expected stock returns. Pastor and Stambaugh (2000) explains that once model uncertainty and margin constraints are taken into account, the difference between characteristic-based and risk-factor-based

[3]This out-of-sample analysis also alleviates the data-mining concerns raised in Fama (1991), Kogan and Tian (2013), Harvey et al. (2015), Bryzgalova (2015), McLean and Pontiff (2016), Linnainmaa and Roberts (2018), and Chordia, Goyal, and Saretto (2017).

models is small from an investment perspective. In addition, Kozak, Nagel, and Santosh (2018) argues that there is no clear distinction between risk-factor pricing and behavioral asset pricing. Therefore, we consider 50 firm-specific characteristics and are agnostic about whether a particular characteristic is a proxy for the loading on a common risk factor or not; instead, we account for risk directly through the mean-variance utility of the investor.[4]

1.2 Relation to literature on asset pricing

The asset pricing literature can be classified by the following three methodologies: cross-sectional regression, time-series regression, and the stochastic discount factor approach. In this section, we discuss how our portfolio approach relates to these three approaches.

One of the most popular *cross-sectional approaches* is the Fama and MacBeth (1973) procedure, which runs a cross-sectional regression of stock returns on firm-specific characteristics at each date, and then tests the significance of the risk premia, defined as the average of the regression slopes over time. One advantage of cross-sectional regressions when studying the dimension of the cross section is that it allows one to test which characteristics are *jointly* significant; see Green et al. (2017), Freyberger et al. (2016), and Messmer and Audrino (2017). Freyberger et al. (2016), in particular, uses *nonparametric* cross-sectional regressions to study the dimension of the cross section.

The main difference between these papers and our work is that, while these papers ignore transaction costs, we focus on the effect of transaction costs on the dimension of the cross section of stock returns. Another important difference is that while cross-sectional regressions focus on mean returns, our portfolio approach accounts for *both* mean and variance of returns. Analytically, we show that our approach based on the parametric portfolios produces results that are different from those of Fama-MacBeth regressions, even in the absence of transaction costs, *unless* the covariance matrix of asset returns is

[4]In addition to the 50 firm-specific characteristics, we consider also beta (i.e., the exposure of each stock to the market-return factor) because of its importance for investment management, as shown in Frazzini and Pedersen (2014). Although beta is a risk-factor loading rather than a characteristic, for expositional ease we refer to all 51 variables as characteristics.

diagonal.[5] For example, Fama-MacBeth regressions find that while return volatility is significant, the beta characteristic is not. However, because the cross-sectional slopes of return volatility and beta are highly correlated over time, our portfolio approach, which accounts for risk, shows that the investor optimally goes short return volatility and goes long beta to reduce risk; hence, we find that both characteristics are jointly significant.[6]

The *time-series approach* regresses the return of a characteristic-based long-short portfolio on the returns of a few commonly accepted factors, such as the Fama and French (1993) and Carhart (1997) four factors. If the intercept of this time-series regression is statistically significant, then the return on the characteristic is not fully explained by the commonly accepted factors. Gibbons, Ross, and Shanken (1989) shows that testing the significance of the intercept is equivalent to testing whether the characteristic long-short portfolio can improve the Sharpe ratio of a mean-variance investor who already has access to the commonly accepted factors. Consequently, this approach captures the tradeoff between mean return and risk. Recently, Novy-Marx and Velikov (2016) develops a "generalized alpha" that extends the time-series regression to capture the impact of transaction costs.

A disadvantage of the time series approach is that it focuses on the significance of the *intercept*, and therefore, tests the significance of a single characteristic when it is added to a set of commonly accepted factors.[7] This is a limitation because the result of the statistical inference depends on the *sequence in which variables are selected*. For instance, a time-series regression of the return on the beta characteristic onto the returns

[5]The slopes in cross-sectional regressions can be estimated using either ordinary least squares (OLS) or generalized least squares (GLS). Lewellen, Nagel, and Shanken (2010) recommends using GLS because its R^2 captures the mean-variance efficiency of the model's factor-mimicking portfolios. Our analytical results show that both OLS and GLS cross-sectional regressions produce results that are different from those of our portfolio approach.

[6]Our out-of-sample analysis is also related to Lewellen (2015), which shows that Fama-MacBeth regressions provide good out-of-sample estimates of stock *expected returns*. Our out-of-sample analysis, however, focuses on estimating directly *portfolio weights*, which incorporate information about expected returns as well as risk and transaction costs.

[7]Note that one can also regress the returns of *multiple* assets with respect to the commonly accepted factors. Gibbons, Ross, and Shanken (1989) shows that in this case, testing whether the intercepts of these regressions are jointly equal to zero is equivalent to testing whether the *multiple* assets can improve the Sharpe ratio of an investor who already has access to the commonly accepted factors. The Gibbons, Ross, and Shanken test, however, does not indicate *how many* of the multiple assets are significant.

of the Fama and French (1993) and Carhart (1997) four factors finds that beta is not significant, but a time-series regression of the beta return onto these four factors *and* the return of the return-volatility characteristic finds that beta is significant. We show analytically that our approach of testing the significance of the characteristics for mean-variance parametric portfolios is *equivalent* to testing the significance of the *slopes* of a particular time-series regression in the absence of transaction costs. The advantage of our approach based on slope significance is that it allows one to consider *all* characteristics simultaneously rather than sequentially. This is crucial because both portfolio risk and transaction costs depend critically on how characteristics are combined.

There are also papers that combine elements from both cross-sectional and time-series regressions. Back, Kapadia, and Ostdiek (2015) first runs cross-sectional regressions to estimate risk premia and then runs time-series regressions of these risk premia on factors. The advantage of this procedure is that it avoids the errors-in-variables problem. Feng, Giglio, and Xiu (2017) combines the double-selection lasso in Belloni, Chernozhukov, and Hansen (2014) with two-pass regressions to estimate risk prices and evaluate the marginal contribution of a new factor with respect to an existing high-dimensional set of factors. The advantage of this approach is that it explicitly accounts for potential model-selection errors, and thus, avoids the biases associated with omitted variables. Nevertheless, the inference in the two aforementioned approaches depends on the *sequence* in which characteristics are tested, just like in time-series regressions.

Baker, Luo, and Taliaferro (2017) studies the relevance of cross-sectional and time-series regressions for a mean-variance investor. The paper shows that a risk-neutral investor facing quadratic transaction costs cares only about characteristics that are significant in cross-sectional regressions, a mean-variance investor facing no transaction costs cares only about time-series regressions, and a mean-variance investor facing quadratic transaction costs cares about both types of regressions. We sidestep the choice between cross-sectional and time-series regressions by focusing directly on the parametric portfolio problem of a mean-variance investor facing transaction costs.

Finally, the *stochastic discount factor* (SDF) approach is the most closely related to our portfolio approach because one can show that for every mean-variance efficient portfolio there is an SDF that is an affine function of the portfolio return. Ghosh, Julliard, and Taylor (2016a,b) uses a model-free robust approach to estimate the SDF that fits a cross section of asset returns by minimizing its entropy relative to the physical probability measure. Using this approach, Ghosh et al. (2016a) identifies a novel source of risk not captured by the Fama and French (1993) and Carhart (1997) factors.

Kozak, Nagel, and Santosh (2017) proposes a robust SDF by imposing an economically-motivated prior on SDF coefficients that can shrink the contributions of both low-variance principal components of characteristics as well as individual characteristics with low risk prices. They find that principal-component-sparse SDFs explain the cross section better than characteristic-sparse SDFs. The main difference between our paper and the aforementioned papers using the SDF approach is that we study the impact of *transaction costs* on the dimensionality of the cross section of stock returns. Our main finding is that transaction costs *increase* the number of characteristics that are significant for portfolio construction, and thus, transaction costs provide an economic rationale for non-sparse characteristic-based asset-pricing models.

1.3 Relation to literature on transaction costs

Several papers study the transaction costs associated with trading *particular* characteristics: Korajczyk and Sadka (2004) studies the market-impact costs associated with exploiting momentum and finds that this characteristic can be exploited on only a relatively modest scale. Novy-Marx and Velikov (2016) considers 23 different anomalies and finds that simple strategies to mitigate transaction costs significantly reduce the impact of transaction costs on the profitability of anomaly-based trading strategies. Chen and Velikov (2017) considers 135 anomalies and shows that if, in addition to transaction costs, one takes into account the post-publication decay, the profitability of anomaly-based trading strategies is substantially diminished. The aforementioned papers use publicly available datasets to estimate the costs of an average investor. Frazzini, Israel, and Moskowitz (2015), using proprietary data from an institutional money manager, finds that the trading costs associated with exploiting size, momentum, and book to market can be quite small for large institutional investors, and that these managers can exploit these characteristics to a much larger extent than previously thought.

Very few papers consider the transaction costs associated with trading multiple characteristics *jointly*. Hanna and Ready (2005) shows that the long-short stock-selection strategy considered in Haugen and Baker (1996), which is based on a combination of more than 50 characteristics, does not outperform the portfolios based solely on book to market and momentum once transaction costs are taken into account. Hand and Green (2011) considers parametric portfolios with three accounting-based characteristics in addition to size, book to market, and momentum and finds that accounting-based characteristics can improve performance substantially, but transaction costs reduce the benefits from exploiting accounting-based characteristics. We show that by combining a large number of characteristics the investor can alleviate the impact of transaction costs significantly because of trading diversification.

Other papers have also found that combining characteristics helps to reduce transaction costs. For instance, Frazzini, Israel, and Moskowitz (2015) considers size, value, and momentum and explains that "value and momentum trades tend to offset each other, resulting in lower turnover which has real transaction costs benefits." Barroso and Santa-Clara (2015) considers currency portfolios based on six characteristics and explains that "transaction costs depend crucially on the time-varying interaction between characteristics." Novy-Marx and Velikov (2016) studies "filtering," a cost mitigation technique that allows investors trading one strategy to opportunistically take small positions in another at effectively negative trading costs. We build on these three papers and show how to quantify precisely the reduction in transaction costs when an investor *optimally* rebalances a portfolio based on a *large* number of characteristics.

The rest of this paper is organized as follows. Section 2 describes the data. Section 3 explains how we extend the methodology of parametric portfolios. Section 4 characterizes the benefits from trading diversification. Sections 5 and 6 study how many characteristics matter in the absence and presence of transaction costs, respectively. Section 7 evaluates the out-of-sample performance of different portfolios. Section 8 concludes. Appendix A studies how our portfolio approach relates to the cross-sectional and time-series regression approaches. Appendix B contains proofs for all analytical results in the manuscript.

The Internet Appendix investigates the robustness of our main finding that transaction costs increase the number of significant characteristics to: considering quadratic instead of proportional transaction costs, excluding microcaps, applying the reality check in White (2000), expanding our dataset to consider also characteristics with a large number of missing observations, different subperiods, risk-aversion, and using different methods to standardize firm characteristics. In addition, the Internet Appendix checks the robustness of our out-of-sample results to: firm size, shortsale constraints, and the constraint on maximum turnover.

2 Data

We combine U.S. stock-market information from CRSP, Compustat, and I/B/E/S, covering the period from January 1980 to December 2014. We start by compiling data on the 100 firm-specific characteristics considered in Green et al. (2017),[8] but drop characteristics with a large proportion of missing observations.[9] Specifically, we drop characteristics with more than 5% of missing observations for more than 5% of firms with CRSP returns available for the entire sample from 1980 to 2014. In addition, we drop characteristics without any observations for more than 1% of these firms. Table 1 lists the resulting 51 characteristics together with their definitions, the name of the author(s) who identified them, and the date and journal of publication.

Our initial database contains every firm traded on the NYSE, AMEX, and NASDAQ exchanges. We then remove firms with negative book-to-market ratios. As in Brandt et al. (2009), we also remove firms below the 20th percentile of market capitalization because these are very small firms that are difficult to trade. Our final dataset

[8]As in Green et al. (2017), when constructing monthly characteristics at time t, we assume that annual (quarterly) accounting data is available at the end of month $t-1$ if the firm's fiscal year ended at least six (four) months earlier.

[9]To ensure that our results are reliable, in our main analysis we have considered only characteristics with a small proportion of missing observations. However, in Section IA.4 of the Internet Appendix, we run our experiments using all 100 characteristics and find that our main finding that transaction costs increase the dimension of the cross section is robust.

contains 51 firm-specific characteristics for a total of 17,930 firms of which an average of 3,071 firms have return data in a particular month.

As in Green et al. (2017), we cross-sectionally winsorize each characteristic; that is, we replace extreme observations that are beyond a certain threshold with the value of the threshold. Specifically, we set equal to the third (first) quartile plus (minus) three times the interquartile range any observations that are above (below) this threshold.

Finally, as in Brandt et al. (2009), we standardize each characteristic so that it has a cross-sectional mean of zero and standard deviation of one. The resulting standardized characteristic is a long-short portfolio that goes long stocks whose characteristic is above the cross-sectional average, and short stocks whose characteristic is below the cross-sectional average.

3 Methodology

To study how many characteristics matter jointly from a portfolio perspective, this section explains how we extend the parametric portfolio methodology in Brandt et al. (2009). We also describe a *screen and clean* method to test whether the parametric portfolio weights corresponding to the different characteristics are significantly different from zero. Appendix A compares analytically and empirically our methodological approach based on the parametric portfolios with the cross-sectional (Section A.1) and time-series (Sections A.2 and A.3) regression approaches.

3.1 Mean-variance parametric portfolios

Parametric portfolios use a set of firm-specific characteristics to *tilt* the benchmark portfolio toward stocks that help to increase the investor's utility. The portfolios are obtained by adding to the benchmark portfolio a linear combination of long-short portfolios obtained by standardizing K firm-specific characteristics cross sectionally so that they have zero mean and unit standard deviation. The resulting *parametric portfolio* at time t, $w_t(\theta) \in \mathbb{R}^{N_t}$, can be written as

$$w_t(\theta) = w_{b,t} + (x_{1,t}\theta_1 + x_{2,t}\theta_2 + \ldots + x_{K,t}\theta_K)/N_t, \tag{1}$$

where $w_{b,t} \in \mathbb{R}^{N_t}$ is the *benchmark portfolio* at time t, $x_{k,t} \in \mathbb{R}^{N_t}$ is the long-short portfolio obtained by standardizing the kth firm-specific characteristic at time t, θ_k is the weight of the kth characteristic in the parametric portfolio, and N_t is the number of firms at time t.[10] As in Brandt et al. (2009), we consider a portfolio that is fully invested in risky assets.[11] The parametric portfolio can also be written in compact matrix notation by defining $X_t \in \mathbb{R}^{N_t \times K}$ to be the matrix whose kth column is $x_{k,t}$:

$$w_t(\theta) = w_{b,t} + X_t\theta/N_t, \tag{2}$$

where $\theta \in \mathbb{R}^K$ is the *parameter vector*, whose kth component is the weight of the kth characteristic θ_k, and $X_t\theta/N_t$ is the characteristic portfolio at time t.

The return of the parametric portfolio at time $t+1$, which we denote as $r_{p,t+1}(\theta)$, can thus be rewritten as

$$r_{p,t+1}(\theta) = \; w_{b,t}^{\top} r_{t+1} + \theta^{\top} X_t^{\top} r_{t+1}/N_t = \; r_{b,t+1} + \theta^{\top} r_{c,t+1}, \tag{3}$$

where $r_{t+1} \in \mathbb{R}^{N_t}$ is the return vector at time $t+1$, $r_{b,t+1} = w_{b,t}^{\top} r_{t+1}$ is the benchmark portfolio return at time $t+1$, and $r_{c,t+1} = X_t^{\top} r_{t+1}/N_t$ is the *characteristic return vector* at time $t+1$, which contains the returns of the long-short portfolios corresponding to the K characteristics scaled by the number of firms N_t.[12] Equation (3) shows that the parametric-portfolio return is the benchmark-portfolio return plus the return of the characteristic portfolio.

We assume that the investor optimizes mean-variance utility. The advantages of mean-variance utility, as we will show below, are that it allows us to identify the marginal contribution of each characteristic to the investor's utility and to compare analytically

[10] The weights of the characteristics in the parametric portfolio are scaled by the number of stocks N_t so that they are meaningful for the case with a varying number of stocks. Without this scaling parameter, increasing the number of stocks while keeping the weights fixed would result in more aggressive portfolio allocations.

[11] Consequently, the parametric portfolio weights on the stocks need to sum to one. Because the weights on the stocks in the long-short portfolios sum to zero, this implies that the parametric weight on the benchmark portfolio must equal one.

[12] Note that we use only lagged values of characteristics to build portfolios; thus, the returns of the characteristic portfolio formed at time t, $X_t\theta/N_t$ are evaluated using the return at time $t+1$; that is, $\theta^{\top} X_t^{\top} r_{t+1}/N_t$.

the parametric portfolio weights to the results from time-series and cross-sectional regressions.[13] In particular, we assume the investor solves the following problem:

$$\min_{\theta} \quad \frac{\gamma}{2}\text{var}_t[r_{p,t+1}(\theta)] - E_t[r_{p,t+1}(\theta)], \tag{4}$$

where γ is the risk-aversion parameter and $\text{var}_t[r_{p,t+1}(\theta)]$ and $E_t[r_{p,t+1}(\theta)]$ are the variance and mean of the parametric portfolio return, respectively.

Given T historical observations of returns and characteristics, the following proposition shows that the parametric portfolio problem can be formulated as a tractable quadratic optimization problem.

Proposition 1. *The mean-variance parametric portfolio problem in* (4) *is equivalent to*

$$\min_{\theta} \quad \underbrace{(\gamma/2)\theta^\top \widehat{\Sigma}_c \theta}_{var(char)} + \underbrace{\gamma\theta^\top \widehat{\sigma}_{bc}}_{cov(bench)} - \underbrace{\theta^\top \widehat{\mu}_c}_{mean}, \tag{5}$$

where $\widehat{\Sigma}_c$ and $\widehat{\mu}_c$ are the sample covariance matrix and mean of the characteristic-return vector r_c, and $\widehat{\sigma}_{bc}$ is the sample vector of covariances between the benchmark portfolio return r_b and the characteristic-return vector r_c.

Proposition 1 shows that the mean-variance parametric portfolio problem is to find the parameter vector θ that offers the optimal tradeoff between the variance of the characteristic portfolio return, $(\gamma/2)\theta^\top \widehat{\Sigma}_c \theta$; the covariance of the characteristic portfolio return with the benchmark portfolio return, $\gamma\theta^\top \widehat{\sigma}_{bc}$; and the mean characteristic portfolio return, $\theta^\top \widehat{\mu}_c$.

3.2 Transaction costs

As in Brandt et al. (2009) and Hand and Green (2011), we consider an investor who faces proportional transaction costs that decrease with firm size and over time. Proportional transaction costs are a reasonable assumption for the average investor, as explained in Novy-Marx and Velikov (2016) and Chen and Velikov (2017). Nevertheless, Section IA.1

[13] We have run our empirical analysis also for power utility, as in Brandt et al. (2009), and the main insights are unchanged.

of the Internet Appendix shows that our main findings are robust to using quadratic transaction costs that are often used to model the price impact costs of large investors; see, for instance, Korajczyk and Sadka (2004).

Let the proportional transaction cost parameter for the ith stock at time t be

$$\kappa_{i,t} = y_t z_{i,t}, \tag{6}$$

where y_t and $z_{i,t}$ capture the variation of the transaction cost parameter with time and firm size, respectively. Following Brandt et al. (2009) and Hand and Green (2011), we assume y_t decreases linearly from 3.3 in January 1980 to 1.0 in January 2002, and after that it remains at 1.0.[14] We set $z_{i,t} = 0.006 - 0.0025 \times me_{i,t}$, where $me_{i,t}$ is the market capitalization of firm i at time t after being normalized cross-sectionally so that it takes values between zero and one. This functional form results in proportional transaction costs in the 1980s of about 180 basis points for the smallest firms and 100 basis points for the largest firms, and after 2002 of about 60 basis points for the smallest firms and 35 basis points for the largest firms.

Given T historical observations of returns and characteristics, the transaction cost associated with implementing the parametric portfolios can be estimated as

$$\text{TC}(\theta) = \frac{1}{T-1} \sum_{t=1}^{T-1} \|\Lambda_t (w_{t+1}(\theta) - w_t^+(\theta))\|_1, \tag{7}$$

where the transaction cost matrix at time t, Λ_t, is the diagonal matrix whose ith diagonal element contains $\kappa_{i,t}$, $\|a\|_1 = \sum_{i=1}^N |a_i|$ is the 1-norm of the N-dimensional vector a, and w_t^+ is the portfolio before rebalancing at time $t+1$, that is,

$$w_t^+ = (w_{b,t} + X_t \times \theta / N_t) \circ (e_t + r_{t+1}), \tag{8}$$

where e_t is the N_t-dimensional vector of ones and $x \circ y$ is the Hadamard or component-wise product of vectors x and y. Combining (5) and (7), the mean-variance parametric portfolio problem *with* transaction costs is

$$\min_\theta \quad \underbrace{(\gamma/2)\theta^\top \widehat{\Sigma}_c \theta}_{var(char)} \; + \; \underbrace{\theta^\top \gamma \widehat{\sigma}_{bc}}_{cov(bench)} \; - \; \underbrace{\theta^\top \widehat{\mu}_c}_{mean} \; + \; \underbrace{\text{TC}(\theta).}_{transaction\ costs} \tag{9}$$

[14] Brandt et al. (2009) defines y_t so that transaction costs in 1974 are four times larger than in 2002. Therefore, if we decrease y_t uniformly until 1980, we would have a starting value for y_t approximately equal to 3.3. See also French (2008, p. 1553) for a discussion of the time evolution of transaction costs.

3.3 Understanding why a characteristic matters

To understand why particular characteristics are significant from a portfolio perspective, it is useful to consider the first-order optimality conditions for the mean-variance parametric portfolio problem with transaction costs, that is, the problem in (9).

By decomposing the variance of the characteristic portfolio return, $\theta^\top \widehat{\Sigma}_c \theta$, into a term associated with the characteristic *own-variances*, $\theta^\top \text{diag}(\widehat{\Sigma}_c)\theta$, and a term associated with the characteristic covariances, $\theta^\top(\widehat{\Sigma}_c - \text{diag}(\widehat{\Sigma}_c))\theta$, where $\text{diag}(\widehat{\Sigma}_c)$ is the diagonal matrix whose kth diagonal element contains the variance of the kth characteristic return, the mean-variance parametric portfolio problem with transaction costs can be rewritten as

$$\min_{\theta} \quad \underbrace{(\gamma/2)\theta^\top \text{diag}(\widehat{\Sigma}_c)\theta}_{own-var(char)} + \underbrace{(\gamma/2)\theta^\top(\widehat{\Sigma}_c - \text{diag}(\widehat{\Sigma}_c))\theta}_{cov(char)} + \underbrace{\theta^\top \gamma \widehat{\sigma}_{bc}}_{cov(bench)} - \underbrace{\theta^\top \widehat{\mu}_c}_{mean} + \underbrace{\text{TC}(\theta).}_{tran.\ costs} \tag{10}$$

Note that the transaction cost term $\text{TC}(\theta)$ is a convex function of the parameter θ, but it is not differentiable at values of θ for which $w_{i,t+1}(\theta) = w^+_{i,t}(\theta)$ for some i and t. Therefore, the optimality conditions must be formally defined in terms of the subdifferential $\partial\text{TC}(\theta)$.[15]

Proposition 2. *The first-order optimality conditions for problem* (10) *are*

$$0 \in \underbrace{\gamma\, diag(\widehat{\Sigma}_c)\theta}_{own-var(char)} + \underbrace{\gamma(\widehat{\Sigma}_c - diag(\widehat{\Sigma}_c))\theta}_{cov(char.)} + \underbrace{\gamma\widehat{\sigma}_{bc}}_{cov(bench.)} - \underbrace{\widehat{\mu}_c}_{mean} + \underbrace{\partial\, TC(\theta)}_{\text{costs}}, \tag{11}$$

where the kth component of the subdifferential of the transaction cost term is

$$\partial_{\theta_k} TC(\theta) = \frac{1}{T-1}\sum_{t=1}^{T-1} sign(w_{t+1}(\theta) - w^+_t(\theta))^\top(\Lambda_t[(X_{t+1})_{\bullet,k} - (X_t)_{\bullet,k} \circ (e_t + r_{t+1})]), \tag{12}$$

where $A_{\bullet,k}$ is the kth column of matrix A, and

$$sign(w_{i,t+1}(\theta) - w^+_{i,t}(\theta)) = \begin{cases} +1 & if \quad w_{i,t+1}(\theta) > w^+_{i,t}(\theta), \\ -1 & if \quad w_{i,t+1}(\theta) < w^+_{i,t}(\theta), \\ [-1,1] & if \quad w_{i,t+1}(\theta) = w^+_{i,t}(\theta). \end{cases} \tag{13}$$

[15]See Rockafellar (2015) for an extensive treatment of subdifferentials.

The first-order optimality conditions in (11) allow us to identify the *marginal* contribution of each characteristic to the investor's mean-variance utility. Specifically, the kth component of the right-hand side in (11) is the marginal contribution of the kth characteristic to the parametric portfolio mean-variance utility; that is, the marginal change to mean-variance utility associated with a unit increase in the weight that the parametric portfolio assigns to the kth characteristic. Moreover, the five terms on the right-hand side of (11) are: the marginal contributions of the kth characteristic to the characteristic own-variance, $\gamma \text{diag}(\widehat{\Sigma}_c)\theta$; the characteristic covariance with other characteristics, $\gamma(\widehat{\Sigma}_c - \text{diag}(\widehat{\Sigma}_c))\theta$; the covariance between the characteristic and benchmark portfolios, $\gamma\widehat{\sigma}_{bc}$; the characteristic portfolio mean, $-\widehat{\mu}_c$; and the transaction cost, $\partial\, \text{TC}(\theta)$.

Finally, to gauge the size of the trading diversification benefit associated with combining characteristics, it will be useful to compute the marginal contribution to transaction costs of trading the kth characteristic in isolation (that is, without the benchmark or any other characteristics), which is

$$\partial^{iso}_{\theta_k} \text{TC}(\theta) = \frac{1}{T-1} \sum_{t=1}^{T-1} \|\Lambda_t[(X_{t+1})_{\bullet,k} - (X_t)_{\bullet,k} \circ (e_t + r_{t+1})]\|_1. \tag{14}$$

Straightforward algebra shows that the marginal contribution to transaction costs of trading the kth characteristic in isolation, given in (14), is larger in general than that of trading it in combination, given in (12).

3.4 The regularized parametric portfolios

Although the parametric portfolios only require estimating one parameter per characteristic, we will be considering a large number of characteristics. To deal with such a high-dimensional setting, we propose a new class of parametric portfolios, which we term *regularized parametric portfolios*. These portfolios are obtained by imposing a lasso[16]

[16]The term lasso originated as the acronym for *least absolute shrinkage and selection operator*. The lasso was originally proposed in Tibshirani (1996) in the context of statistical learning and has become a prominent tool in the age of machine learning. See Hastie, Tibshirani, and Wainwright (2015) for an in-depth treatment of the lasso and DeMiguel, Garlappi, Nogales, and Uppal (2009a) for a Bayesian interpretation of the lasso constraint in the context of portfolio choice.

constraint on the parametric portfolio to achieve two goals. First, the lasso constraint helps to avoid overfitting, reducing the impact of estimation error. Second, the lasso constraint is a variable-selection method that results in parametric portfolios where only the relevant characteristics receive a nonzero parameter. This is exploited by the screen-and-clean significance test described in Section 3.5 to characterize the dimension of the cross section.

In contrast, the minimum-entropy SDF approach used in Ghosh et al. (2016a) results in SDFs that assign a nonzero weight to every characteristic, and thus, it is not suitable to study the dimension of the cross section. The elastic-net SDF approach in Kozak et al. (2017) shrinks the contributions to the SDF of *both* low-variance principal components of characteristics as well as individual characteristics with low risk prices. Shrinking *only* the contributions of low-risk-price characteristics allows us to study how transaction costs affect the number of characteristics that are significant for an investor.

The regularized parametric portfolios are obtained by solving problem (9) subject to the lasso constraint, that is, by solving

$$\min_{\theta} \quad \frac{\gamma}{2}\theta^\top \widehat{\Sigma}_c \theta + \theta^\top \gamma \widehat{\sigma}_{bc} - \theta^\top \widehat{\mu}_c + \mathrm{TC}(\theta), \tag{15}$$

$$\text{s.t.} \quad \|\theta\|_1 \leq \delta, \tag{16}$$

where $\|\theta\|_1 = \sum_{k=1}^{K} |\theta_k|$ is the 1-norm of the parameter vector, and δ is the threshold parameter. To gain intuition about δ, note that for $\delta = \infty$, we recover the standard parametric portfolios, and for $\delta = 0$, we recover the benchmark portfolio. Thus, as one increases δ, the regularized parametric portfolios move from the benchmark portfolio toward the unregularized parametric portfolio.

3.5 Testing the significance of characteristics considered jointly

We now explain how we test whether the parametric portfolio weights corresponding to the different characteristics are significantly different from zero. Chatterjee and Lahiri (2011) shows that it is challenging to carry out statistical inference in the presence of a lasso constraint, such as the one imposed on the regularized parametric portfolios. To address this issue, we use a two-stage *screen-and-clean* method similar to the methods proposed in Wasserman and Roeder (2009), Meinshausen and Yu (2009), and Meinshausen, Meier, and Buhlmann (2009). In the first stage, we *screen* the characteristics by using the regularized parametric portfolios. Specifically, we use five-fold cross-validation, as explained in Hastie et al. (2015, Section 2.3), to select the lasso threshold δ that optimizes the mean-variance criterion.[17] For the resulting optimal lasso threshold, we compute the regularized parametric portfolios and "screen" or remove any characteristics with a zero parameter.

In the second stage, we *clean* the characteristics that were not removed in the first stage. That is, we compute the parametric portfolios using the characteristics that were not removed in the first stage, but now *without a lasso constraint*, thus circumventing the concerns highlighted in Chatterjee and Lahiri (2011), and apply a bootstrap method to establish which of these characteristics have parametric portfolio weights that are significantly different from zero.[18] Specifically, we apply the percentile-interval method described in Efron and Tibshirani (1993, Section 13.3) and Hastie et al. (2015, Section 6.2) to establish the significance of the selected characteristics.[19]

[17]In particular, we divide the sample of monthly observations into five intervals of equal length. For j from 1 to 5, we remove the jth-interval from the sample and use the remaining sample returns to compute the regularized parametric portfolio for several values of δ. We then evaluate the return of the resulting portfolios on the jth-interval. After completing this process for each of the five intervals, we have out-of-sample portfolio returns for the entire sample for each value of δ. Finally, we compute the mean-variance utility of these out-of-sample returns and select the value of δ that corresponds to the portfolio with the largest mean-variance utility.

[18]Barroso and Santa-Clara (2015) uses a one-stage bootstrap method essentially equivalent to our "clean" stage to test the statistical significance of the different characteristics in a *currency* parametric portfolio. This method is appropriate in the context of that paper because it considers only *five* characteristics and thus does not require a regularization method like lasso.

[19]In detail, we first generate $1,000$ bootstrap samples from the original dataset using sampling with replacement. Second, we estimate the optimal parametric portfolio for the remaining characteristics and for each bootstrap sample. Finally, we declare as significant at the 5% level those characteristics whose estimated parameter is strictly positive (strictly negative) for at least 95% of the bootstrap samples, and compute the p-value as the proportion of bootstrap samples for which the parameter is less than or equal to zero (greater than or equal to zero). Note that the parametric portfolio approach relies on the assumption that, conditional on firm-specific characteristics, stock returns are independently and identically distributed (iid). Therefore, we employ an *iid* bootstrap method. Nevertheless, to gauge the importance of the iid assumption, we have repeated the tests using the stationary bootstrap in Politis and Romano (1994), which takes serial dependence into account, and we have found that the results are robust. In particular, we have run the (nonstudentized) stationary bootstrap with expected block sizes of two and six months, and we have found that this does not affect the significance results.

Other approaches have been considered in the literature to identify characteristics that are jointly relevant when considered simultaneously. For instance, Freyberger et al. (2016) and Messmer and Audrino (2017) use a refinement of the lasso approach known as *adaptive lasso* to select characteristics in the context of cross-sectional regressions. The adaptive lasso is complementary to our approach as it could be used as the variable-selection method for the screen stage of our screen-and-clean approach.

In contrast, one might think of using a *sequential* bootstrap method to test the significance of adding one more characteristic to an existing parametric portfolio. This approach would be similar in spirit to the methodology proposed in Harvey and Liu (2018) in the context of sequential factor selection. However, from a portfolio perspective a sequential significance test would not capture the risk and trading-diversification benefits from adding *several* characteristics simultaneously. This is crucial because both risk and transaction costs depend critically on how characteristics are *combined*.[20]

4 Trading diversification

We now characterize analytically and empirically the magnitude of the trading diversification benefits obtained by combining characteristics. We do this by comparing the average trading volume (turnover) required to exploit characteristics in combination with that required to exploit them in isolation. To simplify the exposition, in this section we focus on the case where the investor holds an *equally weighted* portfolio of the characteristics, but all results can be extended to the case with a generic portfolio of characteristics. Note that the reduction of turnover that we characterize in this section will result in a reduction of transaction costs *regardless* of the particular manner in which transaction costs are modeled. Then, in Sections 5 and IA.1, we show that in the presence of proportional

[20]In results not reported to conserve space, we find that our main finding that transaction costs increase the number of significant characteristics is robust to the choice of significance test. The reason for this is that our main insight is obtained by *comparing* the number of significant characteristics for the cases with and without transaction costs. We find that, independently of the test or data sample used, trading diversification results in an increased number of jointly significant characteristics for the case with transaction costs.

and quadratic transaction costs, respectively, the benefits of trading diversification lead to an increase in the number of characteristics that are jointly significant for an investor.

4.1 Analytical results

Proposition 3 below characterizes the reduction in turnover obtained by combining characteristics. The intuition underlying this proposition is that, just as we get diversification of risk when we combine stocks, we get diversification of trading when we combine characteristics. To see this, note that rebalancing the long-short portfolio associated with *each* characteristic requires trading in the *same* set of underlying stocks. Thus, exploiting multiple characteristics allows one to cancel out some of the trades in the underlying stocks required to rebalance the characteristic long-short portfolios. For instance, if to rebalance a characteristic long-short portfolio we need to buy a particular stock, whereas to rebalance another characteristic we need to sell the same stock, then the net amount of trading required to exploit these two characteristics in combination will be smaller than that required to exploit them in isolation.

Proposition 3. *Assume that the trades in the ith stock required to rebalance $K > 1$ different characteristics, that is, the quantities*

$$trade_{i,k} = (X_{t+1})_{i,k} - (X_t)_{i,k}(1 + r_{i,t+1}), \qquad k = 1, 2, \ldots, K \tag{17}$$

are jointly distributed as a multivariate Normal distribution with zero mean and positive-definite covariance matrix Ω. Then:

1. *The ratio of the average trading volume (turnover) in the ith stock required to rebalance an equally weighted portfolio of the K characteristics to that required to rebalance the K characteristics in isolation is*

$$\frac{turnover(trade_i^{ew})}{turnover(trade_i^{iso})} = \frac{\sqrt{e^\top \Omega e}}{\sum_{k=1}^{K} \sqrt{\Omega_{kk}}} < 1,$$

where $e \in \mathbb{R}^K$ is the vector of ones, Ω_{kk} is the variance of $trade_{i,k}$,

$$turnover(trade_i^{ew}) = E\left[\frac{1}{K}\left|\sum_{k=1}^{K} trade_{i,k}\right|\right], \quad and$$

$$turnover(trade_i^{iso}) = E\left[\frac{1}{K}\sum_{k=1}^{K}|trade_{i,k}|\right].$$

2. *If, in addition, the covariance matrix Ω is symmetric with respect to all K characteristics, that is, if the variances and correlations between the trades in the ith stock required to rebalance the K different characteristics are all equal to σ^2 and ρ, respectively, then*[21]

$$\frac{turnover(trade_i^{ew})}{turnover(trade_i^{iso})} = \sqrt{\frac{1+\rho(K-1)}{K}} < 1. \tag{18}$$

3. *If, in addition, the correlations between the trades in the ith stock required to rebalance the K different characteristics are all zero ($\rho = 0$), then*

$$\frac{turnover(trade_i^{ew})}{turnover(trade_i^{iso})} = \frac{1}{\sqrt{K}} < 1.$$

Part 1 of Proposition 3 shows that, provided the covariance matrix of the rebalancing trades is positive definite (and thus, the rebalancing trades between some of the characteristics are not perfectly correlated), combining characteristics will result in trading diversification and a reduction in turnover. Also, Part 2 of Proposition 3 shows that the benefits of trading diversification increase with the number of characteristics and decrease with the correlation between the rebalancing trades of different characteristics.

4.2 Empirical results

We now evaluate empirically the benefits from trading diversification. Figure 1 compares the monthly turnover required to exploit the 51 characteristics in isolation with that required to exploit them in an equally weighted combination.[22] The figure shows that the

[21] Note that in (18) the term $1+\rho(K-1)$ is strictly positive because of the assumption that Ω is positive definite.

[22] For this section only, we have adjusted the sign of every characteristic so that its associated long-short portfolio produces positive average returns. The marginal contributions to turnover are computed using Equation (12) for the case where the transaction cost matrix Λ_t is replaced by the identity matrix and for an equally weighted portfolio of the 51 characteristics without the benchmark; that is, $w_t = X_t e/(51N_t)$, where e is the vector of ones.

trading diversification benefits of combining characteristics are large empirically. While the average monthly turnover required to exploit the 51 characteristics in isolation is 24.09%, the turnover required to exploit an equally weighted combination of them is only 6.71%; that is, trading diversification delivers a 72.15% reduction in turnover.[23]

Note that this 72.15% reduction in turnover is similar in magnitude to that predicted by Part 3 of Proposition 3 for the symmetric case with *zero correlation* between rebalancing trades across characteristics: $1 - 1/\sqrt{K} = 1 - 1/\sqrt{51} \approx 86\%$. Indeed, Figure 2 gives a heatmap of the correlations between the rebalancing trades for the 51 characteristics for a particular stock and shows that many of the correlations are close to zero.[24] Moreover, we find that the average correlation between rebalancing trades across the 51 characteristics and for the entire universe of stocks is 5.47%, not very different from zero. This explains why the empirical benefits from trading diversification are so large and in line with those predicted by Part 3 of Proposition 3.

In this section, we have shown analytically and empirically that combining characteristics in an equally weighted portfolio results in a substantial reduction in *turnover* compared to trading them in isolation. In the next two sections, we show that combining characteristics optimally in a parametric portfolio also results in a substantial reduction in *transaction costs*.

5 How many characteristics matter without costs?

This section studies how many characteristics matter jointly from a portfolio perspective *in the absence of transaction costs* and Section 6 studies the effect of transaction costs.

We apply the screen-and-clean method described in Section 3.5 to test the significance of the characteristics. We consider a risk-aversion parameter $\gamma = 5$, use the value-weighted portfolio as the benchmark, and run the screen-and-clean test on the 319

[23]In fact, Figure 1 shows that the turnover required to exploit each characteristic in isolation (blue bars) is much larger than the marginal contribution to turnover of each characteristic in an equally weighted combination (yellow bars). Most strikingly, for the volatility of share turnover (*std_turn*) characteristic, the marginal contribution to turnover in an equally weighted combination is *negative*, implying that including this characteristic results in an absolute reduction in turnover of the portfolio.

[24]We have produced heatmaps for several stocks as well as the heatmap for the average correlations across stocks and the insights are similar.

monthly observations from May 1988 to December 2014.[25] The results from the "screen" stage, not reported to conserve space, establish that the optimal lasso threshold is $\delta = 25$, and only 10 characteristics survive the screening. We then run the "clean" stage test, without lasso regularization, for these 10 characteristics and find that, in the absence of transaction costs, six characteristics are significant.

Table 2 reports the significance and marginal contributions of each characteristic in the parametric portfolios. For each characteristic, the last four columns of the table give the marginal contribution of the characteristic to: (i) the characteristic own-variance, (ii) the covariance of the characteristic with the other characteristics in the portfolio, (iii) the covariance of the characteristic with the benchmark portfolio, and (iv) the characteristic mean. Marginal contributions that drive the characteristic to be nonzero are in blue sans serif font, and marginal contributions that drive the characteristic toward zero are in red *italic* font.[26]

We observe from Table 2 that, in the absence of transaction costs, six characteristics are significant. Five are significant at the 5% confidence level: unexpected quarterly earnings (*sue*), return volatility (*retvol*), asset growth (*agr*), 1-month momentum (*mom1m*), and gross profitability (*gma*); and one characteristic, beta, is significant at the 10% level. From a return-prediction perspective, Hou et al. (2014) and Fama and French (2015) show that four and five variables, respectively, are enough to predict expected returns. Our result confirms that, in the absence of transaction costs, a small number of characteristics are sufficient also from a portfolio perspective.

The marginal contributions reported in Table 2 show that the five characteristics significant at the 5% level matter because they help to reduce the risk of the portfolio

[25]Although our dataset covers the period from January 1980 to December 2014, we drop the first 100 months so that the significance test is run on the exact same sample as the out-of-sample analysis in Section 7. Also, in Section IA.6 of the Internet Appendix, we consider other values of risk-aversion: $\gamma = 2$ and 10.

[26]Note that for characteristics with a positive parametric portfolio weight, negative (positive) marginal contributions help to decrease (increase) the objective function in the minimization problem (9) and thus increase (decrease) the investor's mean-variance utility. Therefore for characteristics with positive parametric portfolio weights, negative (positive) marginal contributions are in blue sans serif font (red *italic* font). The opposite color and font convention applies to characteristics with negative parametric portfolio weights.

of characteristics *and* increase its mean return.[27] In contrast, the beta characteristic is significant at the 10% level *only* because of its ability to reduce the risk of the portfolio of characteristics. To see this, note that Table 2 shows that, consistent with the findings in the existing literature (see Black (1993)), the marginal contribution of *beta* to the portfolio's mean return is very small. However, the beta return has a large negative covariance with the returns of the other characteristics (marginal contribution −0.01381), and this is what makes it relevant from a portfolio perspective. This is illustrated in Figure 3, which depicts the marginal contributions of the six significant characteristics, and shows that *beta* has a large marginal contribution to the covariance with the other characteristics that helps to reduce the overall portfolio risk.

Table 2 also explains why size, book to market, and momentum are *not* significant when evaluated from a portfolio perspective. For instance, 12-month momentum (*mom12m*) and book to market (*bm*) are not significant, even though their expected returns are large, because their returns have a very large positive covariance with the returns of the other characteristics in the portfolio. In contrast, market capitalization (*mve*) has only a small mean return, consistent with findings in the literature (see Asness, Frazzini, Israel, Moskowitz, and Pedersen (2018)), and hence, although *mve* helps to diversify the characteristic portfolio, the risk reduction is not sufficient to make it significant.

As discussed above, the contribution of characteristics to portfolio risk plays an important role. Thus, the correlations between the characteristic returns matter from a portfolio perspective. Table 3 reports the correlation matrix for the returns of the six significant characteristics and the three characteristics considered in Brandt et al. (2009): size, book to market, and momentum.

We first observe from Table 3 that the returns of the size, book to market, and momentum characteristics are not highly correlated, with their correlation coefficients being smaller than 20%. On the other hand, the returns of the six significant characteristics we identify are more highly correlated. To understand why these characteristics with highly

[27]For instance, return volatility has large positive mean return (marginal contribution 0.00323) and negative return covariance with the other characteristics (marginal contribution 0.02914).

correlated returns are jointly significant for portfolio choice, consider the case of return volatility and beta. The returns of these two characteristics are highly positively correlated (93%), but the mean return of beta is very small. As a consequence, the investor optimally goes long the beta characteristic to hedge the risk of her short position in the return-volatility characteristic, while preserving most of its mean return. The benefit of this strategy is illustrated in Panel (a) of Figure 4, which shows the cumulative returns of a blended strategy that assigns a −50% weight to return volatility and a +50% weight to beta. This blended strategy has large cumulative returns and very low volatility.[28]

Asness, Moskowitz, and Pedersen (2013) finds that the returns of value and momentum are negatively correlated and a blended strategy of these two characteristics performs well. We compare the return volatility and beta blended strategy with the value and momentum blended strategy. Panel (b) in Figure 4 shows the cumulative returns of these two blended strategies, where we have scaled them so that they have the same volatility. We find that the return-volatility and beta blend attains a cumulative return of 110%, whereas the value and momentum blend attains a cumulative return of around 80%.

Summarizing, we find that, in the absence of transaction costs, only six characteristics are significant and that risk diversification plays an important role in determining which characteristics are significant. We now study the role of trading diversification.

6 What is the effect of transaction costs?

In this section, we examine how transaction costs affect the dimension of the cross section of stock returns. As explained in Section 3.2, we consider an investor who faces propor-

[28]Our finding that, despite the high correlation between the return volatility and beta characteristics, the return-volatility characteristic commands a much higher average return than beta is consistent with results in the existing literature. As explained in Bali et al. (2016), return volatility and idiosyncratic volatility are very similar in the cross section. Therefore, the high average return of the return-volatility characteristic can be traced back to the high average return of the idiosyncratic-volatility characteristic, which is documented in Ang, Hodrick, Xing, and Zhang (2006). Moreover, Bali et al. (2016, Table 15.7) shows that the idiosyncratic risk characteristic commands a high average return mostly when computed from daily data over short horizons, which is how return volatility is computed in our analysis. Beta, on the other hand, is computed from weekly returns over the past three years, and thus delivers much lower average returns; for a detailed analysis of the relation between beta and idiosyncratic volatility, see Liu, Stambaugh, and Yuan (2018).

tional transaction costs that decrease with firm size and over time, as in Brandt et al. (2009) and Hand and Green (2011). Section IA.1 of the Internet Appendix shows that our main findings are robust to transaction costs that are quadratic instead of proportional.

Intuitively, one may expect that in the presence of transaction costs *fewer* characteristics would be significant because transaction costs can only erode the benefits from exploiting characteristics. Indeed, we find that this is the case if one were to consider each characteristic *individually*: 21 characteristics are individually significant in the absence of transaction costs, but only 14 in the presence transaction costs.[29] However, when considered *jointly*, we find that the number of characteristics that are jointly significant at the 5% level *increases* from five in the absence of transaction costs to 15 in the presence of proportional transaction costs.[30]

The explanation for this result can be found in Table 4, which gives the significance and marginal contributions of the characteristics for the parametric portfolios in the presence of transaction costs. Of particular interest are the last two columns of the table, which give (i) the marginal contribution of each characteristic to transaction costs when combined in the optimal parametric portfolio and (ii) the marginal contribution of each characteristic to transaction costs when traded in *isolation*; that is, independently from the benchmark portfolio and the other characteristics. Comparing these two columns reveals that the reason why the number of significant characteristics is *larger* in the presence of transaction costs is that the transaction costs associated with trading combinations of characteristics are *substantially* smaller than those associated with trading characteristics in isolation. We find that the marginal transaction cost associated with trading the 15 significant characteristics is reduced by around 65% on average when they are combined. This reduction is illustrated in Figure 5, which depicts the marginal

[29]To evaluate the significance of the 51 characteristics individually, we solve the problem defined in (9) for the case where only one characteristic is available. Because we are considering a single characteristic at a time, we do not need to use the first step of the screen-and-clean test, and instead we just run the bootstrap significance test on each of the 51 single-characteristic parametric portfolios. Finally, note that here we consider 51 *individual* significance tests and thus, following the suggestion in Harvey et al. (2015), we apply Bonferroni's adjustment. To conserve space, the details of these results are omitted.

[30]For the case of quadratic transaction costs reported in Section IA.1, the number of characteristics that are jointly significant increases from five to 19.

contributions to transaction costs of the 15 significant characteristics for the case when the characteristics are traded jointly and in isolation.

A stark example of the trading diversification benefits from combining characteristics is the short-term reversal characteristic (*mom1m* in the 14th row of Table 4), which has an enormous marginal contribution to transaction costs if traded in isolation (marginal contribution 0.00857), but a dramatically smaller marginal contribution to transaction costs when traded in combination (marginal contribution 0.00211). As a result, the short-term reversal characteristic is significant even in the presence of transaction costs when traded in combination with other characteristics.[31]

In Section 4, we showed analytically and empirically that combining characteristics in an equally weighted portfolio results in a substantial reduction in turnover compared to trading them in isolation. The results in this section confirm that combining characteristics in an *optimal* parametric portfolio results in a substantial reduction in *transaction costs*. The intuition behind these results is that combining a larger number of characteristics is advantageous in the presence of transaction costs because the benefits from trading diversification grow with the number of characteristics exploited as shown in Proposition 3. The main takeaway is that transaction costs *increase* the dimension of the cross-section of stock returns and provide an economic rationale for non-sparse characteristic-based asset-pricing models.

7 Out-of-sample analysis

The previous sections studied the significance of the different characteristics for portfolio choice *in-sample*; that is, for our full sample of observations. In this section, we study whether an investor can improve *out-of-sample* performance net of transaction costs by

[31]This result contrasts sharply with DeMiguel, Nogales, and Uppal (2014) and Novy-Marx and Velikov (2016) that find that the short-term reversal characteristic is not profitable after transaction costs *when traded in isolation*. DeMiguel et al. (2014) finds that a short-term reversal (contrarian) strategy is not profitable in the presence of even modest proportional transaction costs of 10 basis points. Novy-Marx and Velikov (2016) finds that the short-term reversal strategy does not improve the investment opportunity set of an investor with access to the Fama-French factors, even when a buy-and-hold transaction-cost-mitigation strategy is employed.

exploiting a larger set of characteristics than that considered in prominent asset-pricing models.

7.1 Methodology for out-of-sample evaluation

To evaluate the out-of-sample performance of the different portfolios we use a "rolling-horizon" procedure similar to that used in DeMiguel, Garlappi, and Uppal (2009b). First, we choose a window over which to perform the estimation. The total number of monthly observations in the dataset is $T_{tot} = 419$ and we choose an estimation window of $T = 100$. Second, using the return data over the estimation window, we compute the various portfolios we study. Third, we repeat this "rolling-window" procedure for the next month, by including the data for the next month and dropping the data for the earliest month. We continue doing this until the end of the dataset is reached. At the end of this process, we have generated $T_{tot} - T = 319$ portfolio-weight vectors, w_t^j, for $t = T, \ldots, T_{tot} - 1$ and for each strategy j. Holding the portfolio w_t^j for one month gives the *out-of-sample* return net of transaction costs at time $t+1$:

$$r_{t+1}^j = (w_t^j)^\top r_{t+1} - \|\Lambda_t(w_t^j - (w_{t-1}^j)^+)\|_1,$$

where Λ_t is the transaction cost matrix at time t defined in Section 3.2, and $(w_{t-1}^j)^+$ is the portfolio for the jth strategy before rebalancing at time t; that is

$$(w_{t-1}^j)^+ = w_{t-1}^j \circ (e_{t-1} + r_t),$$

where e_{t-1} is the N_{t-1} dimensional vector of ones and $x \circ y$ is the Hadamard or componentwise product of vectors x and y. Then, for each portfolio we study, we compute the monthly turnover, and the out-of-sample annualized mean, standard deviation, and Sharpe ratio of returns net of transaction costs:

$$\text{turnover}^j = \frac{1}{T_{tot} - T} \sum_{t=T}^{T_{tot}-1} \|w_t^j - (w_{t-1}^j)^+\|_1,$$

$$\hat{\mu}^j = \frac{12}{T_{tot} - T} \sum_{t=T}^{T_{tot}-1} (w_t^j)^\top r_{t+1},$$

$$\hat{\sigma}^j = \left(\frac{12}{T_{tot} - T} \sum_{t=T}^{T_{tot}-1} \left((w_t^j)^\top r_{t+1} - \hat{\mu}^j \right)^2 \right)^{1/2}, \quad \text{and}$$

$$\widehat{\mathrm{SR}}^j = \frac{\hat{\mu}_j}{\hat{\sigma}_j}.$$

To test if the out-of-sample performance of the regularized parametric portfolio is statistically significantly better than that of the other portfolios we consider, we use the iid bootstrap method in Ledoit and Wolf (2008), with 10,000 bootstrap samples to construct a one-sided confidence interval for the difference between Sharpe ratios. We use three/two/one asterisks (*) to indicate that the difference is significant at the 0.01/0.05/0.10 level.[32]

7.2 Out-of-sample performance

Table 5 reports the out-of-sample performance of the different portfolios in the presence of transaction costs and risk-aversion parameter $\gamma = 5$. Panel A reports the performance for the portfolios that do not use any characteristics, which are the benchmark value-weighted portfolio (VW) and the equally weighted portfolio ($1/N$). Panel B reports the performance of three parametric portfolios: two portfolios that exploit a small number of characteristics and the regularized portfolio that exploits a large set of 51 characteristics.[33] The first parametric portfolio exploits the *three* characteristics considered in Brandt et al. (2009): size, book to market, and momentum. The second parametric portfolio exploits *four* characteristics: size, book to market, asset growth, and gross

[32]Note that to reduce computation time, we compute the optimal parameter vector θ only in January of each year, and use this parameter vector to compute the parametric portfolios for every month of the year. We use the cross-validation methodology explained in Section 3.5 to calibrate the lasso threshold, but using only the 100 observations in each estimation window so that there is no look-ahead bias. Also, we find that the regularized parametric portfolios that solve problem (15)–(16) result in very large turnovers. Although we find that these portfolios are profitable even after transaction costs (see Section IA.8.3 of the Internet Appendix), they may not be implementable for institutional investors facing turnover constraints. Therefore, we report the results for the parametric portfolios after scaling them to control for turnover. Specifically, we scale the optimal parameter vector θ so that the portfolio monthly turnover is around 100%. Section IA.8.3 of the Internet Appendix reports the results in the absence of turnover controls.

[33]For the regularized parametric portfolio, we calibrate the lasso threshold by using five-fold cross-validation to select each year the lasso threshold that maximizes the mean-variance utility criterion.

profitability, which include the investment and profitability characteristics such as those highlighted in Fama and French (2015) and Hou et al. (2014).

We observe from Table 5 that the gains from exploiting a large set of characteristics are significant: the regularized parametric portfolios achieve an out-of-sample Sharpe ratio that is 100% higher than that of the parametric portfolios based on three characteristics and 25% higher than that of the parametric portfolios based on four characteristics, with the differences being statistically significant. The magnitude of the economic gains is evident also from Figure 6, which depicts the out-of-sample cumulative returns of the value-weighted portfolio and the three parametric portfolios we consider, after scaling them so that they all have the same volatility.

These out-of-sample results confirm that in the presence of transaction costs the cross section of stock returns is not fully explained by a small number of characteristics.

7.3 Can factor models explain regularized portfolio returns?

The previous section demonstrates that the regularized parametric portfolios that exploit a large set of 51 significantly outperform the two parametric portfolios that exploit only small sets of characteristics. To check the robustness of this result, we run a time-series regression of the out-of-sample returns of the regularized parametric portfolio onto three sparse factor models from the literature: the Fama and French (1993) and Carhart (1997) four-factor model (FFC), the Fama and French (2015) five-factor model (FF5), and the Hou et al. (2014) four-factor model (HXZ). All factors are obtained from Kenneth French's and Lu Zhang's websites.

Table 6 shows that none of these three sparse factor models fully explains the returns of the regularized parametric portfolios, which achieve an economically and statistically significant abnormal average monthly return of about $\alpha = 1\%$ for each of the three models. This confirms that sparse factor models cannot fully explain the out-of-sample performance of the regularized parametric portfolios.

8 Conclusion

A multitude of variables have been proposed to predict the cross-section of expected stock returns. The existing literature takes a return-prediction perspective to understand which variables provide independent information about average returns. In contrast, we take a *portfolio perspective* to understand the effects of also risk and transaction costs on the dimension of the cross section.

In response to the question posed by Cochrane, which we highlighted at the start of the manuscript, we find that in the absence of transaction costs, out of the 51 characteristics we consider, only a small number—about six—are jointly significant. In the presence of transaction costs, however, the number of significant characteristics *increases* from six to 15 because combining characteristics helps to reduce transaction costs in trading the stocks underlying the characteristics. Kozak et al. (2017) concludes that "the empirical asset-pricing literature's multi-decade quest for a sparse characteristics-based factor model [...] is ultimately futile." We find that transaction costs increase the number of characteristics that are significant for portfolio construction. Moreover, we find that in the presence of transaction costs, the regularized parametric portfolios that exploit a large set of characteristics outperform out of sample the parametric portfolios that exploit only small sets of characteristics, such as those considered in prominent asset-pricing models. Thus, our results provide an economic rationale based on transaction costs for non-sparse characteristic-based asset-pricing models.

A Relation to regression approaches

In this appendix, we study the relation of our approach based on parametric portfolios with the regression approaches frequently used in the literature. Section A.1 studies the relation to the Fama-MacBeth cross-sectional regressions, Section A.2 to the time-series regressions, and Section A.3 to the generalized alpha approach developed in Novy-Marx and Velikov (2016). Proofs for the propositions and corollary that appear in this section are given in Appendix B.

A.1 Relation to Fama-MacBeth regressions

In this section, we study analytically and empirically the relation between our approach and the Fama-MacBeth regressions in the absence of transaction costs. The Fama-MacBeth procedure can be described as running cross-sectional regressions of stock returns, r_t, onto firm-specific characteristics at each date t:

$$r_t \;=\; X_{t-1}\lambda_t + \epsilon_t, \tag{A1}$$

where $X_{t-1} \in \mathbb{R}^{N_{t-1}\times K}$ is the matrix of firm-specific characteristics at time $t-1$,[34] $\lambda_t \in \mathbb{R}^K$ is the vector of slopes at time t, and $\epsilon_t \in \mathbb{R}^{N_{t-1}}$ is the vector of pricing errors at time t. The Fama-MacBeth approach then tests the significance of the average of the slopes over time, $\overline{\lambda}$.

Most of the existing literature estimates the Fama-MacBeth cross-sectional regressions using ordinary least squares (OLS). Lewellen et al. (2010), however, recommends using generalized least squares (GLS) cross-sectional regressions because their goodness-of-fit metric has a clear economic interpretation. In particular, Lewellen et al. (2010) extends a result in Kandel and Stambaugh (1995) to show that the GLS R^2 measures the mean-variance efficiency of the model's factor-mimicking portfolios.[35] The following

[34]For the sake of simplicity and without loss of generality, we assume that X_{t-1} is divided by the number of firms at time $t-1$, as we do for parametric portfolios.

[35]Lewellen et al. (2010) studies two-pass cross-sectional regressions, rather than Fama-MacBeth regressions; see (Cochrane, 2009, Sections 12.2 and 12.3). For our theoretical analysis, we make the simplifying assumption that the characteristics are time invariant, and in this case the cross-sectional regressions coincide with the Fama-MacBeth regressions. In addition, we use firm-specific characteristic data, rather than factor data, and thus all of our analysis is based on a single pass regression of stock returns onto characteristics.

proposition clarifies the relation between our portfolio approach and the Fama-MacBeth OLS and GLS regressions.

Proposition A1. *Assume that the standardized firm characteristics are constant through time so that $X_t = X$. Then, the OLS and GLS Fama-MacBeth average slopes are*

$$\overline{\lambda}_{OLS} = (X^\top X)^{-1} X^\top \widehat{\mu}_r, \quad \text{and} \tag{A2}$$

$$\overline{\lambda}_{GLS} = (X^\top \widehat{\Sigma}_r^{-1} X)^{-1} X^\top \widehat{\Sigma}_r^{-1} \widehat{\mu}_r, \tag{A3}$$

where $\widehat{\mu}_r \in \mathbb{R}^N$ is the sample mean of stock returns and $\widehat{\Sigma}_r \in \mathbb{R}^{N\times N}$ is the sample covariance matrix of stock returns. Assume also that the sample vector of covariances between the benchmark portfolio return and the characteristic portfolio return vector is zero ($\sigma_{bc} = 0$). Then the optimal mean-variance parametric portfolio is

$$\theta^* = \frac{1}{\gamma} (X^\top \widehat{\Sigma}_r X)^{-1} X^\top \widehat{\mu}_r. \tag{A4}$$

Proposition A1 shows that the OLS and GLS Fama-MacBeth slopes differ in general from the mean-variance parametric portfolio weights; that is, testing the significance of Fama-MacBeth slopes is different from testing the significance of the weights a mean-variance investor assigns to each characteristic. Note, in particular, that the OLS and GLS Fama-MacBeth slopes are different in general from the mean-variance parametric portfolio weights *unless* the sample covariance matrix of asset returns is equal to the identity matrix ($\Sigma_r = I$).

The following corollary provides further insight into the difference between the parametric portfolio weights and the OLS Fama-MacBeth slopes.

Corollary A1. *Let the assumptions in Proposition A1 hold, and assume in addition that the columns of the firm-specific characteristic matrix X are orthonormal; that is, $X^\top X = I$. Then, the optimal mean-variance parametric portfolio is*

$$\theta^* = \frac{1}{\gamma} \widehat{\Sigma}_c^{-1} \overline{\lambda}_{OLS}, \tag{A5}$$

where $\widehat{\Sigma}_c$ is the sample covariance matrix of characteristic returns and γ is the risk-aversion parameter.

Corollary A1 shows that, for the particular case in which the columns of the firm-specific characteristic matrix are orthonormal, there is a componentwise one-to-one relation between mean-variance parametric portfolio weights and OLS Fama-MacBeth slopes *only if* the sample covariance matrix of characteristic returns, $\widehat{\Sigma}_c$, is diagonal.[36] If, on the other hand, characteristic returns are correlated, then a given characteristic k could have a zero OLS Fama-MacBeth slope ($\overline{\lambda}_k = 0$), and yet have a nonzero parametric portfolio weight ($\theta_k^* \neq 0$). This is the case, for instance, when the correlation of the kth characteristic return with the returns on the other characteristics can be exploited by the investor to reduce risk, and thus, improve her overall mean-variance utility.

The above theoretical results demonstrate that testing the significance of Fama-MacBeth slopes will, in general, produce results that are different from those of testing the significance of the weights that a mean-variance investor assigns to each characteristic. We now compare empirically the significance results from OLS Fama-MacBeth regressions with those of our approach.[37] Table A1 reports the significance of the Fama-MacBeth slopes for the six characteristics we found to be significant in Section 5 plus size, book to market, and momentum. The first column lists the name of the characteristics, the second column reports the multiple regression slopes and Newey-West t-statistics (in brackets),[38] and the third column reports the individual regression slopes and Newey-West t-statistics.

We see from Table A1 that the five characteristics that are significant at the 5% level in Section 5 are also jointly significant for cross-sectional regressions. However, in contrast to the finding in Section 5, beta is not significant in the Fama-MacBeth regressions even at the 10% level. This is because, as shown in Proposition A1, Fama-MacBeth slopes differ in general from parametric portfolio weights when the returns on the characteristics are correlated over time and the investor can exploit this to reduce the risk of the mean-variance portfolio. Regarding the book-to-market and momentum

[36]To see this, note that if $\widehat{\Sigma}_c$ is diagonal, then $\theta_k^* = (\overline{\lambda}_{OLS})_k/(\gamma(\widehat{\Sigma}_c)_{kk})$, where $(\widehat{\Sigma}_c)_{kk}$ is the kth element of the diagonal of $\widehat{\Sigma}_c$, and thus there is a one-to-one correspondence between the kth component of θ^* and the kth component of $\overline{\lambda}_{OLS}$.

[37]We do not run GLS Fama-MacBeth regressions because the sample covariance matrix of stock returns is singular for our case with thousands of stocks and only hundreds of monthly dates.

[38]We compute t-statistics with Newey-West adjustments of 12 lags, as in Green et al. (2017).

characteristics, we see from Table A1 that both book to market (*bm*) and 12-month momentum (*mom12m*) are significant for multiple cross-sectional regressions, whereas they were not significant from a portfolio perspective. Intuitively, these characteristics are significant in multiple cross-sectional regressions because these regressions ignore the large contribution of these characteristics to the risk of the overall portfolio of characteristics, which reduces their appeal from a portfolio perspective.

A.2 Relation to time-series regressions

In this section, we study analytically and empirically the relation of our portfolio approach to the time-series regression approach in the absence of transaction costs. The time-series approach may be described as regressing the return of a *new* characteristic long-short portfolio onto the returns of K_c *commonly* accepted characteristic long-short portfolios; that is,

$$r_{n,t} = \alpha_{TS} + \beta_{TS}^{\top} r_{c,t} + \epsilon_t, \tag{A6}$$

where $r_{n,t} \in \mathbb{R}$ is the return of the *new* characteristic long-short portfolio at time t, $r_{c,t} \in \mathbb{R}^{K_c}$ is the return of the commonly accepted characteristic long-short portfolios at time t, the error term $\epsilon_t \in \mathbb{R}$ follows a Normal distribution with zero mean and standard deviation σ_ϵ, $\alpha_{TS} \in \mathbb{R}$ is the intercept of the regression, and $\beta_{TS} \in \mathbb{R}^{K_c}$ is the slope vector. If the intercept in this regression is significant, the return on the new characteristic is not fully explained by the return of the commonly accepted characteristics. Gibbons et al. (1989) shows that a significant intercept implies that the new characteristic-based long-short portfolio improves the investment opportunity set of a mean-variance investor who already has access to the returns of the set of commonly accepted characteristics.

As explained above, the time-series regression approach tests the significance of the intercept. In contrast, the following proposition shows that, in the absence of transaction costs, our approach is equivalent to testing the significance of the *slopes* in a particular constrained time-series multiple regression. Britten-Jones (1999) shows that the tangency mean-variance portfolio can be identified by solving a linear regression. We extend this result to the context of *any* parametric portfolio on the mean-variance efficient frontier by introducing a constraint on the mean return of the portfolio.

Proposition A2. *For a given risk-aversion parameter* γ*, the optimal parameter* θ^* *for the mean-variance parametric portfolio problem without transaction costs* (5) *is equal to the ordinary least square (OLS) estimate of the slope vector in the following time-series regression model:*

$$r_{b,t} = \alpha - \beta^\top r_{c,t} + \epsilon_t, \tag{A7}$$

subject to the constraint that

$$\beta^\top \mu_c = (\theta^*)^\top \mu_c, \tag{A8}$$

where $r_{b,t} \in \mathbb{R}$ *is the return of the benchmark portfolio,* $r_{c,t} \in \mathbb{R}^K$ *is the return on the characteristics,* $\alpha \in \mathbb{R}$ *is the intercept,* $\beta \in \mathbb{R}^K$ *is the slope vector,* μ_c *is the mean characteristic return vector, and* $(\theta^*)^\top \mu_c$ *is the average return of the mean-variance parametric portfolio.*

The advantage of the parametric-portfolio approach is that by focusing on the slopes, it allows one to test the significance of the different characteristics when they are considered *jointly*. The traditional time-series approach, on the other hand, is designed to test only the significance of a single characteristic when it is added to a set of commonly accepted characteristics; see also Footnote 7. This is a limitation of the time-series regression approach because the result of the statistical inference depends on the sequence in which variables are selected. For instance, when regressing the return of each characteristic in our dataset onto the returns of the four Fama and French (1993) and Carhart (1997) factors downloaded from Kenneth French's website, we find that eight characteristics are significant in the absence of transaction costs, but beta *is not* significant.[39] Beta, however, *is* significant when its returns are regressed onto the four Fama and French (1993) and Carhart (1997) factors *plus* the return of the return-volatility long-short portfolio, because beta helps to hedge the return-volatility characteristic.[40] Accordingly, beta *matters* if one controls for return volatility.[41] Our portfolio approach considers all char-

[39] We run 48 significance tests corresponding to the 51 characteristics except size, value, and momentum and thus, following Harvey et al. (2015) we apply Bonferroni's adjustment.

[40] We again apply Bonferroni's adjustment.

[41] This result is analogous to that in Asness et al. (2018), which finds that despite the weak performance of the *size* characteristic when evaluated in isolation, it becomes significant once it is considered in combination with a quality characteristic.

acteristics simultaneously and finds that return volatility and beta are jointly significant together with four other characteristics. These empirical results highlight the importance of considering all characteristics simultaneously. Other advantages of our portfolio approach are that it allows one to consider transaction costs in a straightforward manner and identify the marginal contribution of each characteristic to the investor's utility.

A.3 Relation to generalized alpha

In this section, we compare empirically the results from our portfolio approach in the presence of transaction costs with those from using the generalized alpha developed in Novy-Marx and Velikov (2016), which extends the traditional time-series regression framework to take transaction costs into account. Novy-Marx and Velikov (2016) proposes computing the returns of the mean-variance portfolio in the presence of transaction costs for the commonly accepted characteristics, MVE_X, and the returns of the mean-variance portfolio in the presence of transaction costs for the commonly accepted characteristics plus the new characteristic, $\text{MVE}_{X,y}$. Then it runs the following regression:

$$\text{MVE}_{X,y}/w_y = \alpha + \beta \text{MVE}_X + \epsilon, \tag{A9}$$

where w_y is the weight of the mean-variance portfolio on the new characteristic. Novy-Marx and Velikov (2016) shows that in the absence of transaction costs, the generalized alpha in (A9) equals the alpha from the traditional time-series approach. In the presence of transaction costs, this approach tests the significance of adding the new characteristic to a set of commonly accepted characteristics taking transaction costs into account.[42]

As discussed in Section A.2, the main advantage of our portfolio approach with respect to the time-series approach is that it considers all characteristics simultaneously and tests their significance when considered jointly, whereas the time-series regressions are designed to consider one characteristic at a time; see Footnote 7. To illustrate this, we compute the generalized alpha for each of our characteristics with respect to the four Fama and French (1993) and Carhart (1997) factors downloaded from Kenneth French's

[42]Although the implementation in Novy-Marx and Velikov (2016) considers the transaction cost associated with each characteristic independently, here we extend the approach in Novy-Marx and Velikov (2016) to capture trading diversification.

website. We find that, in the presence of transaction costs, *none* of the characteristic portfolios has a significant generalized alpha with respect to the four factors.[43] However, in the absence of transaction costs, Section A.2 showed that eight characteristics were significant with respect to the four factors. That is, the number of characteristics that are significant with respect to the four factors for the time-series approach *decreases* in the presence of transaction costs when the characteristics are considered in isolation, which is consistent with the results in Novy-Marx and Velikov (2016). In contrast, our portfolio approach shows that the number of significant characteristics *increases* in the presence of transaction costs. This is because our approach allows one to consider all characteristics simultaneously and identify the optimal combination of characteristics that results in substantial trading diversification.

[43]To address the multiple testing problem, we again apply Bonferroni's adjustment because we carry out 48 significance tests corresponding to our 51 characteristics except size, value, and momentum.

B Proofs for all propositions

Proof of Proposition 1

Equation (3) shows that the parametric portfolio is a combination of the benchmark portfolio and the K standardized firm-specific characteristics, scaled by the number of firms N_t. Therefore, we can define this combination as $w = [1,\ \theta] \in \mathbb{R}^{K+1}$ and the vector of benchmark and characteristic returns as $R_t = [r_{b,t}\ ,\ r_{c,t+1}/N_t]$. Under this specification, the mean-variance parametric portfolio problem takes the familiar form:

$$\min_{w} \quad \frac{\gamma}{2} w^\top \widehat{\Sigma} w - w^\top \widehat{\mu}, \tag{B1}$$

$$\text{s.t.} \quad w_1 = 1, \tag{B2}$$

where $w = [w_1,\ \theta] \in \mathbb{R}^{K+1}$ and $\widehat{\Sigma}$ and $\widehat{\mu}$ are the sample covariance matrix and mean of $R_t = [r_{b,t}\ ,\ r_{c,t+1}]$. The result follows by using straightforward algebra to eliminate the decision variable w_1 and the constraint, and then removing terms in the objective function that do not depend on the parameter vector θ.

Proof of Proposition 2

The marginal contributions of the characteristics are given by the subdifferential of the objective function in (10) with respect to θ. Note that the first four terms in (10) are differentiable with respect to θ and thus their subdifferentials coincide with their gradient. It is straightforward to show that the gradients of these four terms are given by the first four terms in the right-hand side of (11).

The only term that is not differentiable is the implied transaction cost from trading asset i at time $t+1$. From expression (7), we can define the transaction cost term for asset i at time $t+1$ as

$$u_{i,t+1} = |\Lambda_{ii,t} \left(w_{i,t+1}(\theta) - w^{+}_{i,t}(\theta) \right)|, \tag{B3}$$

where $\Lambda_{ii,t}$ is the associated transaction cost parameter for asset i at time t. Therefore, it suffices to characterize the subdifferential of expression (B3). Note that the function inside the absolute value is differentiable with respect to θ. Thus, applying the chain rule for subdifferentials, we have that the subdifferential of $u_{i,t+1}$ with respect to the kth parametric portfolio weight θ_k is equal to the subdifferential of the absolute value function times the differential of $\Lambda_{ii,t} \left(w_{i,t+1}(\theta) - w^{+}_{i,t}(\theta) \right)$.

Note that $\Lambda_{ii,t} > 0$ and thus, the subdifferential of the absolute value function is given by the sign function as precisely defined in (13). Finally, the differential of the term $\Lambda_{ii,t}\left(w_{i,t+1}(\theta) - w^{+}_{i,t}(\theta)\right)$ is

$$\frac{d[\Lambda_{ii,t}\left(w_{i,t+1}(\theta) - w^{+}_{i,t}(\theta)\right)]}{d\theta_k} = \Lambda_{ii,t}[(X_{t+1})_{ik} - (X_t)_{ik}(1 + r_{i,t+1})].$$

The result follows by adding the subdifferentials of $u_{i,t+1}$ for $i = 1, 2, \ldots, N_t$, and then combining the subdifferentials with respect to θ_k for $k = 1, 2, \ldots, K$ into a single vector.

Proof of Proposition 3

Part 1. The trade in the ith stock required to rebalance an equally weighted portfolio of K characteristics is:

$$\text{trade}^{ew}_i = \frac{1}{K}\sum_{k=1}^{K}\text{trade}_{i,k} = \frac{1}{K}\sum_{k=1}^{K}[(X_{t+1})_{i,k} - (X_t)_{i,k}(1 + r_{i,t+1})]. \tag{B4}$$

Because $\text{trade}_{i,k}$ for $k = 1, 2, \ldots, K$ are jointly distributed as a multivariate Normal distribution with zero mean and and covariance matrix Ω, we have that trade^{ew}_i is distributed as a Normal distribution with zero mean and standard deviation $\sqrt{e^\top \Omega e}/K$.

By definition, the average trading volume (turnover) in the ith stock required to rebalance an equally weighted portfolio of the K characteristics is the average of the absolute value of trade^{ew}_i. Geary (1935) shows that the mean absolute deviation of a Normally distributed random variable is $\sqrt{2/\pi}$ times its standard deviation. Therefore, the average turnover in the ith stock required to rebalance an equally weighted portfolio of K characteristics is

$$\text{turnover}(\text{trade}^{ew}_i) = \sqrt{2/\pi} \times \sqrt{e^\top \Omega e}/K. \tag{B5}$$

Following a similar argument, the average cost of the trade in the ith stock required to rebalance a quantity $1/K$ of each of the K characteristics in isolation is

$$\text{turnover}(\text{trade}^{iso}_i) = \sqrt{2/\pi} \times \sum_{k=1}^{K}\frac{\sqrt{\Omega_{kk}}}{K}. \tag{B6}$$

Taking the ratio of (B5) to (B6), we get

$$\frac{\text{turnover}(\text{trade}_i^{ew})}{\text{turnover}(\text{trade}_i^{iso})} = \frac{\sqrt{e^\top \Omega e}}{\sum_{k=1}^K \sqrt{\Omega_{kk}}}. \tag{B7}$$

To show that this ratio is strictly smaller than one, we note that the square of the ratio in (B7) is

$$\frac{e^\top \Omega e}{(\sum_{k=1}^K \sqrt{\Omega_{kk}})^2} = \frac{\sum_{k=1}^K \Omega_{kk} + \sum_{l=1}^K \sum_{m\neq l} \rho_{lm} \sqrt{\Omega_{ll}} \sqrt{\Omega_{mm}}}{\sum_{k=1}^K \Omega_{kk} + \sum_{l=1}^K \sum_{m\neq l} \sqrt{\Omega_{ll}} \sqrt{\Omega_{mm}}}, \tag{B8}$$

where ρ_{lm} is the correlation between the rebalancing trade in the ith stock for the lth and mth characteristics. The ratio in (B8) is smaller than one because $\rho_{lm} < 1$ by the assumption that Ω is positive definite.

Part 2. Because Ω is symmetric with respect to the K characteristics, we have that trade_i^{ew} is distributed as a Normal distribution with zero mean and standard deviation $\sqrt{e^\top \Omega e}/K = \sigma(1 + \rho(K-1))/K$. The result follows using arguments identical to those in the proof of Part 1.

Part 3. Because $\rho = 0$, we have that trade_i^{ew} is distributed as a Normal distribution with zero mean and standard deviation $\sqrt{e^\top \Omega e}/K = \sigma/K$. The result follows using arguments identical to those in the proof of Part 1.

Proof of Proposition A1

Let us consider the following cross-sectional regression model:

$$r_t = X\lambda_t + \epsilon_t, \tag{B9}$$

where $r_t \in \mathbb{R}^N$ is the vector of stock returns at time t, $X \in \mathbb{R}^{N\times K}$ is the matrix of standardized firm characteristics, $\lambda_t \in \mathbb{R}^K$ is the vector of slopes at time t, and $\epsilon_t \in \mathbb{R}^N$ is the vector of pricing errors at time t.[44] The OLS and GLS Fama-MacBeth slopes of model (B9) are

$$\overline{\lambda}_{OLS} = (X^\top X)^{-1} X^\top \widehat{\mu}_r \tag{B10}$$

$$\overline{\lambda}_{GLS} = (X^\top \widehat{\Sigma}_r^{-1} X)^{-1} X^\top \widehat{\Sigma}_r^{-1} \widehat{\mu}_r, \tag{B11}$$

[44] Note that we now assume that characteristics X_t and the number of firms N_t are constant through time and therefore we drop the subscript t.

where $\widehat{\mu}_r$ is the vector of sample mean returns. It is straightforward to see that $\overline{\lambda}_{OLS}$ and $\overline{\lambda}_{GLS}$ are identical when $\widehat{\Sigma}_r$ is the identity matrix. On the other hand, we know that the solution of a mean-variance parametric portfolio is

$$\theta^* = \frac{1}{\gamma}\widehat{\Sigma}_c^{-1}\widehat{\mu}_c - \widehat{\Sigma}_c^{-1}\widehat{\sigma}_{bc}. \tag{B12}$$

Now, given the assumption that firm characteristics are constant, we can define the vector of mean characteristic-portfolio returns and the covariance matrix of characteristic-portfolio returns as $\widehat{\mu}_c = X^\top \widehat{\mu}_r$ and $\widehat{\Sigma}_c = X^\top \widehat{\Sigma}_r X$, respectively. Assuming that the covariance between characteristic portfolio returns and the benchmark portfolio is zero, expression (B12) can be then expressed as

$$\theta^* = \frac{1}{\gamma}(X^\top \widehat{\Sigma}_r X)^{-1} X^\top \widehat{\mu}_r. \tag{B13}$$

Therefore, one can see that $\overline{\lambda}_{OLS}$, $\overline{\lambda}_{GLS}$, and θ^* will be equivalent when $\widehat{\Sigma}_r$ is the identity matrix of dimension N and the covariance between characteristic portfolio returns and the benchmark portfolio is zero.

Proof of Corollary A1

The result in Corollary A1 follows from the assumption that $X^\top X = I$, which implies that $\overline{\lambda}_{OLS} = X^\top \widehat{\mu}_r = \widehat{\mu}_c$. Then, if the covariance between characteristic-portfolio returns and the benchmark portfolio is zero, we can define the solution of a mean-variance parametric portfolio as

$$\theta^* = \frac{1}{\gamma}\widehat{\Sigma}_c^{-1}\overline{\lambda}_{OLS}. \tag{B14}$$

Proof of Proposition A2

We can estimate model (A7) with OLS. The corresponding optimization problem, in matrix form, is

$$\begin{aligned} \min_{\alpha,\beta} \quad & r_b^\top r_b + \alpha^2 T + \beta^\top r_c^\top r_c \beta - 2\alpha r_b^\top e_T + 2 r_b^\top r_c \beta - 2\alpha e_T^\top r_c \beta \\ \text{s.t.} \quad & \widehat{\mu}_c^\top \beta = \mu_0, \end{aligned}$$

where e_T is a T-dimensional vector of ones. Now, given that $\widehat{\Sigma}_c = r_c^\top r_c - \widehat{\mu}_c \widehat{\mu}_c^\top$, $\widehat{\sigma}_{bc} = r_b^\top r_c - \widehat{\mu}_b \widehat{\mu}_c^\top$ and $e_T^\top r_c = T\widehat{\mu}_c$, we can write the above problem as

$$\begin{aligned} \min_{\alpha,\beta} \quad & r_b^\top r_b + \alpha^2 T + \beta^\top \widehat{\Sigma}_c \beta + \beta^\top \widehat{\mu}_c \widehat{\mu}_c^\top \beta - 2\alpha r_b^\top e_T + 2(\widehat{\sigma}_{bc} + \widehat{\mu}_b \widehat{\mu}_c)^\top \beta - 2\alpha T \widehat{\mu}_c^\top \beta \\ \text{s.t.} \quad & \widehat{\mu}_c^\top \beta = \mu_0. \end{aligned}$$

Because $\widehat{\mu}_c^\top \beta$ is constant in the feasible region, we can obtain the OLS slopes of (A7) as the solution to the following problem:

$$\begin{aligned} \min_{\beta} \quad & \beta^\top \widehat{\Sigma}_c \beta + 2\widehat{\sigma}_{bc} \beta \\ \text{s.t.} \quad & \widehat{\mu}_c^\top \beta = \mu_0, \end{aligned}$$

which is a quadratic mean-variance optimization problem. If we set μ_0 equal to the solution of the mean-variance parametric portfolio problem times the vector of mean characteristic portfolio returns (that is, $\mu_0 = \theta^{*\top} \widehat{\mu}_c$), the OLS slopes of the time-series model in (A7) coincide with the solution of the mean-variance parametric portfolio problem in (5).

Table 1: List of characteristics considered

This table lists the characteristics we consider, ordered alphabetically by acronym. The first column gives the number of the characteristic, the second column gives the characteristic's definition, the third column gives the acronym, and the fourth and fifth columns give the authors who analyzed them, and the date and journal of publication. Our definitions and acronyms match those in Green et al. (2017).

#	Characteristic and definition	Acronym	Author(s)	Date and Journal
1	Abnormal volume in earnings announcement: Average daily trading volume for 3 days around earnings announcement minus average daily volume for 1-month ending 2 weeks before earnings announcement divided by 1-month average daily volume. Earnings announcement day from Compustat quarterly	aeavol	Lerman, Livnat & Mendenhall	2007, WP
2	Asset growth: Annual percent change in total assets	agr	Cooper, Gulen & Schill	2008, JF
3	Bid-ask spread: Monthly average of daily bid-ask spread divided by average of daily spread	baspread	Amihud & Mendelson	1989, JF
4	Beta: Estimated market beta from weekly returns and equal weighted market returns for 3 years ending month $t-1$ with at least 52 weeks of returns	beta	Fama & MacBeth	1973, JPE
5	Book to market: Book value of equity divided by end of fiscal-year market capitalization	bm	Rosenberg, Reid & Lanstein	1985, JPM
6	Industry adjusted book to market: Industry adjusted book-to-market ratio	bm_ia	Asness, Porter & Stevens	2000, WP
7	Cash productivity: Fiscal year-end market capitalization plus long term debt minus total assets divided by cash and equivalents	cashpr	Chandrashekar & Rao	2009 WP
8	Industry adjusted change in asset turnover: 2-digit SIC fiscal-year mean adjusted change in sales divided by average total assets	chatoia	Soliman	2008, TAR
9	Change in shares outstanding: Annual percent change in shares outstanding	chcsho	Pontiff & Woodgate	2008, JF
10	Industry adjusted change in employees: Industry-adjusted change in number of employees	chempia	Asness, Porter & Stevens	1994, WP
11	Change in 6-month momentum: Cumulative returns from months $t-6$ to $t-1$ minus months $t-12$ to $t-7$	chmom	Gettleman & Marks	2006 WP
12	Industry adjusted change in profit margin: 2-digit SIC fiscal-year mean adjusted change in income before extraordinary items divided by sales	chpmia	Soliman	2008, TAR
13	Change in tax expense: Percent change in total taxes from quarter $t-4$ to t	chtx	Thomas & Zhang	2011 JAR
14	Convertible debt indicator: An indicator equal to 1 if company has convertible debt obligations	convind	Valta	2016 JFQA
15	Dollar trading volume in month $t-2$: Natural log of trading volume times price per share from month $t-2$	dolvol	Chordia, Subrahmanyan & Anshuman	2001, JFE
16	Dividends-to-price: Total dividends divided by market capitalization at fiscal year-end	dy	Litzenberger & Ramaswamy	1982, JF
17	3-day return around earnings announcement: Sum of daily returns in three days around earnings announcement. Earnings announcement from Compustat quarterly file	ear	Kishore, Brandt, Santa-Clara & Venkatachalam	2008, WP
18	Change in common shareholder equity: Annual percent change in book value of equity	egr	Richardson, Sloan, Soliman & Tuna	2005, JAE
19	Earnings to price: Annual income before extraordinary items divided by end of fiscal year market cap	ep	Basu	1977, JF
20	Gross profitability: Revenues minus cost of goods sold divided by lagged total assets	gma	Novy-Marx	2013 JFE
21	Industry sales concentration: Sum of squared percent of sales in industry for each company	herf	Hou & Robinson	2006, JF
22	Employee growth rate: Percent change in number of employees	hire	Bazdresch, Belo & Lin	2014 JPE
23	Idiosyncratic return volatility: Standard deviation of residuals of weekly returns on weekly equal weighted market returns for 3 years prior to month-end	idiovol	Ali, Hwang & Trombley	2003, JFE
24	Industry momentum: Equal weighted average industry 12-month returns	indmom	Moskowitz & Grinblatt	1999, JF

Table 1 continued: List of characteristics considered

#	Characteristic and definition	Acronym	Author(s)	Date and Journal
25	Leverage: Total liabilities divided by fiscal year-end market capitalization	lev	Bhandari	1988, JF
26	Change in long-term debt: Annual percent change in total liabilities	lgr	Richardson, Sloan, Soliman & Tuna	2005, JAE
27	12-month momentum: 11-month cumulative returns ending one month before month-end	mom12m	Jegadeesh	1990, JF
28	1-month momentum: 1-month cumulative return	mom1m	Jegadeesh	1990, JF
29	36-month momentum: Cumulative returns from months $t-36$ to $t-13$	mom36m	De Bondt & Thaler	1985, JF
30	6-month momentum: 5-month cumulative returns ending one month before month-end	mom6m	Jegadeesh & Titman	1990, JF
31	Market capitalization: Natural log of market capitalization at end of month $t-1$	mve	Banz	1981, JFE
32	Industry-adjusted firm size: 2-digit SIC industry-adjusted fiscal year-end market capitalization	mve_ia	Asness, Porter & Stevens	2000, WP
33	Δ% CAPEX - industry Δ% AR: 2-digit SIC fiscal-year mean adjusted percent change in capital expenditures	pchcapx_ia	Abarbanell & Bushee	1998, TAR
34	Δ% gross margin - Δ% sales: Percent change in gross margin minus percent change in sales	pchgm_pchsale	Abarbanell & Bushee	1998, TAR
35	Δ% sales - Δ% AR: Annual percent change in sales minus annual percent change in receivables	pchsale_pchrect	Abarbanell & Bushee	1998, TAR
36	Price delay: The proportion of variation in weekly returns for 36 months ending in month t explained by 4 lags of weekly market returns incremental to contemporaneous market return	pricedelay	How & Moskowitz	2005, RFS
37	Financial-statements score: Sum of 9 indicator variables to form fundamental health score	ps	Piotroski	2000, JAR
38	R&D to market cap: R&D expense divided by end-of-fiscal-year market capitalization	rd_mve	Guo, Lev & Shi	2006, JBFA
39	Return volatility: Standard deviation of daily returns from month $t-1$	retvol	Ang, Hodrick, Xing & Zhanf	2006, JF
40	Return on assets: Income before extraordinary items divided by one quarter lagged total assets	roaq	Balakrishnan, Bartov & Faurel	2010, JAE
41	Revenue surprise: Sales from quarter t minus sales from quarter $t-4$ divided by fiscal-quarter-end market capitalization	rsup	Kama	2009, JBFA
42	Sales to cash: Annual sales divided by cash and cash equivalents	salecash	Ou & Penman	1989, JAE
43	Sales to inventory: Annual sales divided by total inventory	saleinv	Ou & Penman	1989, JAE
44	Sales to receivables: Annual sales divided by accounts receivable	salerec	Ou & Penman	1989, JAE
45	Annual sales growth: Annual percent change in sales	sgr	Lakonishok, Shleifer & Vishny	1994, JF
46	Volatility of dollar trading volume: Monthly standard deviation of daily dollar trading volume	std_dolvol	Chordia, Subrahmanyan & Anshuman	2001, JFE
47	Volatility of share turnover: Monthly standard deviation of daily share turnover	std_turn	Chordia, Subrahmanyan & Anshuman	2001, JFE
48	Cashflow volatility: Standard deviation for 16 quarters of cash flows divided by sales	stdcf	Huang	2009, JEF
49	Unexpected quarterly earnings: Unexpected quarterly earnings divided by fiscal-quarter-end market cap. Unexpected earnings is I/B/E/S actual earnings minus median forecasted earnings if available, else it is the seasonally differenced quarterly earnings before extraordinary items from Compustat quarterly file	sue	Rendelman, Jones & Latane	1982, JFE
50	Share turnover: Average monthly trading volume for most recent 3 months scaled by number of shares outstanding in current month	turn	Datar, Naik & Radcliffe	1998, JFM
51	Zero trading days: Turnover weighted number of zero trading days for most recent month	zerotrade	Liu	2006, JFE

Table 2: Significance and marginal contributions without transaction costs

This table reports the significance and marginal contributions for the parametric portfolios without transaction costs, for risk-aversion parameter $\gamma = 5$. We run a screen-and-clean significance test. For the screen step, we calibrate the regularized parametric portfolios with five-fold cross-validation and find that the lasso threshold that maximizes investor's utility is $\delta = 25$. For the clean step, we run the bootstrap experiment for the parametric portfolios using those characteristics with nonzero θ's from the previous step. Characteristic p-values are computed using the percentile method discussed in Section 3.5. We assign three/two/one asterisks (*) to those characteristics whose p-values are lower than 0.01/0.05/0.1, respectively. To compute the optimal parametric portfolio and marginal contributions, we include all characteristics with nonzero θ's for the screen step plus the three characteristics considered in Brandt et al. (2009): size, book to market, and momentum. For each characteristic, the first column gives the acronym, the second the optimal value of the parameter and the significance asterisks, and the next four columns give the marginal contribution of the characteristic to: (i) the characteristic own-variance, (ii) the covariance of the characteristic with the other characteristics in the portfolio, (iii) the covariance of the characteristic with the benchmark portfolio, and (iv) the characteristic mean. Contributions that drive the characteristic to be nonzero are in blue sans serif font, and contributions that drive the characteristic toward zero are in red *italic* font (cf. Footnote 26).

		Marginal contributions to			
Characteristic	Param.	variance	cov (char.)	cov (bench.)	mean
sue	20.12***	*0.00341*	−0.00068	−0.00019	−0.00254
retvol	−10.85***	*−0.03529*	0.02914	0.00292	0.00323
agr	−10.37**	*−0.00397*	0.00050	0.00057	0.00290
mom1m	−3.10**	*−0.00509*	0.00454	*−0.00109*	0.00164
gma	5.97**	*0.00252*	−0.00255	*0.00069*	−0.00066
beta	2.36*	*0.00971*	−0.01381	*0.00419*	−0.00008
bm_ia	6.49	*0.00337*	−0.00328	*0.00072*	−0.00081
chcsho	−5.89	*−0.00210*	*−0.00111*	0.00092	0.00228
rd_mve	6.01	*0.00215*	−0.00096	*0.00045*	−0.00164
std_turn	8.53	*0.01442*	−0.01576	*0.00214*	−0.00080
bm	3.10	*0.00264*	*0.00023*	−0.00082	−0.00205
mve	−4.02	*−0.00136*	0.00148	*−0.00034*	0.00022
mom12m	−4.42	*−0.00784*	0.01125	*−0.00066*	*−0.00275*

Table 3: Correlations of significant characteristics

This table reports the correlation matrix for the returns of the six characteristics that are most significant in the absence of transaction costs and the returns of the three characteristics considered in Brandt et al. (2009): size (*mve*), book to market (*bm*), and momentum (*mom*12*m*).

Characteristics	sue	retvol	agr	mom1m	gma	beta	bm	mve	mom12m
Unexpected quarterly earnings (sue)	1.00	−0.43	−0.08	0.18	−0.18	−0.36	−0.05	0.41	0.45
Return volatility (retvol)	−0.43	1.00	0.22	−0.18	0.45	0.93	−0.46	−0.63	−0.17
Asset growth (agr)	−0.08	0.22	1.00	−0.33	0.56	0.33	−0.64	0.03	−0.17
1-month momentum (mom1m)	0.18	−0.18	−0.33	1.00	−0.23	−0.26	0.14	0.19	0.28
Gross profitability (gma)	−0.18	0.45	0.56	−0.23	1.00	0.54	−0.62	−0.24	−0.06
Beta (beta)	−0.36	0.93	0.33	−0.26	0.54	1.00	−0.54	−0.52	−0.21
Book to market (bm)	−0.05	−0.46	−0.64	0.14	−0.62	−0.54	1.00	−0.05	−0.08
Market capitalization (mve)	0.41	−0.63	0.03	0.19	−0.24	−0.52	−0.05	1.00	0.20
12-month momentum (mom12m)	0.45	−0.17	−0.17	0.28	−0.06	−0.21	−0.08	0.20	1.00

Table 4: Significance and marginal contributions with transaction costs

This table reports the significance and marginal contributions for the parametric portfolios in the presence of transaction costs, for risk-aversion parameter $\gamma = 5$. We run a screen-and-clean significance test. For the screen step, we calibrate the regularized parametric portfolios with five-fold cross-validation and find that the lasso threshold that maximizes investor's utility is $\delta = 25$. For the clean step, we run the bootstrap experiment for the parametric portfolios using those characteristics with nonzero θ's from the previous step. Characteristic p-values are computed using the percentile method discussed in Section 3.5. We assign three/two/one asterisks (*) to those characteristics whose p-values are lower than 0.01/0.05/0.1, respectively. To compute the optimal parametric portfolio and marginal contributions, we include all characteristics with nonzero θ's for the screen step plus the three characteristics considered in Brandt et al. (2009): size, book to market, and momentum. For each characteristic, the first column gives the acronym, the second the optimal value of the parameter and the significance asterisks, and the next five columns give the marginal contribution of the characteristic to: (i) the characteristic own-variance, (ii) the covariance of the characteristic with the other characteristics in the portfolio, (iii) the covariance of the characteristic with the benchmark portfolio, (iv) the characteristic mean, and (v) the transaction cost. The last column reports the marginal contribution of the characteristic to transaction costs when it is traded in isolation. Contributions that drive the characteristic to be nonzero are in blue sans serif font, and contributions that drive the characteristic toward zero are in red *italic* font (cf. Footnote 26).

		Marginal contributions to					Isolation
Characteristic	Param.	variance	cov (char.)	cov (bench.)	mean	tran. cost	tran. costs
rd_mve	11.85***	*0.00425*	−0.00333	*0.00045*	−0.00164	*0.00027*	*0.00055*
agr	−7.27***	*−0.00278*	*−0.00012*	0.00057	0.00290	*−0.00057*	*0.00125*
sue	3.00***	*0.00051*	*0.00077*	−0.00019	−0.00254	*0.00146*	*0.00240*
turn	−3.41***	*−0.00806*	0.00502	0.00279	0.00068	*−0.00043*	*0.00177*
retvol	−1.92***	*−0.00623*	0.00148	0.00292	0.00323	*−0.00139*	*0.00468*
std_turn	1.28***	*0.00217*	−0.00433	*0.00214*	−0.00080	*0.00082*	*0.00493*
zerotrade	−1.53***	*−0.00129*	0.00284	*−0.00205*	0.00124	*−0.00075*	*0.00235*
chatoia	4.51**	*0.00029*	*0.00008*	−0.00005	−0.00077	*0.00046*	*0.00116*
chtx	1.36**	*0.00026*	−0.00022	*0.00015*	−0.00123	*0.00104*	*0.00232*
beta	3.39**	*0.01398*	−0.01829	*0.00419*	−0.00008	*0.00021*	*0.00126*
ps	4.94**	*0.00156*	−0.00027	−0.00068	−0.00127	*0.00066*	*0.00140*
gma	6.60**	*0.00278*	−0.00298	*0.00069*	−0.00066	*0.00016*	*0.00090*
herf	−5.78**	*−0.00144*	0.00061	0.00041	0.00061	*−0.00019*	*0.00077*
mom1m	−0.62**	*−0.00102*	0.00258	*−0.00109*	0.00164	*−0.00211*	*0.00857*
bm_ia	2.85**	*0.00148*	−0.00168	*0.00072*	−0.00081	*0.00029*	*0.00128*
stdcf	−5.05*	*−0.00259*	0.00101	0.00068	0.00114	*−0.00024*	*0.00067*
pchgm_pchsale	3.46*	*0.00034*	*0.00006*	−0.00003	−0.00079	*0.00042*	*0.00122*
chcsho	−3.11*	*−0.00111*	*−0.00166*	0.00092	0.00228	*−0.00044*	*0.00123*
bm	1.74*	*0.00148*	*0.00122*	−0.00082	−0.00205	*0.00017*	*0.00121*
chmom	−0.67	*−0.00065*	0.00166	*−0.00073*	0.00044	*−0.00072*	*0.00404*
baspread	0.55	*0.00240*	−0.00795	*0.00329*	*0.00279*	−0.00053	*0.00322*
ep	1.27	*0.00206*	*0.00045*	−0.00166	−0.00104	*0.00018*	*0.00125*
idiovol	−1.80	*−0.00680*	0.00194	0.00308	0.00187	*−0.00008*	*0.00109*
roaq	−0.12	*−0.00014*	0.00292	*−0.00114*	*−0.00215*	0.00051	*0.00186*
mve	−2.28	*−0.00077*	0.00092	*−0.00034*	0.00022	*−0.00003*	*0.00045*
mom12m	−0.61	*−0.00109*	0.00418	*−0.00066*	*−0.00275*	0.00031	*0.00265*

Table 5: Out-of-sample performance

This table reports the out-of-sample performance of the different portfolios in the presence of transaction costs, for risk-aversion parameter $\gamma = 5$. Panel A reports the performance for the portfolios that do not use any characteristics, which are the benchmark value-weighted portfolio (VW) and the equally weighted portfolio ($1/N$). Panel B reports the performance of two parametric portfolios that exploit a small number of characteristics, and the regularized parametric portfolio that exploits a large set of 51 characteristics. The first parametric portfolio exploits the size, book-to-market, and momentum characteristics (Size/val./mom.). The second parametric portfolio exploits the size, book-to-market, asset growth, and gross profitability characteristics (Size/val./inv./prof.). The third portfolio is the regularized parametric portfolio that exploits all 51 characteristics (Regularized). The lasso threshold is calibrated using cross-validation over the estimation window. For each portfolio, the first column reports the monthly turnover, and the next three columns report the out-of-sample annualized mean, standard deviation, and Sharpe ratio of returns, net of transaction costs. We test the significance of the difference of the Sharpe ratio of each portfolio with that of the regularized parametric portfolio. Three/two/one asterisks (*) indicate that the difference is significant at the 0.01/0.05/0.1 level.

Policy	Turnover	Mean	SD	SR
Panel A: Portfolios with no characteristics				
VW	0.050	0.085	0.150	0.567***
1/N	0.134	0.085	0.177	0.482***
Panel B: Portfolios with characteristics				
Size/val./mom.	0.754	0.145	0.215	0.675***
Size/val./inv./prof.	0.963	0.236	0.220	1.072**
Regularized	0.979	0.241	0.178	1.356

Table 6: Factor loadings of regularized parametric portfolios

This table reports the intercept, slopes, and t-statistics (in brackets) from regressing the out-of-sample regularized portfolio returns onto three sparse factor models: (1) the Fama and French (1993) and Carhart (1997) four-factor model (FFC) that includes the market, size (SMB), value (HML), and momentum (UMD) factors; (2) the Fama and French (2015) five-factor model (FF5) that includes the market, size, value, profitability (RMW), and investment (CMA) factors; and, (3) the Hou et al. (2014) four-factor model (HXZ) that includes the market, size, investment (I/A), and profitability (ROE) factors. We report t-statistics with Newey-West adjustments of 12 lags. Factors are obtained from Kenneth French's and Lu Zhang's websites.

FFC	Coefficient	FF5	Coefficient	HXZ	Coefficient
α	0.0115	α	0.0102	α	0.0095
	[4.12]		[3.59]		[2.89]
Market	0.8898	Market	0.9747	Market	0.9147
	[15.29]		[15.35]		[11.90]
SMB	0.0745	SMB	0.1212	SMB	0.2547
	[0.49]		[0.84]		[1.37]
HML	0.3697	HML	−0.2640	I/A	0.7491
	[1.84]		[−1.71]		[2.65]
UMD	0.3249	RMW	0.2554	ROE	0.3316
	[2.46]		[1.31]		[1.69]
		CMA	1.0852		
			[3.64]		

Table A1: Fama-MacBeth regressions for significant characteristics

This table reports the slope coefficients from Fama-MacBeth regressions and the corresponding t-statistics (in brackets) with Newey-West adjustments of 12 lags. We report the results for multiple and individual regressions for the six most significant characteristics in the absence of transaction costs, and the three characteristics considered in Brandt et al. (2009): size (mve), book to market (bm), and momentum ($mom12m$).

Characteristic	Multiple	Individual
Unexpected quarterly earnings (sue)	0.0019	0.0027
	[7.38]	[7.10]
Return volatility (retvol)	−0.0037	−0.0032
	[−4.42]	[−2.22]
Asset growth (agr)	−0.0026	−0.0031
	[−5.39]	[−5.09]
1-month momentum (mom1m)	−0.0033	−0.0017
	[−4.67]	[−2.13]
Gross profitability (gma)	0.0020	0.0007
	[3.80]	[1.34]
Beta (beta)	0.0013	0.0001
	[0.99]	[0.04]
Book to market (bm)	0.0016	0.0021
	[2.11]	[2.17]
Market capitalization (mve)	−0.0007	−0.0002
	[−1.76]	[−0.40]
12-month momentum (mom12m)	0.0026	0.0030
	[2.43]	[2.45]

Figure 1: Marginal contribution to turnover of characteristics traded in isolation and in equally weighted portfolio

This figure compares the average trading volume (turnover) required to exploit the 51 characteristics in isolation with that required to exploit them in an equally weighted portfolio. The horizontal axis gives the turnover in percentage and the vertical axis gives the acronyms of the characteristics and the equally weighted portfolio (EW). The blue bars give the turnover required to exploit each of the characteristics in isolation (Isolation), the yellow bars give the marginal contribution to turnover of each characteristic in an equally weighted portfolio (Combined), and the red bar gives the turnover of the equally weighted portfolio of the 51 characteristics (EW portfolio).

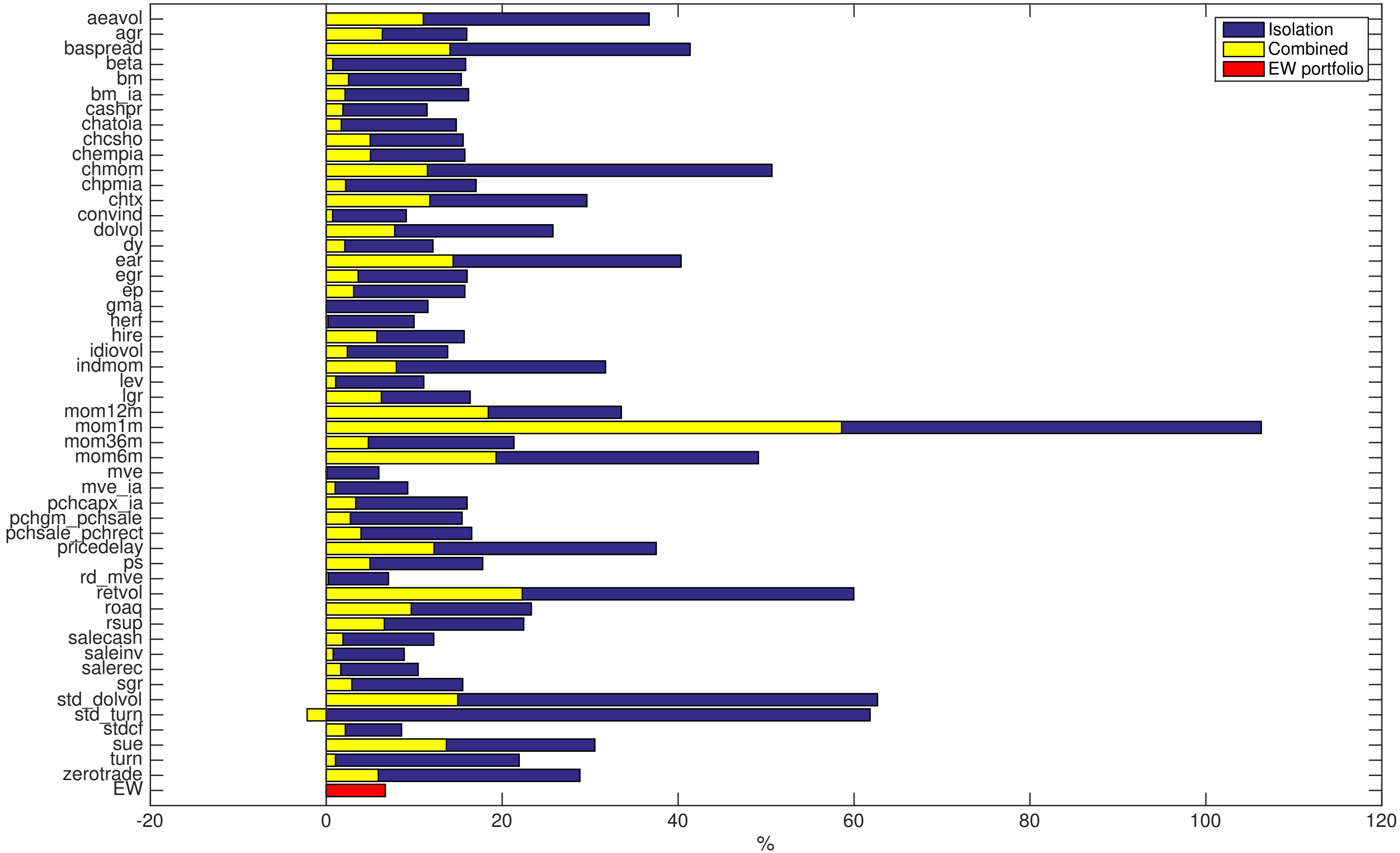


Figure 2: Correlations between rebalancing trades of different characteristics

This figure depicts a heatmap of the correlations between the rebalancing trades for the 51 characteristics for a particular stock.

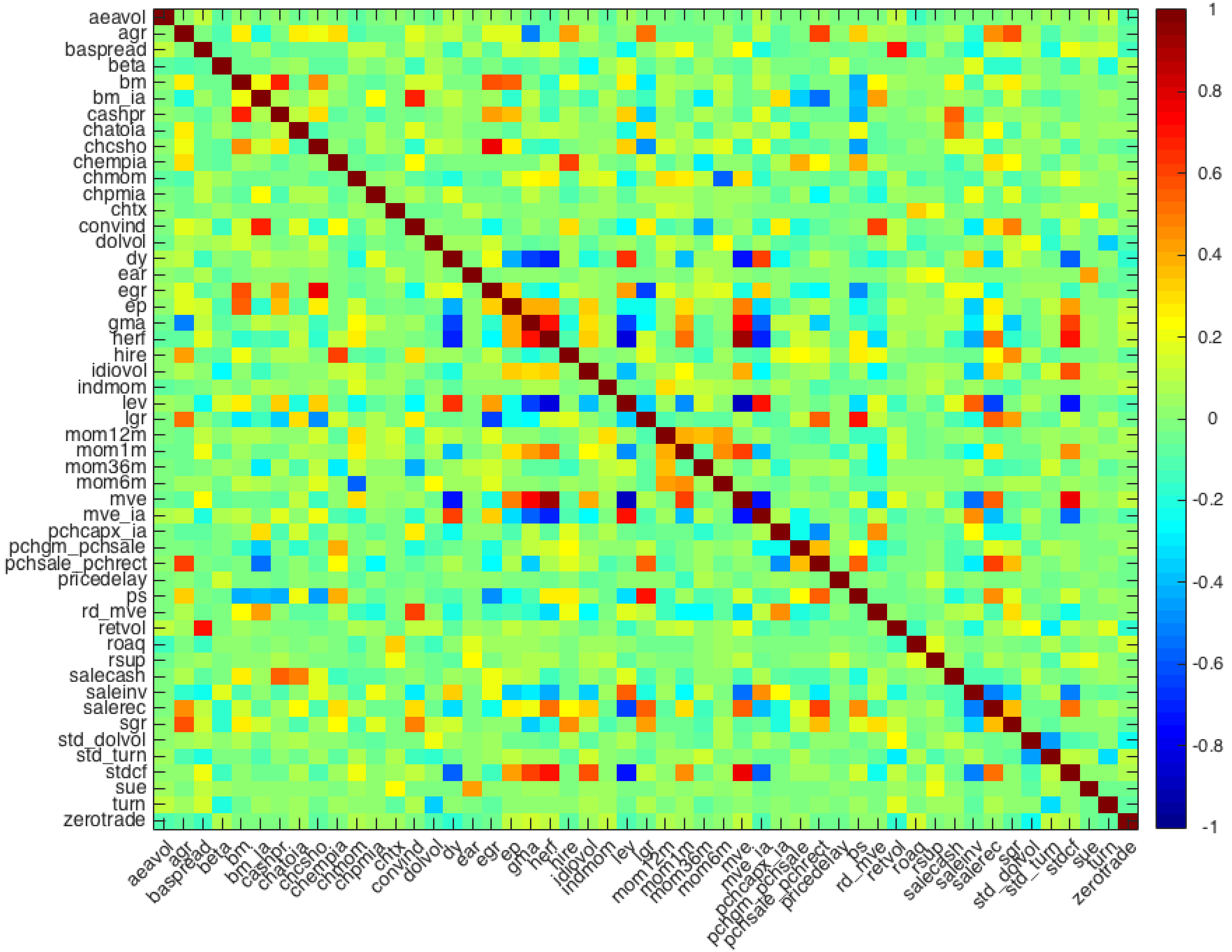


Figure 3: Marginal contributions of significant characteristics without transaction costs

This figure shows the marginal contributions to the investor's utility of the six significant characteristics in the absence of transaction costs. The vertical axis gives the labels of the significant characteristics: unexpected quarterly earnings (*unexp. earn.*), return volatility (*ret. vol.*), asset growth, 1-month momentum (*reversals*), gross profitability (*profit.*), and beta. The horizontal axis gives the marginal contributions of each characteristic to (i) the characteristic own-variance (brown bars, *variance*), (ii) the covariance of the characteristic with the other characteristics in the portfolio (yellow bars, *cov(char.)*), (iii) the covariance of the characteristic with the benchmark portfolio (light-blue bars, *cov(bench.)*), and (iv) the characteristic mean (dark-blue bars, *mean*). Contributions that drive the characteristic to be nonzero are depicted with positive bars, and contributions that drive the characteristic toward zero are depicted with negative bars; cf. Footnote 26.

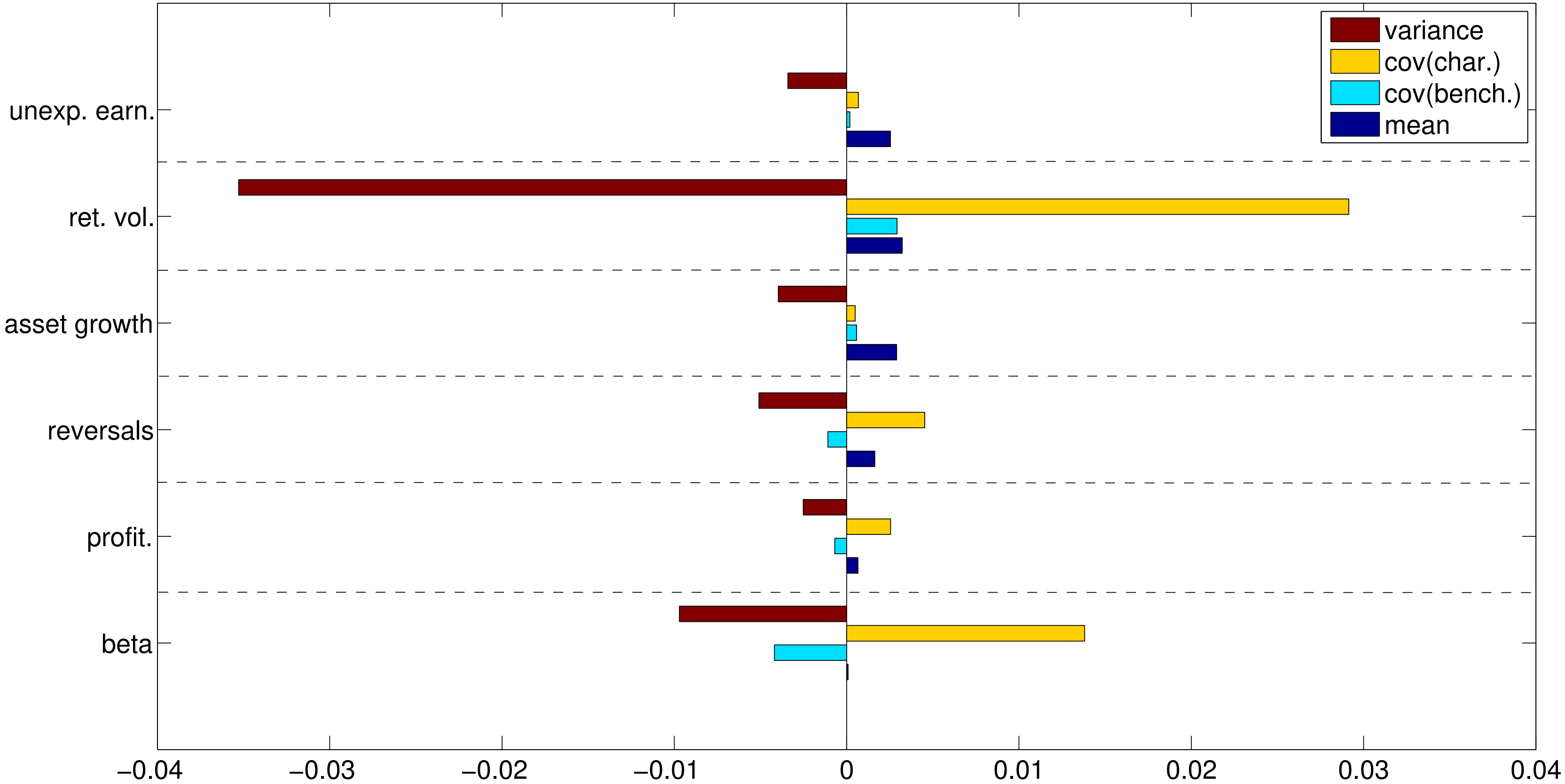
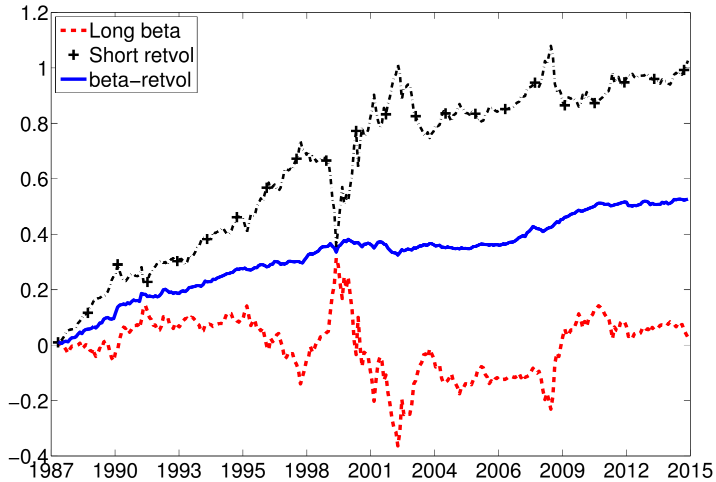


Figure 4: Cumulative returns for beta and return-volatility blended strategy

This figure shows the cumulative returns of a blended strategy that goes long beta and short return volatility. Panel (a) depicts the cumulative returns of going long *beta* (long beta), of going short return volatility (short retvol), and of a blended strategy formed by assigning 50% weight to *beta* and −50% to *retvol*. Panel (b) compares the cumulative returns of the blended strategy with *beta* and *retvol* with those of a blended strategy that assigns 50% to book to market (*bm*) and 50% to 12-month momentum (*mom*12*m*). For comparison purposes, in Panel (b) we normalize both strategies so that they have the same volatility.

(a) Beta and retvol blended strategy



(b) Two blended strategies

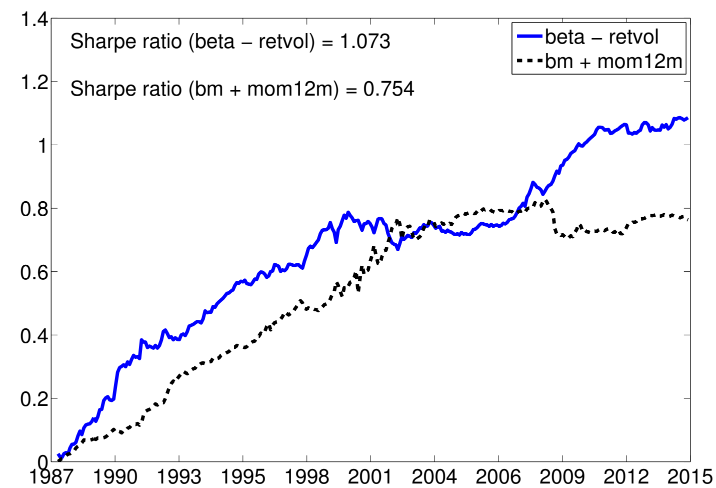


Figure 5: Marginal contribution to transaction costs of characteristics in isolation and in optimal parametric portfolio

This figure shows the marginal contribution to transaction costs when characteristics are traded in isolation and in an optimal parametric portfolio. We plot the marginal contribution to transaction costs of the 15 most significant characteristics in Table 4. The horizontal axis gives the marginal contribution to transaction costs and the vertical axis gives the acronyms of the characteristics. The blue bars give the marginal contribution of each characteristic to transaction costs when traded in isolation (Isolation) and the yellow bars give the marginal contribution of each characteristic to transaction costs when combined in the optimal parametric portfolio (Combined).

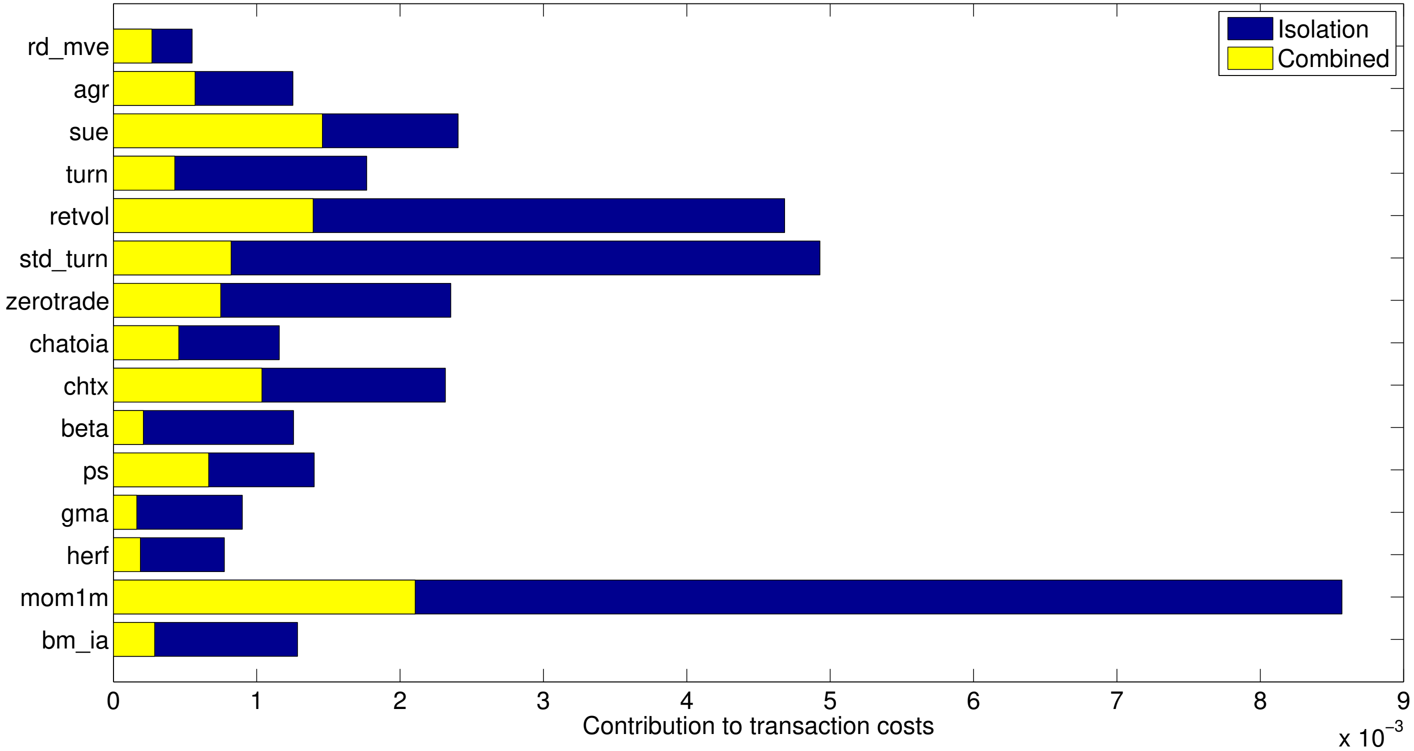
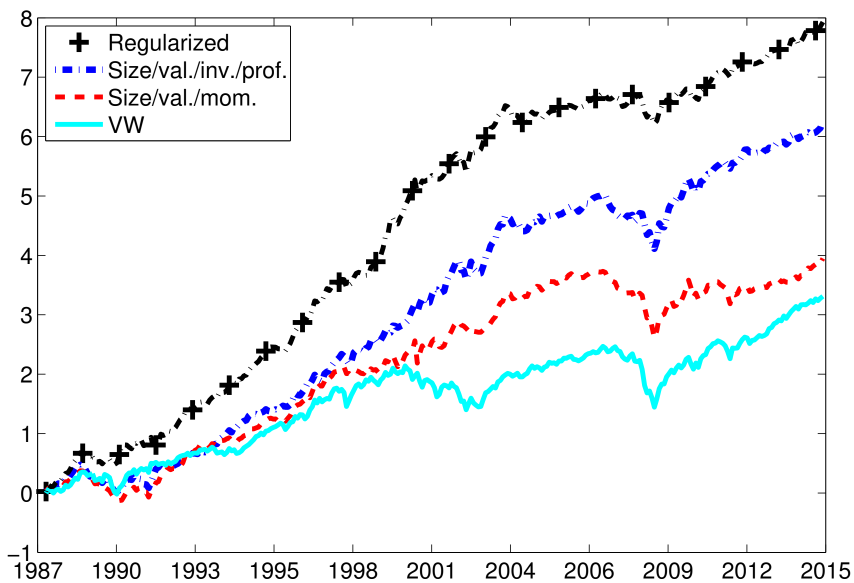


Figure 6: Out-of-sample cumulative returns

This figure shows the out-of-sample cumulative returns of the value-weighted portfolio (VW) and three different parametric portfolios in the presence of transaction costs, for risk-aversion parameter $\gamma = 5$. Two of the parametric portfolios exploit a small number of characteristics. The first parametric portfolio exploits the size, book-to-market, and momentum characteristics (Size/val./mom.). The second parametric portfolio exploits the size, book-to-market, asset growth, and gross profitability characteristics (Size/val./inv./prof.). The third parametric portfolio is the regularized parametric portfolio that exploits all 51 characteristics (Regularized). The lasso threshold is calibrated using cross-validation over the estimation window. For comparison purposes we normalize all portfolio returns so that they have the same volatility.



References

Ang, Andrew, Robert J. Hodrick, Yuhang Xing, and Xiaoyan Zhang, 2006, The cross-section of volatility and expected returns, *The Journal of Finance* 61, 259–299.

Asness, Clifford S., Andrea Frazzini, Ronen Israel, Tobias J. Moskowitz, and Lasse H. Pedersen, 2018, Size matters, if you control your junk, Forthcoming in *Journal of Financial Economics*.

Asness, Clifford S, Tobias J. Moskowitz, and Lasse H. Pedersen, 2013, Value and momentum everywhere, *The Journal of Finance* 68, 929–985.

Back, Kerry, Nishad Kapadia, and Barbara Ostdiek, 2015, Testing factor models on characteristic and covariance pure plays, SSRN working paper.

Baker, Malcolm, Patrick Luo, and Ryan Taliaferro, 2017, Detecting anomalies: The relevance and power of standard asset pricing tests, HBS working paper.

Bali, Turan G., Robert F. Engle, and Scott Murray, 2016, *Empirical Asset Pricing: The Cross Section of Stock Returns* (John Wiley & Sons, Hoboken, New Jersey).

Barroso, Pedro, and Pedro Santa-Clara, 2015, Beyond the carry trade: Optimal currency portfolios, *Journal of Financial and Quantitative Analysis* 50, 1037–1056.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen, 2014, Inference on treatment effects after selection among high-dimensional controls, *The Review of Economic Studies* 81, 608–650.

Black, Fischer, 1993, Beta and return, *The Journal of Portfolio Management* 20, 8–18.

Brandt, Michael W., Pedro Santa-Clara, and Rossen Valkanov, 2009, Parametric portfolio policies: Exploiting characteristics in the cross-section of equity returns, *Review of Financial Studies* 22, 3411–3447.

Britten-Jones, Mark, 1999, The sampling error in estimates of mean-variance efficient portfolio weights, *The Journal of Finance* 54, 655–671.

Bryzgalova, Svetlana, 2015, Spurious factors in linear asset pricing models, Stanford GSB working paper.

Carhart, Mark M., 1997, On persistence in mutual fund performance, *The Journal of Finance* 52, 57–82.

Chatterjee, Arindam, and Soumendra Nath Lahiri, 2011, Bootstrapping lasso estimators, *Journal of the American Statistical Association* 106, 608–625.

Chen, Andrew, and Mihail Velikov, 2017, Accounting for the anomaly zoo: A trading cost perspective, Federal Reserve Board working paper.

Chordia, Tarun, Amit Goyal, and Alessio Saretto, 2017, p-hacking: Evidence from two million trading strategies, HEC Lausanne working paper.

Cochrane, John H., 2009, *Asset Pricing: Revised Edition* (Princeton University Press).

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

Daniel, Kent, and Sheridan Titman, 1997, Evidence on the characteristics of cross sectional variation in stock returns, *The Journal of Finance* 52, 1–33.

DeMiguel, Victor, Lorenzo Garlappi, Francisco J. Nogales, and Raman Uppal, 2009a, A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms, *Management Science* 55, 798–812.

DeMiguel, Victor, Lorenzo Garlappi, and Raman Uppal, 2009b, Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy?, *Review of Financial Studies* 22, 1915–1953.

DeMiguel, Victor, Francisco J. Nogales, and Raman Uppal, 2014, Stock return serial dependence and out-of-sample portfolio performance, *Review of Financial Studies* 27, 1031–1073.

Efron, Brad, and Rob Tibshirani, 1993, *An Introduction to the Bootstrap* (Chapman & Hall).

Fama, Eugene F., 1991, Efficient capital markets: II, *Journal of Finance* 47, 427–465.

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

Fama, Eugene F., and Kenneth R. French, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1–22.

Fama, Eugene F., and James D. MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *The Journal of Political Economy* 81, 607–636.

Feng, Guanhao, Stefano Giglio, and Dacheng Xiu, 2017, Taming the factor zoo, Chicago Booth working paper.

Frazzini, Andrea, Ronen Israel, and Tobias J. Moskowitz, 2015, Trading costs of asset pricing anomalies, Fama-Miller working paper.

Frazzini, Andrea, and Lasse Heje Pedersen, 2014, Betting against beta, *Journal of Financial Economics* 111, 1–25.

French, Kenneth R, 2008, Presidential address: The cost of active investing, *The Journal of Finance* 63, 1537–1573.

Freyberger, Joachim, Andreas Neuhierl, and Michael Weber, 2016, Dissecting characateristics nonparametrically, Chicago Booth working paper.

Geary, Roy C., 1935, The ratio of the mean deviation to the standard deviation as a test of normality, *Biometrika* 27, 310–332.

Ghosh, Anisha, Christian Julliard, and Alex P. Taylor, 2016a, An information-theoretic asset pricing model, London School of Economics working paper.

Ghosh, Anisha, Christian Julliard, and Alex P. Taylor, 2016b, What is the consumption-capm missing? an information-theoretic framework for the analysis of asset pricing models, *The Review of Financial Studies* 30, 442—504.

Gibbons, Michael R., Stephen A. Ross, and Jay Shanken, 1989, A test of the efficiency of a given portfolio, *Econometrica* 57, 1121–1152.

Goyal, Amit, 2012, Empirical cross-sectional asset pricing: A survey, *Financial Markets and Portfolio Management* 26, 3–38.

Green, Jeremiah, John R.M. Hand, and X. Frank Zhang, 2017, The characteristics that provide independent information about average us monthly stock returns, Forthcoming in *The Review of Financial Studies*.

Hand, John R.M., and Jeremiah Green, 2011, The importance of accounting information in portfolio optimization, *Journal of Accounting, Auditing & Finance* 26, 1–34.

Hanna, J. Douglas, and Mark J. Ready, 2005, Profitable predictability in the cross section of stock returns, *Journal of Financial Economics* 78, 463–505.

Harvey, Campbell R., and Yan Liu, 2018, Lucky factors, SSRN working paper.

Harvey, Campbell R., Yan Liu, and Heqing Zhu, 2015, ...and the cross-section of expected returns, *The Review of Financial Studies* 29, 5–68.

Hastie, Trevor, Robert Tibshirani, and Martin Wainwright, 2015, *Statistical Learning with Sparsity: The Lasso and Generalizations* (CRC Press).

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

Hou, Kewei, Chen Xue, and Lu Zhang, 2014, Digesting anomalies: An investment approach, *Review of Financial Studies* 28, 650–705.

Hou, Kewei, Chen Xue, and Lu Zhang, 2017, Replicating anomalies, Technical report, National Bureau of Economic Research.

Kandel, Shmuel, and Robert F. Stambaugh, 1995, Portfolio inefficiency and the cross-section of expected returns, *The Journal of Finance* 50, 157–184.

Kogan, Leonid, and Mary H. Tian, 2013, Firm characteristics and empirical factor models: A data-mining experiment, FRB International Finance discussion paper.

Korajczyk, Robert A., and Ronnie Sadka, 2004, Are momentum profits robust to trading costs?, *The Journal of Finance* 59, 1039–1082.

Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh, 2017, Shrinking the cross section, NBER working paper.

Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh, 2018, Interpreting factor models, Forthcoming in *The Journal of Finance*.

Ledoit, Olivier, and Michael Wolf, 2008, Robust performance hypothesis testing with the Sharpe ratio, *Journal of Empirical Finance* 15, 850–859.

Lewellen, Jonathan, 2015, The cross-section of expected stock returns, *Critical Finance Review* 4, 1–44.

Lewellen, Jonathan, Stefan Nagel, and Jay Shanken, 2010, A skeptical appraisal of asset pricing tests, *Journal of Financial Economics* 96, 175–194.

Linnainmaa, Juhani T., and Michael R. Roberts, 2018, The history of the cross section of stock returns, Forthcoming in *The Review of Financial Studies*.

Liu, Jianan, Robert F. Stambaugh, and Yu Yuan, 2018, Absolving beta of volatility's effects, Forthcoming in *Journal of Financial Economics*.

McLean, R. David, and Jeffrey Pontiff, 2016, Does academic research destroy stock return predictability?, *The Journal of Finance* 71, 5–32.

Meinshausen, Nicolai, Lukas Meier, and Peter Buhlmann, 2009, p-values for high-dimensional regression, *Journal of the American Statistical Association* 104, 1671–1681.

Meinshausen, Nicolai, and Bin Yu, 2009, Lasso-type recovery of sparse representations for high-dimensional data, *Annals of Statistics* 37, 246–270.

Messmer, Marcial, and Francesco Audrino, 2017, The (adaptive) lasso in the zoo-firm characteristic selection in the cross-section of expected returns, University of St.Gallen working paper.

Nagel, Stefan, 2013, Empirical cross-sectional asset pricing, *Annual Review of Financial Economics* 5, 167–199.

Novy-Marx, Robert, and Mihail Velikov, 2016, A taxonomy of anomalies and their trading costs, *Review of Financial Studies* 29, 104–147.

Pastor, Lubos, and Robert F. Stambaugh, 2000, Comparing asset pricing models: An investment perspective, *Journal of Financial Economics* 56, 335–381.

Politis, Dimitris N., and Joseph P. Romano, 1994, The stationary bootstrap, *Journal of the American Statistical Association* 89, 1303–1313.

Richardson, Scott, İrem Tuna, and Peter Wysocki, 2010, Accounting anomalies and fundamental analysis: A review of recent research advances, *Journal of Accounting and Economics* 50, 410–454.

Rockafellar, R. Tyrrell, 2015, *Convex Analysis* (Princeton university press).

Subrahmanyam, Avanidhar, 2010, The cross-section of expected stock returns: What have we learnt from the past twenty-five years of research?, *European Financial Management* 16, 27–42.

Tibshirani, Robert, 1996, Regression shrinkage and selection via the lasso, *Journal of the Royal Statistical Society. Series B (Methodological)* 58, 267–288.

Wasserman, Larry, and Kathryn Roeder, 2009, High-dimensional variable selection, *Annals of Statistics* 37, 2178–2201.

White, Halbert, 2000, A reality check for data snooping, *Econometrica* 68, 1097–1126.

This page is blank.

Internet Appendix to

A Portfolio Perspective on the Multitude of Firm Characteristics*

Victor DeMiguel Alberto Martín-Utrera Francisco J. Nogales Raman Uppal

March 8, 2018

*DeMiguel, London Business School, e-mail: avmiguel@london.edu; Martín-Utrera, Lancaster University Management School, e-mail: a.martinutrera@lancaster.ac.uk; Nogales, Universidad Carlos III de Madrid, e-mail: FcoJavier.Nogales@uc3m.es; Uppal, Edhec Business School and CEPR, e-mail: raman.uppal@edhec.edu.

IA Robustness checks

We investigate the robustness of our main finding that transaction costs increase the number of significant characteristics to: considering quadratic instead of proportional transaction costs, excluding microcaps, applying the reality check in White (2000), expanding our dataset to consider also characteristics with a large number of missing observations, different subperiods, different levels of risk-aversion, and using different methods to standardize firm characteristics. In addition, we check the robustness of our out-of-sample results to: firm size, shortsale constraints, and the constraint on maximum turnover.

IA.1 Quadratic transaction costs

In the main body of the manuscript, we consider an investor who faces proportional transaction costs, as in Brandt, Santa-Clara, and Valkanov (2009). Proportional transaction costs are a reasonable assumption for the average investor; see Novy-Marx and Velikov (2016) and Chen and Velikov (2017). For large investors, a common assumption is that their price impact is linear on the amount traded, and thus, they face quadratic transaction costs; see, for instance, Korajczyk and Sadka (2004). In this section, we show that our main finding is robust to considering quadratic transaction costs; that is, the number of characteristics that are jointly significant at the 5% level increases from five to 19 in the presence of quadratic transaction costs.

IA.1.1 Modeling quadratic transaction costs

To model market impact costs, we need to track absolute portfolio *positions* instead of portfolio weights. To do this, we consider an investor with a wealth of $\$B$ billion, who holds the following parametric portfolio:

$$w_t(\theta) = Bw_{b,t} + X_t\theta/N_t, \tag{1}$$

where the notation is as in Section 3 of the manuscript. We assume the investor maximizes her mean-variance utility of wealth growth net of quadratic transaction costs:

$$\min_{\theta} \quad \frac{\gamma_a}{2}\theta^\top \widehat{\Sigma}_c \theta + \gamma_a B \theta^\top \widehat{\sigma}_{bc} - \theta^\top \widehat{\mu}_c + TC(\theta), \tag{2}$$

where $\gamma_a = \gamma/B$ is the investor's absolute risk-aversion parameter, defined as the ratio of relative risk-aversion parameter to wealth,[1] and $\mathrm{TC}(\theta)$ is the quadratic transaction cost

$$\mathrm{TC}(\theta) = \frac{1}{T-1}\sum_{t=1}^{T-1}(w_{t+1}(\theta) - w_t^+(\theta))^\top \Lambda_{t+1}(w_{t+1}(\theta) - w_t^+(\theta)), \tag{3}$$

where $\Lambda_t = \mathrm{diag}(\lambda_{1,t}, \ldots, \lambda_{N_t,t})$ is the diagonal matrix whose ith element is the Kyle lambda of the ith stock at time t and $w_t^+(\theta)$ is the parametric portfolio before rebalancing at time $t+1$.

To calibrate our quadratic transaction cost function we rely on the empirical results in Novy-Marx and Velikov (2016), which uses Trade and Quote (TAQ) data to estimate the Kyle lambdas of individual stocks. The paper shows that the R-squared of a cross-sectional regression of log Kyle lambda on log market capitalization is 70% and the slope is not statistically distinguishable from minus one. This suggests that a good approximation to the cross-sectional variation of Kyle lambdas is to assume they are inversely proportional to the market capitalization of each firm. Moreover, Novy-Marx and Velikov (2016) shows that the average price elasticity of supply, defined as the product of Kyle lambda and market capitalization, $\lambda_{i,t} \times me_{i,t}$, is about 6.5. Based on this evidence, we assume the Kyle lambda of the ith stock at time t is $\lambda_{i,t} = 6.5/me_{i,t}$, where $me_{i,t}$ is the market capitalization of the ith stock at time t.

IA.1.2 How many characteristics matter jointly with quadratic costs?

Table IA.1 reports the significance and marginal contributions of each characteristic in the parametric portfolios in the presence of quadratic transaction costs. We consider an investor who allocates $B = \$1$ billion to the benchmark portfolio[2] and has an absolute risk-aversion parameter $\gamma_a = 5/10^9$, which corresponds to a relative risk-aversion parameter of $\gamma = 5$ for an investor with wealth of $B = \$1$ billion. Similar to the analysis in Section 6, we run a screen-and-clean significance test.

Table IA.1 reports the significance and marginal contributions of each characteristic in the parametric portfolios. Our main finding is that, similar to the case with

[1]Because the mean-variance optimization problem is defined in terms of absolute portfolio positions, we formulate the problem in terms of absolute risk-aversion instead of relative risk-aversion, as in Gârleanu and Pedersen (2013).

[2]We have also considered the cases with $B = \$10$ and \$100 billion and the results are similar.

proportional transaction costs, the number of significant characteristics with quadratic transaction costs is substantially larger than for the case without transaction costs. In particular, the number of characteristics that are significant at the 5% level increases from five in the absence of transaction costs to 19 in the presence of quadratic transaction costs. The explanation for this result can be found by comparing the last two columns of Table IA.1, which report the marginal contribution of the characteristic to the transaction cost when traded in the optimal parametric portfolio and in isolation, respectively.[3] We observe that combining characteristics reduces the marginal contribution to quadratic transaction costs by an average of 93%; that is, the benefits from trading diversification are very large also in the presence of quadratic transaction costs.

IA.2 Excluding microcaps

In the main body of the manuscript, we exclude stocks that are below the 20th percentile of market capitalization across the NYSE, AMEX and NASDAQ exchanges. We now check the robustness of our significance results to excluding microcap stocks, which are the stocks below the NYSE 20th percentile. Tables IA.2 and IA.3 report the significance results for the cases where we compute the parametric portfolios without and with transaction costs, respectively. The number of characteristics significant at the 5% level increases from seven in the absence of transaction costs to 12 in the presence of transaction costs. Thus, our central insight that transaction costs increase the dimension of the cross section is robust to excluding microcaps.

IA.3 Reality check

Novy-Marx (2016) explains that *overfitting bias* occurs when researchers consider multiple variables that have been shown individually to predict stock returns in sample. We believe that our main finding that transaction costs increase the dimension of the cross section is not driven by overfitting bias because, although overfitting may increase the

[3]To compute the marginal contribution to transaction costs of a characteristic traded in isolation, we assign to the single characteristic a weight equal to the sum of the absolute values of the optimal parametric portfolio weights for the case when characteristics are traded in combination. This allows for a meaningful comparison for the case with quadratic transaction costs.

number of significant characteristics for both cases without and with transaction costs, it is unlikely to reverse the relative size of these two numbers.

Nonetheless, to check whether overfitting bias affects our significance test, we adapt the reality check in White (2000) to the context of parametric portfolios.[4] First, we set the benchmark portfolio equal to the in-sample optimal parametric portfolio instead of the value-weighted portfolio. By doing this we essentially remove the predictability from our dataset while preserving the correlation structure of the 51 characteristics. Second, we implement a variant of the screen-and-clean test.[5] Specifically, we generate 1,000 bootstrap samples from the original dataset using sampling with replacement. For each of the 1,000 bootstrap samples we use five-fold cross-validation to select the lasso threshold that optimizes the mean-variance criterion and we screen any characteristics with zero weight for the resulting regularized parametric portfolio. We then compute the optimal parametric portfolio of the characteristics that have survived the screen stage for each bootstrap sample, but without the lasso constraint. Finally, we use the percentile-interval method to establish the significance of the characteristics across the 1,000 samples.

We perform the reality check for both the cases with and without transaction costs. In results not reported to conserve space, we find that after removing the predictability from our dataset, none of the 51 characteristics are significant either in the absence or the presence of transaction costs.

IA.4 Characteristics with many missing observations

To ensure our results are reliable, we consider in our main analysis only characteristics with a small proportion of missing observations. Specifically, we drop characteristics with more than 5% of missing observations for more than 5% of those firms with CRSP returns available for the entire sample from 1980 to 2014. In addition, we drop characteristics without any observations for more than 1% of these firms. Consequently, our main analysis is based on 51 out of the 100 characteristics listed in Green, Hand, and Zhang

[4] Harvey and Liu (2018) applies the reality check to the context of sequential factor selection.

[5] We cannot implement the plain screen-and-clean test of Section 3.5 because, given that we are using the in-sample optimal parametric portfolio as the benchmark portfolio, none of the characteristics would survive the screen stage.

(2017). However, to check the robustness of our results, we also run the screen-and-clean significance test of Section 3 using all 100 characteristics in Green et al. (2017).

We find that our results are robust to the inclusion of characteristics with a large proportion of missing observations. First, in the absence of transaction costs, out of the 100 characteristics, only seven are significant, compared to six in the case with 51 characteristics. Second, in the presence of transaction costs, the number of significant characteristics increases to 15, just as in the case with 51 characteristics.

IA.5 Pre- and post-January 2003 analysis

Chordia, Subrahmanyam, and Tong (2014) shows that the magnitude of asset return predictability has decreased in the last decade. To understand how our results vary over time, we test the statistical significance of characteristics for two different subperiods with similar number of observations: May 1988 to December 2002 and January 2003 to December 2014. In results not reported to conserve space, we find that for the period before 2003 the number of significant characteristics increases from eight in the absence of transaction costs to 14 in the presence of transaction costs. For the period after 2003, the number of significant characteristics increases from two in the absence of transaction costs to eight in the presence of transaction costs. Our results confirm both, our main finding that transaction costs increase the dimension of the cross section, as well as the finding in the literature that the magnitude of asset return predictability has decreased in the last decade.

IA.6 Risk-aversion

We now study how our results depend on the risk-aversion parameter. Tables IA.4 and IA.5 report the significance of characteristics for the parametric portfolios with risk-aversion parameter $\gamma = 2$ for the cases without and with transaction costs, respectively. Likewise, Tables IA.6 and IA.7 report the significance of characteristics for $\gamma = 10$ for the cases without and with transaction costs, respectively. Our main finding that transaction costs increase the dimension of the cross section is robust to the investor's risk-aversion parameter: For $\gamma = 2$, the number of significant characteristics increases from six in the absence of transaction costs to 14 in the presence of transaction costs, and for $\gamma = 10$, the number of significant characteristics increases from eight in the absence of transaction costs to 13 in the presence of transaction costs.

IA.7 Quintile-standardized characteristics

We now consider characteristic long-short portfolios defined in terms of the top and bottom quintiles instead of standardizing the characteristics by subtracting the cross-sectional mean and dividing by the standard deviation. Specifically, we assign a weight of $1/Q_t$ to firms in the fifth quintile and $-1/Q_t$ to firms in the first quintile, where Q_t is the number of firms per quintile in month t. Tables IA.8 and IA.9 report the significance of characteristics for the parametric portfolios with quintile-standardized characteristics in the absence and presence of transaction costs, respectively. The tables show that the number of significant characteristics increases from six in the absence of transaction costs to 10 in the presence of transaction costs.

IA.8 Out-of-sample analysis

We now run several checks on the robustness of our out-of-sample analysis.

IA.8.1 Firm size

To study how the out-of-sample performance of the regularized parametric portfolios depends on firm size, we classify stocks (including those with market capitalization below the 20th percentile of our sample, which are excluded in our main analysis) into five size quintiles. Table IA.10 reports the out-of-sample performance of the parametric portfolios in the presence of transaction costs applied to each of the five quintiles separately. It is clear from the table that the performance of the regularized parametric portfolios is better for the quintiles with small stocks. Indeed, this table demonstrates that the regularized parametric portfolios outperform the benchmark value-weighted portfolios significantly for the first four quintiles corresponding to the 80% of smallest stocks. These results are consistent with the findings in Hand and Green (2011) and Fama and French (2008). Also, the regularized parametric portfolios significantly outperform the parametric portfolios based on a small number of characteristics for the first three quintiles corresponding to the 60% of smallest stocks.

IA.8.2 Shortsale constraints

Table IA.11 reports the out-of-sample performance of the regularized portfolios subject to shortsale constraints.[6] Panel A reports the performance for the parametric portfolios with no shortselling, and Panel B reports the performance for the parametric portfolios after scaling the optimal parameter θ so that the short positions in the regularized parametric portfolio amount to around 50%. Panel A shows that with shortsale constraints, although the out-of-sample Sharpe ratio of the regularized parametric portfolios is higher than that of the value-weighted benchmark portfolio, the difference is not statistically significant. Panel B, however, shows that the amount of shorting required for the regularized parametric portfolios to significantly outperform the other portfolios is not large. We observe that for the case with around 50% shortselling, the regularized parametric portfolios attain an out-of-sample Sharpe ratio around 87% higher than the benchmark value-weighted portfolio, 48% higher than that of the portfolios that exploit three characteristics, and around 22% higher than that of the portfolios that exploit four characteristics, with the differences being statistically significant.

IA.8.3 Turnover constraints

In Section 7, we evaluate the out-of-sample performance of the regularized parametric portfolios after controlling their turnover to be around 100% per month. Table IA.12 reports the performance of the regularized parametric portfolios in the *absence* of turnover controls. The regularized parametric portfolios without turnover control have a monthly turnover of around 386%. Despite their high turnover, the table shows that the regularized parametric portfolios attain an out-of-sample Sharpe ratio of returns net of transaction costs around 125% higher than the parametric portfolio that exploits three characteristics, and around 29% higher than the parametric portfolio that exploits four characteristics with the difference significant at the 10% level.

[6] As in Brandt et al. (2009), we compute shortsale constrained portfolios by first computing the unconstrained regularized parametric portfolios, then setting all negative *firm* weights equal to zero, and finally normalizing the resulting vector so that its weights sum to one.

Table IA.1: Significance and marginal contributions with quadratic transaction costs

This table reports the significance and marginal contributions for the parametric portfolios in the presence of quadratic transaction costs, for the case where the investor allocates $B = \$1$ billion to the benchmark portfolio and has an absolute risk-aversion parameter $\gamma_a = 5/B$. We run a screen-and-clean significance test. For the screen step, we calibrate the regularized parametric portfolios with five-fold cross-validation and find that the lasso threshold that maximizes investor's utility is $\delta = 1.5 \times B$. For the clean step, we run the bootstrap experiment for the parametric portfolios using those characteristics with nonzero θ's from the previous step. Characteristic p-values are computed using the percentile method discussed in Section 3.5. We assign three/two/one asterisks (*) to those characteristics whose p-values are lower than 0.01/0.05/0.1, respectively. To compute the optimal parametric portfolio and marginal contributions, we include all characteristics with nonzero θ's for the screen step plus the three characteristics considered in Brandt et al. (2009): size, book to market, and momentum. For each characteristic, the first column gives the acronym, the second the optimal value of the parameter (divided by 100 million) and the significance asterisks, and the next five columns give the marginal contribution of the characteristic to: (i) the characteristic own-variance, (ii) the covariance of the characteristic with the other characteristics in the portfolio, (iii) the covariance of the characteristic with the benchmark portfolio, (iv) the characteristic mean, and (v) the transaction cost. The last column reports the marginal contribution of the characteristic to transaction costs when it is traded in isolation. Contributions that drive the characteristic to be nonzero are in blue sans serif font, and contributions that drive the characteristic toward zero are in red italic font (cf. Footnote 26).

		Marginal contributions					Isolation
Characteristic	Param.	variance	cov (char.)	cov (bench.)	mean	tran. cost	tran. costs
rd_mve	2.592***	0.00009	−0.00033	0.00045	−0.00164	0.00143	0.01365
agr	−1.347***	−0.00005	−0.00041	0.00057	0.00290	−0.00302	0.02405
sgr	−1.284***	−0.00004	−0.00046	0.00075	0.00179	−0.00203	0.02672
chatoia	0.898***	0.00001	0.00005	−0.00005	−0.00077	0.00077	0.02745
turn	−2.270***	−0.00054	−0.00091	0.00279	0.00068	−0.00203	0.01855
retvol	−0.668***	−0.00022	−0.00141	0.00292	0.00323	−0.00452	0.11102
std_turn	0.567***	0.00010	−0.00123	0.00214	−0.00080	−0.00020	0.07697
zerotrade	−0.817***	−0.00007	0.00092	−0.00205	0.00124	−0.00004	0.07316
chcsho	−1.104***	−0.00004	−0.00056	0.00092	0.00228	−0.00260	0.01692
ps	1.017***	0.00003	0.00030	−0.00068	−0.00127	0.00162	0.02187
sue	0.310***	0.00001	0.00015	−0.00019	−0.00254	0.00258	0.07646
egr	−0.819***	−0.00003	−0.00035	0.00041	0.00231	−0.00234	0.02703
idiovol	−1.781***	−0.00067	−0.00109	0.00308	0.00187	−0.00319	0.03448
gma	1.408***	0.00006	−0.00047	0.00069	−0.00066	0.00038	0.01359
ep	0.718**	0.00012	0.00089	−0.00166	−0.00104	0.00169	0.03735
convind	−1.187**	−0.00002	−0.00032	0.00071	0.00051	−0.00088	0.01024
roaq	0.582**	0.00007	0.00078	−0.00114	−0.00215	0.00244	0.03971
cashpr	−0.814**	−0.00004	−0.00063	0.00091	0.00139	−0.00163	0.01514
indmom	1.534**	0.00026	0.00037	−0.00050	−0.00222	0.00209	0.02111
herf	−1.044*	−0.00003	−0.00002	0.00041	0.00061	−0.00098	0.01354
pchcapx_ia	−0.676*	−0.00001	−0.00009	0.00018	0.00093	−0.00100	0.02260
mom12m	0.853*	0.00015	0.00043	−0.00066	−0.00275	0.00282	0.03359
stdcf	−0.828*	−0.00004	−0.00046	0.00068	0.00114	−0.00131	0.01370
lgr	0.329	0.00001	−0.00044	0.00064	0.00182	−0.00203	0.02653
saleinv	0.587	0.00001	0.00027	−0.00064	−0.00005	0.00041	0.00975
hire	−0.288	−0.00001	−0.00051	0.00065	0.00197	−0.00211	0.02518
beta	−1.003	−0.00041	−0.00164	0.00419	−0.00008	−0.00205	0.01902
rsup	0.170	0.00000	0.00015	−0.00017	−0.00054	0.00056	0.04223
mve	−9.661	−0.00033	0.00041	−0.00034	0.00022	0.00003	0.00165
mom36m	−0.627	−0.00003	−0.00034	0.00022	0.00125	−0.00110	0.01653
ear	0.082	0.00000	0.00013	0.00004	−0.00137	0.00120	0.07008
bm_ia	0.498	0.00003	−0.00044	0.00072	−0.00081	0.00051	0.02151
mom6m	0.305	0.00006	0.00068	−0.00093	−0.00247	0.00266	0.05988
baspread	−0.192	−0.00008	−0.00181	0.00329	0.00279	−0.00418	0.08273
chtx	0.115	0.00000	−0.00002	0.00015	−0.00123	0.00110	0.04466
bm	−0.442	−0.00004	0.00084	−0.00082	−0.00205	0.00207	0.02369
salerec	0.482	0.00001	−0.00006	0.00016	−0.00044	0.00033	0.01198
dy	0.591	0.00006	0.00084	−0.00161	−0.00029	0.00100	0.01251
pchgm_pchsale	−0.132	−0.00000	0.00000	−0.00003	−0.00079	0.00082	0.02849
lev	0.661	0.00008	0.00088	−0.00123	−0.00092	0.00119	0.01234
mom1m	0.083	0.00001	0.00045	−0.00109	0.00164	−0.00102	0.22528
std_dolvol	−0.040	−0.00000	0.00061	−0.00150	0.00003	0.00086	0.08448
dolvol	−0.130	−0.00001	−0.00057	0.00139	−0.00025	−0.00056	0.03294
mve_ia	0.318	0.00001	−0.00004	0.00016	−0.00013	−0.00000	0.00613

Table IA.2: Significance test without transaction costs: No microcaps

This table reports, for the case where microcaps are excluded from the dataset, the significance and marginal contributions for the parametric portfolios without transaction costs, for risk-aversion parameter $\gamma = 5$. We run a screen-and-clean significance test. For the screen step, we calibrate the regularized parametric portfolios with five-fold cross-validation and find that the lasso threshold that maximizes investor's utility is $\delta = 40$. For the clean step, we run the bootstrap experiment for the parametric portfolios using those characteristics with nonzero θ's from the previous step. Characteristic p-values are computed using the percentile method discussed in Section 3.5. We assign three/two/one asterisks (*) to those characteristics whose p-values are lower than 0.01/0.05/0.1, respectively. To compute the optimal parametric portfolio and marginal contributions, we include all characteristics with nonzero θ's for the screen step plus the three characteristics considered in Brandt et al. (2009): size, book to market, and momentum. For each characteristic, the first column gives the acronym, the second the optimal value of the parameter and the significance asterisks, and the next five columns give the marginal contribution of the characteristic to: (i) the characteristic own-variance, (ii) the covariance of the characteristic with the other characteristics in the portfolio, (iii) the covariance of the characteristic with the benchmark portfolio, and (iv) the characteristic mean. Contributions that drive the characteristic to be nonzero are in blue sans serif font, and contributions that drive the characteristic toward zero are in red italic font (cf. Footnote 26)

		Marginal contributions			
Characteristic	Param.	variance	cov (char.)	cov (bench.)	mean
ear	9.43***	*0.00173*	−0.00086	*0.00025*	−0.00112
indmom	−4.12***	*−0.00698*	0.00843	*−0.00052*	*−0.00093*
retvol	−6.74***	*−0.02308*	0.01817	0.00338	0.00153
std_turn	9.57***	*0.01983*	−0.02187	*0.00223*	−0.00019
bm_ia	5.63***	*0.00315*	−0.00312	*0.00075*	−0.00078
sue	8.48**	*0.00120*	*0.00002*	−0.00015	−0.00107
ps	6.87**	*0.00175*	−0.00033	−0.00076	−0.00066
rd_mve	5.96*	*0.00213*	−0.00137	*0.00043*	−0.00119
std_dolvol	3.15*	*0.00088*	*0.00001*	−0.00061	−0.00028
chmom	−2.14*	*−0.00241*	0.00208	*−0.00076*	0.00110
agr	−8.33	*−0.00439*	0.00134	0.00098	0.00207
roaq	1.89	*0.00138*	*0.00086*	−0.00077	−0.00148
mom1m	−0.96	*−0.00146*	0.00152	*−0.00095*	0.00089
egr	−4.74	*−0.00218*	*−0.00059*	0.00083	0.00194
gma	1.46	*0.00085*	−0.00136	*0.00070*	−0.00019
bm	0.85	*0.00080*	*0.00090*	−0.00065	−0.00106
mve	−0.19	*−0.00007*	0.00023	*−0.00053*	0.00037
mom12m	−0.76	*−0.00150*	0.00346	*−0.00045*	*−0.00151*

Table IA.3: Significance test with transaction costs: No microcaps

This table reports, for the dataset without microcaps, the significance and marginal contributions for the parametric portfolios in the presence of transaction costs, for risk-aversion parameter $\gamma = 5$. We run a screen-and-clean significance test. For the screen step, we calibrate the regularized parametric portfolios with five-fold cross-validation and find that the lasso threshold that maximizes investor's utility is $\delta = 20$. For the clean step, we run the bootstrap experiment for the parametric portfolios using those characteristics with nonzero θ's from the previous step. Characteristic p-values are computed using the percentile method discussed in Section 3.5. We assign three/two/one asterisks (*) to those characteristics whose p-values are lower than 0.01/0.05/0.1, respectively. To compute the optimal parametric portfolio and marginal contributions, we include all characteristics with nonzero θ's for the screen step plus the three characteristics considered in Brandt et al. (2009): size, book to market, and momentum. For each characteristic, the first column gives the acronym, the second the optimal value of the parameter and the significance asterisks, and the next five columns give the marginal contribution of the characteristic to: (i) the characteristic own-variance, (ii) the covariance of the characteristic with the other characteristics in the portfolio, (iii) the covariance of the characteristic with the benchmark portfolio, (iv) the characteristic mean, and (v) the transaction cost. The last column reports the marginal contribution of the characteristic to transaction costs when it is traded in isolation. Contributions that drive the characteristic to be nonzero are in blue sans serif font, and contributions that drive the characteristic toward zero are in red italic font (cf. Footnote 26)

		Marginal contributions					Isolation
Characteristic	Param.	variance	cov (char.)	cov (bench.)	mean	tran. cost	tran. costs
std_turn	0.28***	0.00057	−0.00300	0.00223	−0.00019	0.00039	0.00472
rd_mve	7.13***	0.00255	−0.00205	0.00043	−0.00119	0.00026	0.00051
ear	0.33***	0.00006	−0.00019	0.00025	−0.00112	0.00100	0.00324
bm_ia	2.53***	0.00142	−0.00181	0.00075	−0.00078	0.00043	0.00125
mom36m	1.17***	0.00069	−0.00256	0.00046	0.00107	0.00034	0.00177
chmom	−1.01**	−0.00114	0.00181	−0.00076	0.00110	−0.00101	0.00421
egr	−4.09**	−0.00188	−0.00032	0.00083	0.00194	−0.00057	0.00124
gma	3.84**	0.00223	−0.00283	0.00070	−0.00019	0.00009	0.00089
ps	2.72**	0.00069	0.00020	−0.00076	−0.00066	0.00053	0.00141
retvol	−0.24**	−0.00081	−0.00318	0.00338	0.00153	−0.00092	0.00492
zerotrade	−0.23**	−0.00008	0.00094	−0.00108	0.00064	−0.00042	0.00182
aeavol	−0.17**	−0.00004	−0.00050	0.00070	0.00021	−0.00037	0.00293
indmom	−0.59*	−0.00101	0.00293	−0.00052	−0.00093	−0.00047	0.00252
saleinv	2.80*	0.00040	0.00023	−0.00058	−0.00020	0.00016	0.00069
chatoia	2.45*	0.00023	0.00020	−0.00016	−0.00067	0.00040	0.00118
herf	−2.40*	−0.00067	0.00024	0.00049	0.00011	−0.00017	0.00076
sue	0.21	0.00003	0.00039	−0.00015	−0.00107	0.00080	0.00219
stdcf	−2.31	−0.00105	−0.00008	0.00063	0.00073	−0.00023	0.00065
roaq	0.64	0.00047	0.00107	−0.00077	−0.00148	0.00071	0.00186
rsup	0.40	0.00010	0.00012	−0.00015	−0.00048	0.00042	0.00169
pchcapx_ia	−1.46	−0.00039	−0.00034	0.00034	0.00069	−0.00030	0.00123
pchgm_pchsale	1.22	0.00014	−0.00018	−0.00000	−0.00027	0.00031	0.00120
chtx	0.16	0.00005	−0.00048	0.00030	−0.00052	0.00065	0.00233
baspread	−0.28	−0.00122	−0.00323	0.00380	0.00128	−0.00062	0.00344
cashpr	−2.07	−0.00092	−0.00051	0.00063	0.00095	−0.00016	0.00085
pricedelay	0.06	0.00001	0.00093	−0.00079	−0.00033	0.00018	0.00286
sgr	−1.69	−0.00099	−0.00107	0.00100	0.00127	−0.00021	0.00120
mom6m	0.63	0.00131	0.00066	−0.00076	−0.00101	−0.00020	0.00397
turn	−0.53	−0.00156	−0.00169	0.00288	0.00054	−0.00017	0.00171
beta	0.83	0.00341	−0.00810	0.00413	0.00052	0.00004	0.00122
salerec	1.41	0.00038	0.00015	−0.00020	−0.00045	0.00011	0.00080
lev	1.84	0.00219	−0.00072	−0.00071	−0.00081	0.00005	0.00081
chcsho	−0.66	−0.00026	−0.00205	0.00098	0.00175	−0.00042	0.00126
convind	−0.72	−0.00013	−0.00082	0.00068	0.00035	−0.00008	0.00076
agr	−1.15	−0.00060	−0.00185	0.00098	0.00207	−0.00060	0.00123
ep	0.84	0.00112	0.00078	−0.00153	−0.00050	0.00013	0.00119
idiovol	−0.01	−0.00002	−0.00407	0.00335	0.00081	−0.00006	0.00114
mom1m	0.08	0.00013	0.00111	−0.00095	0.00089	−0.00117	0.00850
bm	−2.05	−0.00194	0.00361	−0.00065	−0.00106	0.00004	0.00114
mve	−2.27	−0.00079	0.00100	−0.00053	0.00037	−0.00005	0.00059
mom12m	−0.34	−0.00068	0.00251	−0.00045	−0.00151	0.00013	0.00274

Table IA.4: Significance test without transaction costs: Risk-aversion of $\gamma = 2$

This table reports the significance and marginal contributions for the parametric portfolios without transaction costs, for risk-aversion parameter $\gamma = 2$. We run a screen-and-clean significance test. For the screen step, we calibrate the regularized parametric portfolios with five-fold cross-validation and find that the lasso threshold that maximizes investor's utility is $\delta = 75$. For the clean step, we run the bootstrap experiment for the parametric portfolios using those characteristics with nonzero θ's from the previous step. Characteristic p-values are computed using the percentile method discussed in Section 3.5. We assign three/two/one asterisks (*) to those characteristics whose p-values are lower than 0.01/0.05/0.1, respectively. To compute the optimal parametric portfolio and marginal contributions, we include all characteristics with nonzero θ's for the screen step plus the three characteristics considered in Brandt et al. (2009): size, book to market, and momentum. For each characteristic, the first column gives the acronym, the second the optimal value of the parameter and the significance asterisks, and the next four columns give the marginal contribution of the characteristic to: (i) the characteristic own-variance, (ii) the covariance of the characteristic with the other characteristics in the portfolio, (iii) the covariance of the characteristic with the benchmark portfolio, and (iv) the characteristic mean. Contributions that drive the characteristic to be nonzero are in blue sans serif font, and contributions that drive the characteristic toward zero are in red *italic* font (cf. Footnote 26).

		Marginal contributions			
Characteristic	Param.	variance	cov (char.)	cov (bench.)	mean
sue	51.10***	*0.00347*	−0.00085	−0.00008	−0.00254
retvol	−27.05***	*−0.03520*	0.03081	0.00117	0.00323
bm	9.54**	*0.00325*	−0.00087	−0.00033	−0.00205
gma	14.88**	*0.00251*	−0.00213	*0.00028*	−0.00066
agr	−25.44**	*−0.00389*	0.00077	0.00023	0.00290
mom1m	−7.75**	*−0.00508*	0.00388	*−0.00043*	0.00164
bm_ia	16.42*	*0.00341*	−0.00289	*0.00029*	−0.00081
beta	6.69*	*0.01103*	−0.01262	*0.00167*	−0.00008
rd_mve	14.66	*0.00210*	−0.00064	*0.00018*	−0.00164
std_turn	18.71	*0.01266*	−0.01271	*0.00085*	−0.00080
chcsho	−13.50	*−0.00192*	*−0.00073*	0.00037	0.00228
zerotrade	−6.41	*−0.00216*	0.00174	*−0.00082*	0.00124
mve	−9.17	*−0.00124*	0.00115	*−0.00014*	0.00022
mom12m	−10.59	*−0.00752*	0.01054	*−0.00026*	*−0.00275*

Table IA.5: Significance test with transaction costs: Risk-aversion of $\gamma = 2$

This table reports the significance and marginal contributions for the parametric portfolios in the presence of transaction costs, for risk-aversion parameter $\gamma = 2$. We run a screen-and-clean significance test. For the screen step, we calibrate the regularized parametric portfolios with five-fold cross-validation and find that the lasso threshold that maximizes investor's utility is $\delta = 75$. For the clean step, we run the bootstrap experiment for the parametric portfolios using those characteristics with nonzero θ's from the previous step. Characteristic p-values are computed using the percentile method discussed in Section 3.5. We assign three/two/one asterisks (*) to those characteristics whose p-values are lower than 0.01/0.05/0.1, respectively. To compute the optimal parametric portfolio and marginal contributions, we include all characteristics with nonzero θ's for the screen step plus the three characteristics considered in Brandt et al. (2009): size, book to market, and momentum. For each characteristic, the first column gives the acronym, the second the optimal value of the parameter and the significance asterisks, and the next five columns give the marginal contribution of the characteristic to: (i) the characteristic own-variance, (ii) the covariance of the characteristic with the other characteristics in the portfolio, (iii) the covariance of the characteristic with the benchmark portfolio, (iv) the characteristic mean, and (v) the transaction cost. The last column reports the marginal contribution of the characteristic to transaction costs when this is traded in isolation. Contributions that drive the characteristic to be nonzero are in blue sans serif font, and contributions that drive the characteristic toward zero are in red *italic* font (cf. Footnote 26).

		Marginal contributions					Isolation
Characteristic	Param.	variance	cov (char.)	cov (bench.)	mean	tran. cost	tran. costs
rd_mve	31.87***	*0.00457*	−0.00332	*0.00018*	−0.00164	*0.00022*	*0.00047*
agr	−18.14***	*−0.00278*	0.00016	0.00023	0.00290	*−0.00051*	*0.00115*
sue	8.40***	*0.00057*	*0.00062*	−0.00008	−0.00254	*0.00143*	*0.00224*
turn	−9.17***	*−0.00866*	0.00727	0.00111	0.00068	*−0.00040*	*0.00168*
retvol	−5.26***	*−0.00685*	0.00378	0.00117	0.00323	*−0.00133*	*0.00445*
std_turn	3.45***	*0.00234*	−0.00320	*0.00085*	−0.00080	*0.00081*	*0.00478*
zerotrade	−5.25***	*−0.00177*	0.00220	*−0.00082*	0.00124	*−0.00085*	*0.00218*
beta	10.98***	*0.01811*	−0.01992	*0.00167*	−0.00008	*0.00023*	*0.00111*
chtx	3.85**	*0.00030*	−0.00016	*0.00006*	−0.00123	*0.00103*	*0.00222*
mom1m	−1.77**	*−0.00116*	0.00207	*−0.00043*	0.00164	*−0.00211*	*0.00833*
ps	11.93**	*0.00151*	−0.00056	−0.00027	−0.00127	*0.00060*	*0.00130*
chatoia	12.45**	*0.00032*	*0.00004*	−0.00002	−0.00077	*0.00043*	*0.00107*
gma	16.17**	*0.00273*	−0.00248	*0.00028*	−0.00066	*0.00013*	*0.00081*
herf	−13.82**	*−0.00138*	0.00073	0.00017	0.00061	*−0.00012*	*0.00065*
pchgm_pchsale	9.58*	*0.00037*	*0.00004*	−0.00001	−0.00079	*0.00039*	*0.00112*
bm_ia	7.30*	*0.00152*	−0.00125	*0.00029*	−0.00081	*0.00026*	*0.00116*
stdcf	−14.34*	*−0.00294*	0.00175	0.00027	0.00114	*−0.00021*	*0.00060*
bm	5.70*	*0.00194*	*0.00029*	−0.00033	−0.00205	*0.00015*	*0.00104*
chcsho	−7.07	*−0.00101*	*−0.00127*	0.00037	0.00228	*−0.00038*	*0.00114*
chmom	−1.98	*−0.00077*	0.00135	*−0.00029*	0.00044	*−0.00073*	*0.00393*
ear	1.09	*0.00007*	*0.00059*	*0.00002*	−0.00137	*0.00070*	*0.00305*
baspread	1.34	*0.00233*	−0.00594	*0.00131*	*0.00279*	−0.00049	*0.00299*
idiovol	−4.59	*−0.00693*	0.00386	0.00123	0.00187	*−0.00004*	*0.00091*
ep	3.57	*0.00233*	−0.00077	−0.00066	−0.00104	*0.00014*	*0.00107*
roaq	−0.18	*−0.00009*	0.00219	*−0.00046*	*−0.00215*	0.00050	*0.00171*
mve	−3.63	*−0.00049*	0.00045	*−0.00014*	0.00022	*−0.00004*	*0.00038*
mom12m	−1.71	*−0.00121*	0.00393	*−0.00026*	*−0.00275*	0.00030	*0.00255*

Table IA.6: Significance test without transaction costs: Risk-aversion of $\gamma = 10$

This table reports the significance and marginal contributions for the parametric portfolios without transaction costs, for risk-aversion parameter $\gamma = 10$. We run a screen-and-clean significance test. For the screen step, we calibrate the regularized parametric portfolios with five-fold cross-validation and find that the lasso threshold that maximizes investor's utility is $\delta = 15$. For the clean step, we run the bootstrap experiment for the parametric portfolios using those characteristics with nonzero θ's from the previous step. Characteristic p-values are computed using the percentile method discussed in Section 3.5. We assign three/two/one asterisks (*) to those characteristics whose p-values are lower than 0.01/0.05/0.1, respectively. To compute the optimal parametric portfolio and marginal contributions, we include all characteristics with nonzero θ's for the screen step plus the three characteristics considered in Brandt et al. (2009): size, book to market, and momentum. For each characteristic, the first column gives the acronym, the second the optimal value of the parameter and the significance asterisks, and the next four columns give the marginal contribution of the characteristic to: (i) the characteristic own-variance, (ii) the covariance of the characteristic with the other characteristics in the portfolio, (iii) the covariance of the characteristic with the benchmark portfolio, and (iv) the characteristic mean. Contributions that drive the characteristic to be nonzero are in blue sans serif font, and contributions that drive the characteristic toward zero are in red *italic* font (cf. Footnote 26).

		Marginal contributions			
Characteristic	Param.	variance	cov (char.)	cov (bench.)	mean
agr	−6.82***	*−0.00522*	0.00118	0.00115	0.00290
ps	7.84***	*0.00494*	−0.00232	−0.00135	−0.00127
sue	9.17***	*0.00311*	−0.00019	−0.00038	−0.00254
mom1m	−2.20***	*−0.00723*	0.00776	*−0.00217*	0.00164
std_turn	8.63***	*0.02919*	−0.03265	*0.00427*	−0.00080
dolvol	−5.98**	*−0.00898*	0.00645	0.00278	*−0.00025*
retvol	−4.58**	*−0.02980*	0.02074	0.00583	0.00323
bm_ia	3.70**	*0.00385*	−0.00447	*0.00144*	−0.00081
gma	2.37*	*0.00200*	−0.00272	*0.00138*	−0.00066
rd_mve	3.29	*0.00236*	−0.00161	*0.00089*	−0.00164
chcsho	−1.34	*−0.00095*	*−0.00317*	0.00184	0.00228
bm	1.13	*0.00192*	*0.00176*	−0.00164	−0.00205
mve	2.35	*0.00159*	−0.00112	−0.00068	*0.00022*
mom12m	−2.50	*−0.00886*	0.01294	*−0.00132*	*−0.00275*

Table IA.7: Significance test with transaction costs: Risk-aversion of $\gamma = 10$

This table reports the significance and marginal contributions for the parametric portfolios in the presence of transaction costs, for risk-aversion parameter $\gamma = 10$. We run a screen-and-clean significance test. For the screen step, we calibrate the regularized parametric portfolios with five-fold cross-validation and find that the lasso threshold that maximizes investor's utility is $\delta = 15$. For the clean step, we run the bootstrap experiment for the parametric portfolios using those characteristics with nonzero θ's from the previous step. Characteristic p-values are computed using the percentile method discussed in Section 3.5. We assign three/two/one asterisks (*) to those characteristics whose p-values are lower than 0.01/0.05/0.1, respectively. To compute the optimal parametric portfolio and marginal contributions, we include all characteristics with nonzero θ's for the screen step plus the three characteristics considered in Brandt et al. (2009): size, book to market, and momentum. For each characteristic, the first column gives the acronym, the second the optimal value of the parameter and the significance asterisks, and the next five columns give the marginal contribution of the characteristic to: (i) the characteristic own-variance, (ii) the covariance of the characteristic with the other characteristics in the portfolio, (iii) the covariance of the characteristic with the benchmark portfolio, (iv) the characteristic mean, and (v) the transaction cost. The last column reports the marginal contribution of the characteristic to transaction costs when this is traded in isolation. Contributions that drive the characteristic to be nonzero are in blue sans serif font, and contributions that drive the characteristic toward zero are in red *italic* font (cf. Footnote 26).

		Marginal contributions					Isolation
Characteristic	Param.	variance	cov (char.)	cov (bench.)	mean	tran. cost	tran. costs
rd_mve	5.99***	*0.00429*	−0.00375	*0.00089*	−0.00164	*0.00020*	*0.00047*
gma	4.09***	*0.00345*	−0.00432	*0.00138*	−0.00066	*0.00015*	*0.00081*
ps	3.38***	*0.00213*	−0.00017	−0.00135	−0.00127	*0.00066*	*0.00130*
sue	1.69***	*0.00057*	*0.00091*	−0.00038	−0.00254	*0.00143*	*0.00224*
retvol	−0.77***	*−0.00501*	*−0.00288*	0.00583	0.00323	*−0.00117*	*0.00445*
std_turn	0.58***	*0.00197*	−0.00633	*0.00427*	−0.00080	*0.00090*	*0.00478*
zerotrade	−0.81***	*−0.00136*	0.00495	*−0.00410*	0.00124	*−0.00074*	*0.00218*
herf	−3.98**	*−0.00198*	0.00072	0.00083	0.00061	*−0.00018*	*0.00065*
chtx	0.72**	*0.00028*	−0.00035	*0.00030*	−0.00123	*0.00100*	*0.00222*
turn	−1.07**	*−0.00505*	*−0.00088*	0.00557	0.00068	*−0.00032*	*0.00168*
chatoia	3.13**	*0.00040*	*0.00006*	−0.00011	−0.00077	*0.00042*	*0.00107*
agr	−2.61**	*−0.00200*	*−0.00154*	0.00115	0.00290	*−0.00051*	*0.00115*
stdcf	−2.26**	*−0.00232*	0.00003	0.00135	0.00114	*−0.00020*	*0.00060*
mom1m	−0.37*	*−0.00122*	0.00376	*−0.00217*	0.00164	*−0.00200*	*0.00833*
chmom	−0.45*	*−0.00088*	0.00268	*−0.00146*	0.00044	*−0.00078*	*0.00393*
mve	−1.63	*−0.00110*	0.00160	*−0.00068*	0.00022	*−0.00004*	*0.00038*
pchgm_pchsale	1.49	*0.00029*	*0.00016*	−0.00006	−0.00079	*0.00040*	*0.00112*
bm_ia	1.21	*0.00126*	−0.00209	*0.00144*	−0.00081	*0.00021*	*0.00116*
sgr	−2.25	*−0.00154*	*−0.00159*	0.00150	0.00179	*−0.00015*	*0.00111*
chcsho	−1.39	*−0.00099*	*−0.00277*	0.00184	0.00228	*−0.00037*	*0.00114*
bm	0.81	*0.00138*	*0.00217*	−0.00164	−0.00205	*0.00013*	*0.00104*
pchcapx_ia	−1.11	*−0.00047*	*−0.00060*	0.00036	0.00093	*−0.00021*	*0.00118*
roaq	−0.34	*−0.00083*	0.00488	*−0.00228*	*−0.00215*	0.00038	*0.00171*
ep	0.74	*0.00243*	*0.00176*	−0.00331	−0.00104	*0.00017*	*0.00107*
dolvol	−0.25	*−0.00037*	*−0.00200*	0.00278	*−0.00025*	*−0.00015*	*0.00195*
idiovol	0.17	*0.00131*	−0.00930	*0.00615*	*0.00187*	−0.00003	*0.00091*
mom12m	−0.74	*−0.00264*	0.00663	*−0.00132*	*−0.00275*	0.00008	*0.00255*

Table IA.8: Significance without transaction costs: Quintile-standardized characteristics

This table reports the significance and marginal contributions for the parametric portfolios without transaction costs, for risk-aversion parameter $\gamma = 5$. We sort firms by each characteristic every month into quintiles, assign a weight of $1/Q_t$ to firms in the fifth quintile, a weight of $-1/Q_t$ to firms in the first quintile, and a zero weight to the remaining firms, where Q_t is the number of firms in each quintile in month t. We run a screen-and-clean significance test. For the screen step, we calibrate the regularized parametric portfolios with five-fold cross-validation and find that the lasso threshold that maximizes investor's utility is $\delta = 10$. For the clean step, we run the bootstrap experiment for the parametric portfolios using those characteristics with nonzero θ's from the previous step. Characteristic p-values are computed using the percentile method discussed in Section 3.5. We assign three/two/one asterisks (*) to those characteristics whose p-values are lower than 0.01/0.05/0.1, respectively. To compute the optimal parametric portfolio and marginal contributions, we include all characteristics with nonzero θ's for the screen step plus the three characteristics considered in Brandt et al. (2009): size, book to market, and momentum. For each characteristic, the first column gives the acronym, the second the optimal value of the parameter and the significance asterisks, and the next four columns give the marginal contribution of the characteristic to: (i) the characteristic own-variance, (ii) the covariance of the characteristic with the other characteristics in the portfolio, (iii) the covariance of the characteristic with the benchmark portfolio, and (iv) the characteristic mean. Contributions that drive the characteristic to be nonzero are in blue sans serif font, and contributions that drive the characteristic toward zero are in red *italic* font (cf. Footnote 26).

		Marginal contributions			
Characteristic	Param.	variance	cov (char.)	cov (bench.)	mean
rd_mve	2.11***	*0.03183*	−0.02588	*0.00284*	−0.00879
agr	−5.32***	*−0.01694*	0.00844	0.00075	0.00775
gma	2.79***	*0.01346*	−0.01336	*0.00247*	−0.00257
sue	9.03***	*0.01184*	−0.00350	−0.00051	−0.00783
mom1m	−1.08***	*−0.01310*	0.01112	*−0.00330*	0.00528
std_turn	4.37***	*0.06424*	−0.06666	*0.00681*	−0.00439
ep	3.21*	*0.04109*	−0.03176	−0.00464	−0.00469
stdcf	−2.26*	*−0.02143*	0.01442	0.00325	0.00376
retvol	−3.00	*−0.07674*	0.06176	0.00901	0.00597
roaq	−1.14	*−0.01051*	0.02086	*−0.00321*	*−0.00714*
bm	0.04	*0.00033*	*0.00895*	−0.00292	−0.00636
chcsho	−0.21	*−0.00081*	*−0.00885*	0.00317	0.00649
mve	−2.16	*−0.01538*	0.01612	0.00020	*−0.00094*
mom12m	−1.27	*−0.02163*	0.03398	*−0.00312*	*−0.00924*

Table IA.9: Significance test with transaction costs: Quintile-standardized characteristics

This table reports the significance and marginal contributions for the parametric portfolios in the presence of transaction costs, for risk-aversion parameter $\gamma = 5$. We sort firms by each characteristic every month into quintiles, assign a weight of $1/Q_t$ to firms in the fifth quintile, a weight of $-1/Q_t$ to firms in the first quintile, and a zero weight to the remaining firms, where Q_t is the number of firms in each quintile in month t. We run a screen-and-clean significance test. For the screen step, we calibrate the regularized parametric portfolios with five-fold cross-validation and find that the lasso threshold that maximizes investor's utility is $\delta = 5$. For the clean step, we run the bootstrap experiment for the parametric portfolios using those characteristics with nonzero θ's from the previous step. Characteristic p-values are computed using the percentile method discussed in Section 3.5. We assign three/two/one asterisks (*) to those characteristics whose p-values are lower than 0.01/0.05/0.1, respectively. To compute the optimal parametric portfolio and marginal contributions, we include all characteristics with nonzero θ's for the screen step plus the three characteristics considered in Brandt et al. (2009): size, book to market, and momentum. For each characteristic, the first column gives the acronym, the second the optimal value of the parameter and the significance asterisks, and the next five columns give the marginal contribution of the characteristic to: (i) the characteristic own-variance, (ii) the covariance of the characteristic with the other characteristics in the portfolio, (iii) the covariance of the characteristic with the benchmark portfolio, (iv) the characteristic mean, and (v) the transaction cost. The last column reports the marginal contribution of the characteristic to transaction costs when this is traded in isolation. Contributions that drive the characteristic to be nonzero are in blue sans serif font, and contributions that drive the characteristic toward zero are in red *italic* font (cf. Footnote 26).

		Marginal contributions					Isolation
Characteristic	Param.	variance	cov (char.)	cov (bench.)	mean	tran. cost	tran. costs
rd_mve	2.46***	*0.03707*	−0.03198	*0.00284*	−0.00879	*0.00086*	*0.00160*
agr	−3.56***	*−0.01133*	0.00400	0.00075	0.00775	*−0.00117*	*0.00223*
gma	2.58***	*0.01243*	−0.01266	*0.00247*	−0.00257	*0.00033*	*0.00141*
sue	4.01***	*0.00525*	*0.00003*	−0.00051	−0.00783	*0.00306*	*0.00368*
mom1m	−0.55***	*−0.00668*	0.00864	*−0.00330*	0.00528	*−0.00395*	*0.01160*
std_turn	1.08***	*0.01595*	−0.02139	*0.00681*	−0.00439	*0.00303*	*0.00651*
chmom	−0.53**	*−0.00442*	0.00664	*−0.00230*	0.00200	*−0.00192*	*0.00608*
ep	2.04**	*0.02609*	−0.01738	−0.00464	−0.00469	*0.00061*	*0.00206*
ps	2.05**	*0.00388*	*0.00014*	−0.00133	−0.00364	*0.00095*	*0.00201*
bm	0.53**	*0.00414*	*0.00457*	−0.00292	−0.00636	*0.00057*	*0.00348*
retvol	−0.64	*−0.01639*	0.00310	0.00901	0.00597	*−0.00168*	*0.00635*
stdcf	−1.08	*−0.01020*	0.00353	0.00325	0.00376	*−0.00034*	*0.00136*
chcsho	−1.06	*−0.00415*	*−0.00488*	0.00317	0.00649	*−0.00063*	*0.00180*
roaq	0.16	*0.00145*	*0.00807*	−0.00321	−0.00714	*0.00083*	*0.00290*
mve	−1.08	*−0.00767*	0.00885	0.00020	*−0.00094*	*−0.00044*	*0.00195*
mom12m	−0.58	*−0.00987*	0.02257	*−0.00312*	*−0.00924*	*−0.00034*	*0.00421*

Table IA.10: Out-of-sample performance: Size quintiles

This table reports the out-of-sample annualized Sharpe ratio of returns net of transaction costs for the regularized parametric portfolios applied to each of the five quintiles of stocks sorted by size, for risk-aversion parameter $\gamma = 5$. Panel A reports the performance for the portfolios that do not use any characteristics, which are the benchmark value-weighted portfolio (VW) and the equally weighted portfolio ($1/N$). Panel B reports the performance of two parametric portfolios that exploit a small number of characteristics, and the regularized parametric portfolio that exploits a large set of 51 characteristics. The first parametric portfolio exploits the size, book-to-market, and momentum characteristics (Size/val./mom.). The second parametric portfolio exploits the size, book-to-market, asset growth, and gross profitability characteristics (Size/val./inv./prof.). The third portfolio is the regularized parametric portfolio that exploits all 51 characteristics (Regularized). The lasso threshold is calibrated using cross-validation over the estimation window. We test the significance of the difference of the Sharpe ratio of each portfolio with that of the regularized parametric portfolio. Three/two/one asterisks (*) indicate that the difference is significant at the 0.01/0.05/0.1 level.

Policy	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
Panel A: Portfolios with no characteristics					
VW	0.341***	0.402***	0.458***	0.546***	0.568
1/N	0.442***	0.391***	0.438***	0.530***	0.558
Panel B: Portfolios with characteristics					
Size/val./mom.	0.852***	0.889***	0.666***	0.601*	0.456
Size/val./inv./prof.	0.933***	1.072***	0.856**	0.796	0.360
Regularized	1.734	1.456	1.008	0.769	0.497

Table IA.11: Out-of-sample performance with shortsale constraints

This table reports the out-of-sample performance of the regularized parametric portfolios in the presence of transaction costs and shortsale constraints, for risk-aversion parameter $\gamma = 5$. Panel A reports the performance for the parametric portfolios with no shortselling, and Panel B reports the performance for the parametric portfolios with 50% shortselling. Each panel reports the results for four portfolios: the benchmark value-weighted portfolio (VW), which has zero shortselling in both panels, two parametric portfolios that exploit a small number of characteristics, and the regularized parametric portfolio that exploits a large set of 51 characteristics. The first parametric portfolio exploits the size, book-to-market, and momentum characteristics (Size/val./mom.). The second parametric portfolio exploits the size, book-to-market, asset growth, and gross profitability characteristics (Size/val./inv./prof.). For the regularized parametric portfolio (Regularized), the lasso threshold is calibrated using cross-validation over the estimation window. For each portfolio, the first column reports the monthly turnover, and the next three columns report the out-of-sample annualized mean, standard deviation, and Sharpe ratio of returns, net of transaction costs. We test the significance of the difference of the Sharpe ratio of each portfolio with that of the regularized parametric portfolio. Three/two/one asterisks (*) indicate that the difference is significant at the 0.01/0.05/0.1 level.

Policy	Turnover	Mean	SD	SR
Panel A: Portfolios with no shortselling				
VW	0.050	0.085	0.150	0.567
Size/val./mom.	0.233	0.102	0.177	0.576^{**}
Size/val./inv./prof.	0.186	0.109	0.186	0.586^{**}
Regularized	0.301	0.125	0.187	0.669
Panel B: Portfolios with 50% shortselling				
VW	0.050	0.085	0.150	0.567^{***}
Size/val./mom.	0.429	0.119	0.165	0.721^{***}
Size/val./inv./prof.	0.319	0.132	0.152	0.868^{***}
Regularized	0.451	0.155	0.147	1.059

Table IA.12: Out-of-sample performance without turnover constraint

This table reports the out-of-sample performance of the regularized parametric portfolios that do not control for turnover in the presence of transaction costs, for risk-aversion parameter $\gamma = 5$. Panel A reports the performance for the portfolios that do not use any characteristics, which are the benchmark value-weighted portfolio (VW) and the equally weighted portfolio ($1/N$). Panel B reports the performance of two parametric portfolios that exploit a small number of characteristics, and the regularized parametric portfolio that exploits a large set of 51 characteristics. The first parametric portfolio exploits the size, book-to-market, and momentum characteristics (Size/val./mom.). The second parametric portfolio exploits the size, book-to-market, asset growth, and gross profitability characteristics (Size/val./inv./prof.). The third portfolio is the regularized parametric portfolio that exploits all 51 characteristics (Regularized). The lasso threshold is calibrated using cross-validation over the estimation window. For each portfolio, the first column reports the monthly turnover, and the next three columns report the out-of-sample annualized mean, standard deviation, and Sharpe ratio of returns, net of transaction costs. We test the significance of the difference of the Sharpe ratio of each portfolio with that of the regularized parametric portfolio. Three/two/one asterisks (*) indicate that the difference is significant at the 0.01/0.05/0.1 level.

Policy	Turnover	Mean	SD	SR
Panel A: Portfolios with no characteristics				
VW	0.050	0.085	0.150	0.567**
1/N	0.134	0.085	0.177	0.482***
Panel B: Portfolios with characteristics				
Size/val./mom.	1.167	0.161	0.300	0.537***
Size/val./inv./prof.	1.863	0.358	0.381	0.939*
Regularized	3.859	0.738	0.611	1.209

References

Brandt, Michael W., Pedro Santa-Clara, and Rossen Valkanov, 2009, Parametric portfolio policies: Exploiting characteristics in the cross-section of equity returns, *Review of Financial Studies* 22, 3411–3447.

Chen, Andrew, and Mihail Velikov, 2017, Accounting for the anomaly zoo: A trading cost perspective, Federal Reserve Board working paper.

Chordia, Tarun, Avanidhar Subrahmanyam, and Qing Tong, 2014, Have capital market anomalies attenuated in the recent era of high liquidity and trading activity?, *Journal of Accounting and Economics* 58, 41–58.

Fama, Eugene F., and Kenneth R. French, 2008, Dissecting anomalies, *The Journal of Finance* 63, 1653–1678.

Gârleanu, Nicolae, and Lasse H. Pedersen, 2013, Dynamic trading with predictable returns and transaction costs, *The Journal of Finance* 68, 2309–2340.

Green, Jeremiah, John R.M. Hand, and X. Frank Zhang, 2017, The characteristics that provide independent information about average us monthly stock returns, Forthcoming in *The Review of Financial Studies*.

Hand, John R.M., and Jeremiah Green, 2011, The importance of accounting information in portfolio optimization, *Journal of Accounting, Auditing & Finance* 26, 1–34.

Harvey, Campbell R., and Yan Liu, 2018, Lucky factors, SSRN working paper.

Korajczyk, Robert A., and Ronnie Sadka, 2004, Are momentum profits robust to trading costs?, *The Journal of Finance* 59, 1039–1082.

Novy-Marx, Robert, 2016, Backtesting strategies based on multiple signals, Simon Graduate School of Business working paper.

Novy-Marx, Robert, and Mihail Velikov, 2016, A taxonomy of anomalies and their trading costs, *Review of Financial Studies* 29, 104–147.

White, Halbert, 2000, A reality check for data snooping, *Econometrica* 68, 1097–1126.