Nonlinearities Everywhere: Sparse Supervised Learning of Market Anomalies*

Gabriel Kaiser †

March 29, 2020

Abstract

This paper compares non-parametric multiple-step approaches to identify a reduced information set that produces expected return spreads in the cross-section. I show that sparse models can respond to the multidimensional challenge imposed by the 'zoo of anomalies'. The models employed are outlier resistant, computationally efficient on big data sets and capable to handle more features than observations, i.e. $p \ll N$. The determined information subset discovered ranges between 2 and 8 out of 90 characteristics of a novel market anomaly database. The selected set of characteristics with predictive power is not time-invariant. Non-linearity's seems to matter in a multivariate setting for the cross-section of expected returns. Out-of-sample results suggest that Elastic Net penalties tend to overfit the data, while SCAD penalties suggest sparse models with remarkable Sharpe ratios.

JEL Classification: C14, C52, C58, G12.

Keywords: Anomalies, Cross-Sectional Return Predictability, LASSO, Model Selection.

*I would like to thank my supervisors Tibor Neugebauer, Julien Penasse and Andrea Tamoni. Further, I am grateful to Arnaud Dufays, Simon Rottke, Ulf von Lilienfeld-Toal, Christian von Drathen, Denis Gromb for insightful conference discussions and seminar and conference participants at Luxembourg School of Finance, the London School of Economics, HEC Liège. I also thank Thummim Cho, Kenneth French and Tyler Muir for providing data. The estimations presented in this paper were carried out using the HPC facilities of the University of Luxembourg Varrette, Bouvry, Cartiaux, and Georgatos (2014) − see `https://hpc.uni.lu`. All errors are the sole responsibility of the author.

†Correspondence: Department of Finance, LSF, University of Luxembourg, Tel.: (+352) 46 66 44 5948, Email: gabriel.kaiser@uni.lu, Web: `https://gabrielkaiserqfin.github.io`

Contents

1 Introduction

In financial economics, a market anomaly is typically a significant nonzero intercept with respect to a factor model, such as the Capital Asset Pricing Model (CAPM) of Black, Jensen, and Scholes (1972). Given the efficient market hypothesis, this intercept should be indistinguishable from zero at least most of the time. Nonetheless, fifty years of comprehensive research in empirical asset pricing led to a remarkably large number of cross-sectional risk factors.

However, in times of distress, hedge returns, obtained from portfolio sorting on firm characteristics, have substantial commonality. This implies low pricing kernel dimensions and hence, a small set of characteristics proxying for latent risk factors. Certainly, some covariates are just a product of others.

Given this 'zoo of anomalies', a simple portfolio sorting is unable to answer this multivariate question of dimension reduction as recommended by Cochrane (2011). Alternative econometric techniques that deal with the curse of dimensionality are necessary. With this mindset, I employ non-parametric cross-sectional regressions and propose penalization terms that are novel in the finance literature. The suggested single-stage technique detects a reduced and independently important information subset of around 13 firm characteristics from a new dataset of 85 market anomalies. The determined information subset discovered is time-varying and hence state-dependent.

In particular, I show that single-stage models with a minimax concave penalty (MCP) or smoothly clipped absolute deviations (SCAD) penalty on group level, introduce enough sparsity to answer the multidimensional challenge imposed by the substantial set of return predictors. What is more, the suggested methodology allows the estimated coefficients to reach large values quicker than the adaptive group LASSO. Important firm characteristics have therefore a greater chance to enter the reduced subset. A positive side effect of a one-stage approach is computational efficiency.

After selecting a model, I suggest a supervised learning approach to identify the nonlinear functional form of the partial contribution of the selected variables to the conditional expected returns. In principal, I use a generalized adaptive model (GAM) which derives

the unknown functional form that maps characteristics to expected returns during model estimation. This non-parametric element recovers marginal nonlinearities, which a parametric model would miss. Furthermore, GAMs have the ability to control the smoothness of the explanatory variables, which is of great importance whenever observations are noisy. In contrast to many other machine learning techniques, the additive strucutre models the marginal contribution of each variable independently and therefore is easy to interpret.
The motivation for the existence of additional risk factors besides the systematic component that are suggested by the countless number of empirical results, is presented by the ICAPM by Merton (1969). It has great intuition to understand the importance of risk sharing. Agents hedge their correlated outside income exposure and therefore generate time-varying risk exposure that materializes in market mispricings. These mispricings occasionally translate in risk premiums, once the prices converged to an equitable spread. As a result, empirical distributions of hedge returns are typically left-skewed and leptokurtic, representing an insurance-like payoff. We can think of this, as writing out-of-the money call options on a characteristic. Daniel, Jagannathan, and Kim (2012) provide characteristic-specific evidence in line with these arguments. They link state dependency to the 12 month momentum anomaly. Momentum crashes when the market sharply recovers, since stocks in the short leg win more than stocks in the long leg. These tail events are most probably the reason that momentum yields a premium and that it is classified as a market anomaly in the first place. Similarly, Dobrynskaya (2017) presents evidence that the profitability of momentum strategies result from the exposure to downside risk and thus rejects the hypothesis that momentum is an anomaly or a market mispricing. Vuolteenaho and Tuomo (2004) go one step further and decompose the covariance structure of value and size with the market into discount rate and cash flow betas. Since they do this individually for the short and long leg of the anomaly strategy, they are able to prove that only exposure to cash flow news earn high risk prices. All these papers provide evidence to consider hedge returns as insurances.
Given this evidence, I define a market anomaly as a non-identified object resulting from statistical biases, market frictions, latent risk factors or as actual mispricing. McLean and Pontiff (2016) provide evidence that extending the sample periods identifies statisti-

cal biases.

The market anomalies considered in this paper last from January 1965 to December 2018, which is on average twice the original sample period as presented in Table 2. Hence, anomalies subject to data mining shall not survive the extended sample. On the other hand, if the market anomaly is an adequate risk premium, mean returns should line up with covariances, which results in insurance-like payoffs of hedge returns. Distributional-wise, these returns should show a slightly positive mean, a larger median, and a leptokurtic body. Exogenous shocks, such as the Great Depression, will lead to heavy tails on the downside that cause the negative skewness in the joint distribution of returns. Conversely, I classify hedge returns of positive skewness with a mode above limits to arbitrage as market mispricing.

Reading the Hansen and Jagannathan (1991) bound from left to right, allows us to conclude that tradable Sharpe ratios form the lower bound of the coefficient of variation of the stochastic discount factor. Considering the inequality as binding, restricts the set of possible discount factors that can price a given set of hedge returns. Following Mehra and Prescott (1985), an equity premium of 8% and a standard deviation of about 16% for annual US equity returns, leads to tradeable Sharpe ratios of around 0.5. The average risk-free rate is about one percent, and hence the expected kernel is slightly below one. Consequently, the standard deviation of the pricing kernel is above 50% per year. Assuming an orthonormal linear factor model, the sum of the volatilities of a factor mimicking portfolio needs to satisfy this lower bound. While this fact suggests a property that needs to be satisfied, it does not advise on the dimensionality of possible risk factors but allows for reconciliation in this respect.

I structure the remainder of this paper as follows. In section 2, I review related literature. Section 3 introduces the methodology. Section 4 describes the data sources used in this paper. In section 5, I present the in- and out-of-sample comparison. Finally, section 6 summarizes the findings.

2 Related Literature

With the Modern Portfolio Theory (MPT), Markowitz (1952) unveils the trade-off between anticipated returns and variance of returns. The subsequent Capital Asset Pricing Model (CAPM) of Sharpe (1964), Lintner (1965) and Mossin (1966) illustrates risk as the systematic component stemming from the covariance with the market portfolio. Fama and French (1992, 1993) identify specific firm characteristics (C), in particular valuation ratios and market capitalization, to provide independent risk exposures additional to the market risk factor, leading to the Fama-French three-factor (FF3) model.

The FF3 model has become a successful tool in the academic world as well as in the industry, although it is rejected in explaining average returns due to its significant nonzero intercepts. As a consequence, later research started fishing for factors that jointly eliminate any anomalies. The exercise of this paper is mainly motivated by the EMH. In a world where the EMH holds, the intercept of an asset pricing model should be indistinguishable from zero, at least most of the time.

Fifty years of comprehensive research later, empirical asset pricing has thus led to a fairly large number of anomalies in the cross-section of average equity returns. Harvey, Liu, and Zhu (2016) summarize 316 of such factors in the published literature and show that specifically after 2000 this number increases at an exponential rate. They also find that fifty percent of anomalies result from statistical biases in the analysis, a common concern intensively documented in the literature, as for example by Barras, Scaillet, and Wermers (2010). Other contributions that address the factor zoo are Green, Hand, and Zhang (2017), McLean and Pontiff (2016) and many others. Green et al. (2017) for instance encounter that only the minority of characteristics present independent information after adjusting p-values and controlling for micro-cap stocks in their Fama-MacBeth regressions. In contrast, Novy-Marx and Velikov (2016) as well as A. Y. Chen and Velikov (2019) argue that fifty percent of remaining anomalies are in fact subject to transaction costs. Accordingly, their paper suggests that the entire zoo vanishes after publication. Other causes for anomalies can be the limit of arbitrage, in particular deteriorating funding conditions, as suggested by Adrian, Etula, and Muir (2014).

A different approach to overcome these adverse effects of multiple null hypothesis testing is to estimate the rate of type I errors or false discovery rate (FDR). However, this approach is only applicable if the econometrician possess information regarding the number of multiple testings. Another concept to overcome data mining as well as overfitting is to introduce regularization terms that penalize parameter estimates and thus, propose sparsity. What is more, regularization extracts independent firm characteristics of the dozens of variables found in the literature that produce expected return spreads.

One common stylized fact within this research field is to assume that "means and covariances are stable functions of characteristics not of individual securities", as examined by Cochrane (2011). We can denote such a relationship by

$$R_{i,t+1} = g(C_{i,t}, \theta) + \epsilon_{i,t+1}, \tag{1}$$

where $g(.)$ represents an unknown function with parameters θ and a individual firm characteristic $C_{i,t}$, which formulates the conditional expectation of excess returns $\mathbb{E}(R^e_{t+1}|C_t)$. Individual excess or long-short returns are denoted as $R_{i,t+1}$. This paper focuses on the functional form that relates the multidimensional vector of firm characteristics to individual expected returns and covariances, thereby allowing for non-linearity in the cross-section.

Whereas the majority of above-mentioned papers suggest a static model, time-varying expected returns demand dynamic models, leading to a relatively new strand in the finance literature and the involvement of machine learning in asset pricing. This literature is essentially motivated by leading theoretical asset pricing models such as Campbell and Cochrane (1999), Bansal and Yaron (2004) and Santos and Veronesi (2004). An example of its empirical implementation is for instance a paper by Haddad, Kozak, and Santosh (2018), which identifies the time-varying function of the conditional expectation of excess returns in terms of characteristics using Principal Component Analysis (PCA). DeMiguel, Martin-Utrera, Nogales, and Uppal (2019) take a portfolio perspective extending Brandt, Santa-Clara, and Valkanov (2009) and include a trading cost term in the objective function adhered by a L1 regularization term. They find that the number of significant characteristics increases from six to fifteen after including transaction costs due to a cancellation effect in rebalancing. Kozak, Nagel, and Santosh (2019) construct a robust stochastic discount factor (SDF) in a Bayesian setting and find that sparsity (L1 norm) is elusive, whereas L2 norm penalization works well. Freyberger, Neuhierl, and Weber (2020) apply a non-parametric adaptive group Least Absolute Shrinkage and Selection Operator (LASSO) on ranked firm characteristics and observe high Sharpe-ratios. From 62 firm characteristics around 13 seem to provide independent information in a non-linear framework. L. Chen, Pelger, and Zhu (2019) employ a Deep Neural Network (DNN) with long-short-term-memory and find that non-linearities are essential in a multidimensional factor world. Kelly, Pruitt, and Su (2019) suggest using an Instrumented Principal Component Analysis (IPCA) in order to incorporate both time-varying factor loadings and latent factors. Their suggested five-factor model outperforms previous well-established models. Gu, Kelly, and Xiu (2018) compare a set of techniques including generalized linear models, dimension reduction, boosted regression trees, random forests, and neural networks (NN), and conclude that NN and trees outperform the rest. Chinco, Clark-Joseph, and Ye (2019) successfully predict high-frequency expected returns employing LASSO regressions.

This paper extends the literature of machine learning in asset pricing by introducing a generalized additive model (GAM) enhanced with regularization. Within this framework I am able to model the cross-section, in particular function $g(.)$, and can thus allow for flexible semi-parametric modeling. Besides, GAMs have several advantages compared to a Generalized Linear Model (GLM) or FamaMacBeth regression. First, allowing for non-linear relationships between characteristics and expected returns makes any individual transformation of characteristic variables redundant. Secondly, the use of natural splines increases prediction accuracy and decreases the confident intervals of the forecasts. Third, the additive structure enables tractability and interpretability of individual characteristics similar to a standard OLS model. Thereby, interactions can still be added manually via two-dimensional smoothers. All in all, GAMs are therefore a suitable alternative to both non-parametric and parametric models. Furthermore, in comparison to decision trees, GAMs are supervised learning approaches whenever parameters are selected non-parmetrically and have typically higher prediction accuracy.

3 Model and Methodology

In general, regularization methods try to optimize the trade-off between model bias and efficiency. With an increasing number of explanatory variables, the dimension of parameter estimates grows and hence, its variance rises. Eventually, a large variance can lead to inefficiencies and over-fitting. Contrarily, reducing the number of explanatory variables might produce a larger bias with higher efficiency, thereby under-fitting the true model. This is called the curse of dimensionality. Consequently, one of the major challenges in empirical asset pricing is to identify an independent subset of variables that produces expected return spreads in the cross-section.

Assumption *Return moments such as mean and covariance are stable and well-behaved functions of firm characteristics not of individual stocks.*

$$g(C) = \mathbb{E}[R_{i,t+1}|C_{1,it} = size_{it}, \ldots, C_{Z,it} = mom_{it}] \tag{2}$$

I denote this set of firm characteristics or anomalies by the matrix $C_t \in \mathbb{R}^{N \times Z}$, where N is the number of firms and Z is the set of 90 firm characteristics of a novel market anomaly database. Model 2 relates a unknown function $g(\cdot)$ with characteristics as arguments, to conditional expected returns of individual stocks.

Cochrane (2011) argues that portfolio sorts, the most common methodology to study risk premia, are equivalent to non-parametric cross-sectional predictive regressions. Freyberger et al. (2020) establish the theoretical relationship and provide empirical evidence of the superiority relative to portfolio sorts and linear regressions. In their extensive empirical section, they employ the adaptive group LASSO approach as in Huang, Horowitz, and Wei (2010) to select nonlinear mappings of rank-normalized firm characteristics.

I pick up on their findings and replicate the adaptive group LASSO approach as benchmark against other regularization approaches as discussed in the next subsection. Similarly, I rank-normalize each firm characteristic across firms, which I denote by $\tilde{C}_{i,t}$. The predictive regression is then,

$$R_{i,t+1} = \sum_{z=1}^{Z} g_z(\tilde{C}_{i,t}, \theta) + \epsilon_{i,t+1}, \tag{3}$$

In general, identifying any unknown function $g(\cdot)$ is a difficult task, specifically if covariates are slightly collinear and thus challenging to separate. Additionally, the signal-to-noise ratio in empirical asset pricing is exceptionally low compared to other disciplines such as electrical engineering. Another aspect, is the univariate evidence of non-linear functional forms that map certain characteristics to expected returns as in Cattaneo, Crump, Farrell, and Schaumburg (2018). These unfortunate aspects motivate the usage of various feature subset selection techniques to benefit from the contrasting model properties.

3.1 Models and Regularization

As a consequence, I choose a model that allows for $K + 2$ group-levels in order to cluster interpolations of each characteristic and is flexible enough to allow for various regularization terms of distinctive properties. In particular, I consider four distinct penalties for the linear case: LASSO, Adaptive LASSO, SCAD, Elastic Net. For the non-linear case, I utilize their grouped analogues: **Group LASSO, Adaptive Group LASSO, Group SCAD, Group Elastic Net**. The word 'group' indicates that penalization is applied on group level. Without loss of generality, I describe all regularization terms including the group-levels since it is straightforward to obtain the original form by setting $(K + 2) = 1$ and hence dropping the corresponding sum in equation 4.
The baseline model utilized for table 5, 6, 7 and 8 is then given for each step $s \in \{1, 2\}$ by:

$$\hat{\beta}_s = \underset{\beta_{s,z}}{\arg\min} \sum_{i=1}^{N} \left(R_{i,t+1} - \sum_{z=1}^{Z} \beta'_{s,z} \cdot g_z(\tilde{C}_{i,t,z}, \theta) \right)^2 + \lambda_s \sum_{z=1}^{Z} R_{type}(\beta_{s,z}) \tag{4}$$

where $\hat{\beta}_s$ is a $Z \times (K+2)$ matrix of estimated coefficients with row columns $\beta_{s,z}$ of dimension $(K+2) \times 1$ and the hyperparameter λ_s is 10-fold cross-validated. The subscript of λ and β indicates the step of the model, since for the adaptive group model a second step and thus a second tuning parameter λ_2 exists. The OLS component can be broken down into future returns and a vectorized sum. The latter part can also be written as

$$\sum_{z=1}^{Z} \sum_{k=1}^{K+2} \beta_{s,z,k} \cdot g_z(\tilde{C}_{i,t,z}, \theta) \tag{5}$$

I now describe the various types of regularization terms studied in this paper.
First, the group LASSO algorithm utilizes the ℓ_1 norm, which controls the complexity of the model. Although the ℓ_1 norm is convex, it is non-differentiable around zero and thus harder to optimize via quadratic programming. A unique solution assuming that all variable pairs are not perfectly collinear exists. Hence, highly correlated covariates are problematic. Otherwise, when $N \ll p$ its solution is not uniquely defined and can select at most N variables before it saturates due to the nature of ℓ_1 optimization. In other words, it tends to select few correlated variables and not necessarily the optimal subset. Nonetheless, if a unique solution exists, it pushes the entries towards zero and thus enforces sparsity. Hence, it is an appropriate methodology for variable subset selection as suggested by Tibshirani (1996).
A disadvantage of LASSO is that it does not satisfy the weak Oracle properties, which implies that the parameter estimates obtained from variable selection are not unbiased. As a countermeasure, we can employ adaptive LASSO which is the extension that tries to fix this shortcoming and leads us to the second method. The adaptive component realizes in a second estimation step, subsequently to a standard LASSO estimation. In the second stage, additional weights w_z, one per group z, are introduced in the regularization term, which leads to equation 7. This weight w_z is just the inverse of the root of the squared sum per group of step one estimated coefficients. Consequently, adaptive LASSO penalizes those coefficients with lower initial estimates more and in the extreme case of zero predictive power per group with infinity. In other words, the regularization term pushes all coefficients of an entire group of each characteristic to zero whenever the predictive

power is little and thus introduces sparsity. The oracle procedure ensures an optimal estimation rate and is thus more efficient as illustrated in Zou (2006). Nevertheless, it yields improved but still biased parameter estimates.

Therefore a third alternative, the Smoothly Clipped Absolute Deviation (SCAD) penalty is employed to compare results. It is a single step approach that extends the concept of LASSO up to a certain threshold, and sets the penalization term to zero thereafter. Hence, substantial coefficients are nearly unbiased as described in more detail in Fan and Li (2001). The switch from the penalized to unpenalized form is smooth and corresponds to a quadratic spline function with knots at λ and $\gamma\lambda$.

The last method, introduces shrinkage in form of a RIDGE penalty in combination to the LASSO regularization term. Its authors Zou and Hastie (2005) argue that it is more stable than LASSO regarding feature selection. Furthermore, it yields improved predictions than LASSO when variables are correlated. The intuition behind Elastic Nets are that unknown groups of variables, such as related genes in microarrays are grouped together while producing sparse solutions.

$$R_{\text{LASSO}} = \sum_{k=1}^{K+2} \|\boldsymbol{\beta}_{z,k}\|_{2,1} \tag{6}$$

$$R_{\text{Adaptive}} = w_z \cdot R_{\text{LASSO}} \tag{7}$$

$$R_{\text{SCAD}} = \sum_{k=1}^{K+2} p_{\lambda,\gamma}^{SCAD} \|\boldsymbol{\beta}_{z,k}\|_{2,1} \tag{8}$$

$$R_{\text{Elastic Net}} = \sum_{j=1}^{p} d_j |\beta_j| + \frac{(1-\alpha)}{\alpha} \lambda \sum_{j=1}^{p} d_j \beta_j^2 \tag{9}$$

The $\ell_{2,1}$ norm or group LASSO regularizer controls the complexity of the non-linearity's in the model as defined in equation 10. It pushes the entries towards zero and thus enforces sparsity at the group level. Group LASSO drops or selects the smoothing spline of each characteristic simultaneously. Hence, it encourages sparsity in the number of characteristics included, rather than in the number of non-linear expansions of characteristics.

I set the knot parameter of the quadratic spline function of the SCAD term in equation

12 to three, $\lambda = 3$, a common choice in the machine learning literature.

$$\|\boldsymbol{\beta}_{z,k}\|_{2,1} = \left(\sum_{i \in z,k} \beta_j^2\right)^{\frac{1}{2}} \tag{10}$$

$$w_z = \begin{cases} \left(\sum\limits_{k=1}^{K+2} \hat{\beta}_{zk}^2\right)^{-\frac{1}{2}} & \text{if } \sum_{k=1}^{K+2} |\hat{\beta}_{zk}| \neq 0 \\ \infty & \text{if } \sum_{k=1}^{K+2} |\hat{\beta}_{zk}| = 0 \end{cases} \tag{11}$$

$$p_{\lambda,\gamma}^{SCAD}(\beta) = \begin{cases} \lambda|\beta| & \text{if} \quad |\beta| \leq \lambda \\ \frac{|\beta|^2 - 2\gamma\lambda|\beta| + \lambda^2}{2(1-\gamma)} & \text{if} \quad \lambda < |\beta| \leq \gamma\lambda \\ \frac{(\gamma+1)\lambda^2}{2} & \text{if} \quad |\beta| \geq \gamma\lambda. \end{cases} \tag{12}$$

Each group is necessary to cluster smooth functions of a particular characteristic C_z together. Hence, each group is just a mapping of each characteristic, i.e.

$$g(\tilde{C}_{t,z}) : \mathbb{R}^N \mapsto \mathbb{R}^{(K+2)\times N}$$

where K can be interpreted as the number of portfolio sorts. The most direct way to represent this transformation and to capture non-linearities in the cross-section is by using quadratic polynomials in an additive form, i.e. $1, \tilde{C}_{t,z}, \tilde{C}_{t,z}^2$ and by adding a truncated quadratic basis function $h(.)$ per portfolio sort K. For K sorts we require $K-1$ cutoff points, typically referred as knots.

$$h(\tilde{C}_{t,z}, \theta) = (\tilde{C}_{t,z} - \theta)_+^2 = \begin{cases} (\tilde{C}_{t,z} - \theta)^2 & \text{if } \tilde{C}_{t,z} > \theta \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

where θ comprises percentile values of the empirical distribution of the rank-normalized firm characteristics $\tilde{C}_{t,z}$.

In both steps and with all regularization types, I select the penalization parameter based on the minimum 10-fold cross validated Bayesian Information Criterion (BIC) following Yuan and Lin (2006).

3.2 Time-variation

Another question that arises is whether this set of variables is time invariant or not. McLean and Pontiff (2016) for instance, test 97 cross-sectional anomalies of 79 studies and find that long-short portfolio returns are 32% smaller out-of-sample due to publication-informed trading, implying a significant variation in the informative set of variables over time. Penasse (2017) goes one step further and shows that ignoring time variation, whenever present, can even lead to false discoveries.

Given the ambiguity of many dynamics found to impact the time variation of financial market anomalies such as publication dates or funding constraints, the aim of this paper is to identify a subset of characteristics with predictive power per time point. I do so by generalizing the baseline model of equation 4, to a time-varying as follows:

$$\hat{\beta}_{t,s} = \underset{\beta_{t,s,z}}{\arg\min} \sum_{i=1}^{N} \left(R_{i,t+1} - \sum_{z=1}^{Z} \beta'_{t,s,z} \cdot g_{t,z}(\tilde{C}_{i,t,z}, \theta) \right)^2 + \lambda_{t,s} \sum_{z=1}^{Z} R_{type}(\beta_{t,s,z}) \tag{14}$$

where $\hat{\beta}_t$ is a $T \times Z(K+2)$ matrix of estimated coefficients with row columns $\beta_{t,z}$ of dimension $T \times (K+2)$. For this reason, Table 9 presents a rolling window model for a constant estimation window of 26 years and a holding period of one year.

4 Data and Market Anomalies

I obtain daily, monthly, and annual equity returns from the Center for Research in Security Prices database. Following standard conventions, I use domestically incorporated ordinary common stocks listed on NYSE, AMEX or NASDAQ and include delisting returns following the methodology of Shumway and Warther (1999).

Firm characteristics include quarterly, and annual balance sheet information from the Standard & Poor's Capital IQ Compustat industrial database. To mitigate a potential survivorship bias due to backfilling, I exclude firms with less than three years of Compustat data. I delete duplicates based on the global company identifier, fiscal year and stock level. The time horizon spans from January 1965 to December 2018 and is a result of data availability. These restrictions lead to an average of 1.4 million observations per market anomaly or 127 million data points in total.

The risk-free rate is the 1-month TBill return from Ibbotson and Associates. EDHEC Risk Institute provides the long-short equity benchmark, which is the weighted average of the following six hedge fund indexes: CSFB, HFR, HF Net, Barclay, CISDM, and Altvest. In the appendix, Table 2 includes the identifier, anomaly abbreviation, authors and year of publication of the corresponding anomaly that I analyze in this paper. Subsection C in the appendix states the details regarding the factor construction following the previously mentioned literature. I do not investigate any arbitrage opportunities that result from violations of the law of one price or any parity violations since these rare events are either of short life or occur in different asset classes.

Successive anomaly portfolio construction uses the cross-sectional heterogeneity in specific factor characteristics. This dispersion allows me to compute decile factor portfolios from $P1$ to $P10$ for each anomaly factor. Accordingly, the difference in the first and the last decile forms a long-short self-financing portfolio. I invert the factor signal whenever the original paper considers the lowest factor characteristic to construct the long position. Hence, the long leg of the portfolio includes the most undervalued and the short leg the most overvalued securities. All portfolios are value-weighted if not specified else to ensure that results are not driven by illiquid firms. Table 3 reports these portfolio decile returns. In order to generate comparative statics I measure these returns per month and adjust annual or quarterly accounting data as if they were monthly. For quarterly or yearly observations I do not re-balance monthly, but express the returns per month. Table 1 presents summary statistics of the joint distribution of both, value-weighted and equally-weighted, dollar-neutral portfolio returns and benchmark portfolio returns.

4.1 Supervised Learning

In the second step, a predictive model is estimated using the subset-selected firm characteristics denoted by $\tilde{Z}$, which were obtained through single-stage group modeling. I use different forms of Generalized Additive Models (GAMs), a modern semi-parametric approach that is classified as supervised machine learning. While focusing on the technicalities of GAMs in this section, section 2 discusses the advantages compared to OLS. A GAM is a generalized version of Generalized Linear Models (GLMs) and hence, GLMs and OLS regressions are actually special cases of GAMs. As in the GLM case, the word 'generalized' represents the fact that the response variable can follow any distribution of the exponential family and not only the standard Gaussian one. As before, we can write the predictive model as a GAM, i.e.

$$R_{i,t+1} = \sum_{z=1}^{\tilde{Z}} g_{tz}(\tilde{C}_{i,t}) + \epsilon_{i,t+1}, \qquad R_{i,t+1} \sim EF(\mu_i, \sigma^2), \tag{15}$$

where g(.) is again a nonlinear transformation on each characteristic.

The main variations across the models presented differ in their functional form and its corresponding constraints. While the two-step model selection procedure used truncated power functions or B-splines, I also compare results obtained from *P-splines* (penalized), *smoothing splines* and *natural splines*. The latter smoother has the largest number of constraints but does the best job in modeling returns. I therefore go one step further, relax the distributional assumption of normality and incorporate natural splines. Hence, I obtain an even more general version of GAM models, which assumes any parametric distribution for the response variable. By considering higher moments, hedge portfolio returns that are clearly heavy-tailed, negatively skewed and have heterogeneous variance, are now more accurately estimated.

In this context, *smoothing splines* as in Reinsch (1967), are basically all functions g(.) that are continuous on the closed interval of the rank-normalized firm characteristics, have continuous first and second derivatives, and minimize:

$$\arg\min_{\beta_z} \sum_{i=1}^{N} \left(R_{i,t+1} - \sum_{z=1}^{\tilde{Z}} \beta_z' \cdot g_z(\tilde{C}_{i,t,z}, \theta) \right) + \zeta \int g''(\tilde{C}_{i,t,z})^2 d\tilde{C}_{i,t,z}, \tag{16}$$

where is a ζ non-negative tuning parameter. We can again interpret the latter component as a penalization term that penalizes the variability of $g(.)$, since the second derivative of g measures the extent to which the slope (first derivative g') is changing. For instance, the second derivative of a straight line is zero and so is the integral of the regularization term. On the other hand, the penalty term becomes relatively large for flexible functions such as $\tilde{C}^{6}_{i,tz}$. These continuous constraints avoid the knot selection issues that B-Splines, or local, and quantile regressions suffer from.

In addition to that, I employ *natural splines* that are extensions of smoothing splines in the sense that they have additional boundary constraints. These constraints force the function $g(.)$ to be linear at the boundaries and thus resistant to extreme values. As the number of firms with extreme characteristic values is scarce, non-linear functions at the extremes might lead to over-fitting. I therefore employ natural splines which have more stable estimates at the boundaries.

5 Empirical Results

I start the analysis by reporting monthly average returns for ten value-weighted portfolios for each market anomaly, as seen in Table 3. Parentheses indicate standard errors and significant average returns at a confidence interval of 95% and are printed in bold. All 90 market anomalies explain expected returns in the long leg (P10). On the short leg, around 90% are significant. Moreover, approximately half of the hedge returns show a t-statistic above two and eight have a monthly average return above 1%.

5.1 Identification of Mispricing

Since this paper focuses on the identification of a subset of truly predictive characteristics and thus, a reduction of the anomaly zoo, it is important to distinguish them from anomalies that arise through other sources such as statistical biases, risk factors, or limits to arbitrage. Harvey et al. (2016) suggest a multiple testing framework, such as estimating the rate of type I errors, false discovery rate (FDR) or to consider a new t-statistic threshold of 3. Since I am unable to verify the truthful number of repetitions per anomaly which I need in such a multiple testing framework, I estimate Jensen alphas with respect to a benchmark risk model using an extended sample period.

Anomalies identified through data mining or risk premiums should resolve in insignificant intercepts. Table 4 presents unconditional Information ratios as well as spanning tests, using two common factor-mimicking portfolios (FMP). The statistics are reported on unadjusted hedge returns and in excess of the CAPM or FF3 factor model according to

$$\alpha_z = \mathbb{E}[R_z^e] - \sum_{f=1}^{F} \beta_{f,z} \cdot \mathbb{E}[\text{FMP}_f^e]. \tag{17}$$

Sharpe ratios and Information ratios are annualized by the approximation $\sqrt{1}2 \cdot \hat{\alpha}/\hat{\sigma}_\epsilon$. The standard errors of the Information ratios are estimated by moving blocks bootstrapping the data 10,000 times per month to retain the autocorrelation structure in the returns. What is surprising, the majority of anomalies does not outperform the market in terms of Sharpe ratios, implying a statistical bias or time variation of expected returns. Depending on the periods measured, the value-weighted buy-and-hold Sharpe ratio for US stocks is around 0.5, given an equity premium of 8% and a market volatility of 16% per year. Moreover, the t-statistic threshold of 3 is exceeded in 13 of the cases. Considering risk-adjusted returns and thus Information ratios, it seems that more than half of all anomalies have market or factor loadings of negligible size. Figure 4 presents these estimated factor betas per anomaly sorted by CAPM betas in order to compare factor loadings across risk factor models. The anomaly "Illiquidity" (ill) is the perfect example of the joint hypothesis problem. It lacks a CAPM beta but significantly loads on the SMB and HML factors. This indicates that returns are higher due to greater risk premiums on small or illiquid firms and thus, "Illiquidity" is not (entirely) a mispricing.
However, the big challenge is to understand the multivariate counterparts of univariate regressions as Cochrane (2011) advocates in his presidential address. This indicates that we have to move from univariate to multivariate regressions with the ability to filter for the important covariates, given that they are multicollinear as suggested by the pairwise correlation plot, see figure 1. These heat maps present pairwise Pearson correlations between rank-normalized firm characteristics. Dark blue (red) indicates positively (negatively) correlated features and uncorrelated variables are transparent. Features are ordered by complete hierarchical clustering. The correlation across characteristics is relatively strong for some clusters and even after controlling for micro stocks at the NYSE 10% level in 2 as well as at the NYSE 20% level in 3 these co-movements seem to even increase. I try

solve this problem by applying penalized multivariate regression as discussed in the next subsection.

5.2 In-Sample Comparison

The baseline setting includes all firms of the Compustat North America data. In addition, to that I restrict the set of firms to a subset with a market capitalization above the NYSE 10%, respectively NYSE 20% percentile in order to exclude effects which manifest due to limits of arbitrage. Hou, Mo, Xue, and Zhang (2019), for instance, replicate a set of well-established anomalies and find that most returns are concentrated in micro-caps or illiquid stocks.

Table 5 reports selected characteristics filtered by the penalized regressions from a total universe of 90 market anomalies. The columns report model selection results for different steps, firms and regularization terms. Characteristics are rank-normalized per month and the predictive model is estimated via pooling omitting the time index as suggested by Cochrane (2011).

I present four different regularization approaches that turned out to be useful in the asset pricing literature: LASSO, Elastic Net, Adaptive LASSO and SCAD. I set the α parameter to 0.5 which weights the LASSO and the RIDGE penalty, and hence equally weights sparsity and shrinkage. For robustness, I also set α to 0.25 (0.75) and thus allow for more (less) shrinkage than sparsity. However, results are similar and Elastic Net converges to LASSO already at an α level of 0.5 given the low signal-to-noise ratio. As a consequence, I omit the LASSO results in this table and present LASSO in the online appendix. Additionally to that but not represented in this paper, I find that MCP penalties result in almost identical feature selections as the SCAD penalties. This finding is intuitive given their similarities within their functional form.

Model (1) and (3) yield similar in-sample Sharpe ratios with less selected features compared to model (2). Surprisingly, once controlled for micro stocks as suggested by Asness, Frazzini, Israel, Moskowitz, and Pedersen (2018), Sharpe ratios increase across all models. Moreover, in terms of Sharpe ratios the SCAD and the adaptive LASSO model exceed the Elastic Net model although the number of covariates is 40% smaller.

Furthermore, by comparing the multivariate in-sample results with the univariate regressions of Table 4, we observe that 38 characteristics have significant intercepts after controlling for the FF3 model. Out of those, 31 variables are selected in at least one of the in-sample linear models. However, only 6 characteristics are selected by all nine linear models, namely. **Net Working Capital Changes, Dividend-to-Price ratio, Earnings per Share, Industry Relative Reversals, Net Issuance and Return volatility**. This circumstance once more highlights the difference between a univariate analysis and a multivariate analysis that conditions on multiple firm characteristics enhanced with regularization.

All in all, I conclude that Elastic Net as well as LASSO does not introduce enough sparsity to sufficiently reduce features within individual stock forecasting models. What is more, controlling for micro-cap stocks within both thresholds increases the in-sample explanatory power measured by the Sharpe Ratio independent of the number of selected covariates.

Given the evidence that sparser models have greater explanatory power once mirco stocks are excluded, I account for non-linearity's in the cross-section of firm characteristics to further increase model accuracy as well as sparsity as represented in Table 6. Section 3 describes the spline approach in more detail and states the equivalence to local regressions. The main intuition behind spline interpolation is to split the cross-section into subsets and to model each subset individually. Therefore, extreme values in certain subset do not effect other areas of the cross-section and thus reduce biases. This methodology is the regression equivalent to multivariate independent portfolio sortings but allows to present results for more than three characteristics. Knots are the percentile cutoffs that are necessary to form portfolio sorts. Given nine knots for instance implies sorting into 10 portfolios. I employ 19 cutoffs or knots in the in-sample analysis which is equivalent to sorts into 20 portfolios and for robustness I increase the number of knots to 25. The full set of results can be found in the online appendix. The magnitude of knots depends on the sample size and since I omit the time index, each firm-time observation becomes an independent cross-sectional observation. Hence, pooling the entire sample period allows for detailed non-parametric estimation of the cross-section with a larger number of knots.

In the choice of magnitude of knots, I follow Freyberger et al. (2020).
Table 6 represents again the four main regularization models allowing for non-linear relationships in the cross-section. Despite the use of different model specifications, information sets selected show very little variation. Column 1 represents the SCAD model which selects only three variables: **52-Week High, Log of stock price** and **Industry Relative Reversals**. This finding is identical to the NYSE 10% restriction, which is therefore omitted form the table. The former two characteristics are selected across all models linear or non-linear. In 14 out of 15 cases, the variable 52-Week High is selected. However, the last column present the SCAD model for the most restricted sample which excludes the feature: Industry Relative Reversals. Although this model consists of only two variables its in-sample Sharpe ratio is only slightly lower. In comparison, the linear SCAD model of table 5 selects 65 characteristics with only slightly higher Sharpe ratio. Column (3) to (5) have identical results for NYSE 10% restrictions and are thus only presented in the online appendix. As in the linear case, Elastic Net selects far more variables without significantly increasing the explanatory power in form of in-sample Sharpe ratios. Another interesting fact is the difference between the non-linear LASSO model and its adaptive counterpart. The only variable that is excluded through the second and more comprehensive estimation procedure is the Earnings-to-Price ratio with an only slightly lower Sharpe ratio.
The linear in-sample model comparison yield that sufficiently sparser models explain almost as well as the larger models. This finding is even more pronounced for the non-linear model in-sample results. In summary, between 2 to 9 variables are enough the explain in-sample variation and are hence a tremendous reduction of features.
The literature suggests an upper bound of around 13 risk factors, which seems to be in line with the set of features found by the conditional models presented in this subsection. Nevertheless, in order to understand whether the excluded characteristics lead to overfitting as theoretically predicted by the curse of dimensionality, we need to consider out-of-sample analysis, which we discuss in the next section.

5.3 Out-of-Sample Comparison

While the previous subsection selected firm characteristics that are independently meaningful for the entire sample period, the out-of-Sample comparison should provide information concerning estimation uncertainty. The previous model fits might be subject to overfitting.

For this reason, I perform model selection for a certain period, as for instance from 1965 to 1990. I re-estimate the coefficients for the selected set of significant variables within the last 10 years of the sample period to obtain unpenalized parameter estimates. With these beta coefficients and new data starting from January 1991 for instance, I predict individual returns for the next month. Subsequently, I estimate the coefficients rolling forward one month and forecast using the updated values of the previously selected information set and therefore conduct expanding return prediction. I sort predicted returns in 10 portfolios and form value-weighted hedge portfolios going long the stocks in the highest sort and shorting the stocks of the lowest sort. The different out-of-sample periods, number of selected characteristics and regularization terms are presented for the linear case in Table 7 and for the non-linear model Table 8 presents the results. Both tables also report the first four moments of monthly percentage returns, annualized Sharpe ratios, turnover, and predictive slopes and R^2 for the hedge portfolios in Panel A, as well as for the long legs in Panel B and the short legs in Panel C. To calculate turnover I follow Koijen, Moskowitz, Pedersen, and Vrugt (2018) for both turnover measures, which is defined as $TO_t = \frac{1}{4}\sum_t^{N_t} |w_{i,t} - w_{i,t-1}|$ or as $TO_{R_t} = \frac{1}{4}\sum_t^{N_t} |w_{i,t} - w_{i,t-1}(1 + r_{i,t})|$, where $w_{i,t}$ is the value-weight of stock i at time t.

In the first to third column of Table 7, the model selection period starts in 1965 until 1982. Thereafter, given the selected variables, out-of-sample predictions are made for the next successive 120 months. The SCAD model selects 19 variables within this period, 17 less than the Elastic Net model and 5 more compared to the two step LASSO penalization. However, its out-of-sample Long-Short portfolio suggest a monthly return of 2.9% and a Sharpe ratio of 2.52, which is slightly higher than its competitors.

In the same period, I also estimate a pure LASSO model, which as in the in-sample period yields identical selected features as the Elastic Net and is thus omitted from the table. Furthermore, I omit the presentation of the Elastic Net for further estimation periods due to its overfitting tendency and its consequent poor out-of-sample performance.

Comparing the model fits for the remaining estimation periods reveals a decaying out-of-sample performance. Indeed, the Adaptive LASSO model of column (6) yields an out-of-sample Sharpe ratio similar to market portfolio of 0.5, if we assume an equity premium of 8% and a market volatility around 16%. Overall, the skewness is mostly positive for hedge and long positions, whereas it is negative for the short leg.

To further validate the quality of individual return predictions, I use average betas and R^2 as suggested by Lewellen (2014). I estimate these measures by regressing individual ex-post returns on individual predicted returns per month. In both cases, a value of one represents a perfect one-to-one relationship. R^2 ranges between 3 and 5% indicating that the SCAD penalty has slightly greater predictive power.

Decomposing hedge returns into the long and short returns reveals that most of its performance comes from the long leg. Indeed, beside for column (3), the SCAD model, returns are almost zero or even positive, leading to a smaller Long-Short portfolio return. On the other hand, volatility is always larger on the long leg.

All in all, the SCAD model seems to select the right features more often than its competing regularization terms. Despite the use of expanding prediction for the remaining 10 years, average percentage returns and Sharpe ratios of Panel A are surprisingly large and above the mean-variance efficient tangency portfolio of 0.5.

In the non-linear case, Table 8 presents the out-of-sample results. I chose identical data sample restrictions as for the linear model in order to compare both approaches. We can observe that once we allow for quadratic terms, the model selects far less characteristics. In the most recent period, both the Adaptive LASSO as well as the SCAD model have higher Sharpe Ratios with only 3, respectively 2 features: **Industry Relative Reversals, Log of stock price, (Lagged Momentum)** compared to 18 and 30 active characteristics. Moreover, out-of-sample betas and R^2 are larger throughout the non-parametric estimation given similar turnover. For the various estimation periods, the Group SCAD model selects 1, 4 and 2 characteristics, respectively 2, 3, 3 for the Adaptive

Group LASSO model. As a result, we can conclude that the predictive power of market anomalies as well as the selected subset of the first 20 years relative to the entire sample varies across time.

Consequently, I conduct a rolling estimation with a variable selection window of 20 years. Thereafter, I predict returns for each consecutive month, fixing the selected variable set for one year before I redo variable selection. Table 9 show the out-of-sample results for rolling variable selection. All are significantly larger than their time-invariant complement of table 8. Since the number of selected variables per column changes from estimation to estimation, I present the selected characteristics per year of the SCAD regularization term in figure 5. Some features seem to be very consistent over time, while others such as the Bid and Ask spread vary tremendously through time.

To sum up, differences across models are substantial. The number of variables selected range from 1 to 30 out of 90 characteristics. Nonetheless, a stable intersection between the characteristics selected across models exists. SCAD regularization terms are more stable in terms of parameter choices than their counterparts. It also selects the most sparest model with economically significant out-of-sample Sharpe ratios. Adaptive group LASSO seems to be a consistent alternative, whereas Elastic Net tends to select many variables and the out-of-sample results suggest that it overfits the model. Quadratic terms seems to matter substantially performance wise as well as in terms of explanatory power.

6 Conclusion

Given the multidimensional challenge proposed by Cochrane (2011), I suggest an alternative econometric technique that manages the curse of dimensionality and dissect the independent information content that firm characteristics provide for expected returns. In particular, I recommend a single-stage regression enhanced by a smoothly clipped absolute deviations (SCAD) penalty on group level. This regularization term does not penalize critical firm characteristics and therefore allows them survive the dimension reduction.

Starting with a novel database of 90 market anomalies, I select between 2 to 8 characteristics in an in-sample comparison which considers non-linear relationships between expected returns and rank-normalized firm characteristics. Three are consistent across various model specifications such as penalization terms and number of steps. The determined information subset discovered is time-varying. Out of 26 different predictors only 4 are consistent through time.

After selecting a model, I suggest a supervised learning approach to identify the optimal number of knots non-parametricly. The multivariate additive model recovers the partial contribution of the previously selected variables with respect to the conditional expected returns. In particular, I employ natural splines that are extensions of smoothing splines in the sense that they have additional boundary constraints. These constraints force the function $g(.)$ to be linear at the boundaries and thus resistant to extreme values which are common in the tails.

References

Adrian, T., Etula, E., & Muir, T. (2014). Financial Intermediaries and the Cross-Section of Asset Returns. *Journal of Finance*, *69*(6), 2557–2596. doi: 10.1111/jofi.12189

Asness, C., Frazzini, A., Israel, R., Moskowitz, T. J., & Pedersen, L. H. (2018). Size matters, if you control your junk. *Journal of Financial Economics*, *129*(3), 479–509.

Bansal, R., & Yaron, A. (2004). Risks for the long run: A potential resolution of asset pricing puzzles. *The journal of Finance*, *59*(4), 1481–1509.

Barras, L., Scaillet, O., & Wermers, R. (2010). False discoveries in mutual fund performance: Measuring luck in estimated alphas. *The journal of finance*, *65*(1), 179–216.

Black, F., Jensen, M. C., & Scholes, M. (1972). The capital asset pricing model: Some empirical tests. *Studies in the theory of capital markets*, *81*(3), 79–121.

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

Campbell, J. Y., & Cochrane, J. H. (1999). By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of political Economy*, *107*(2), 205–251.

Cattaneo, M. D., Crump, R. K., Farrell, M. H., & Schaumburg, E. (2018). Characteristic-sorted portfolios: Estimation and inference. *Review of Economics and Statistics*, 1–47.

Chen, A. Y., & Velikov, M. (2019). Accounting for the anomaly zoo: A trading cost perspective. *Finance Down Under 2019 Building on the Best from the Cellars of Finance*.

Chen, L., Pelger, M., & Zhu, J. (2019). Deep learning in asset pricing. *Available at SSRN 3350138*.

Chinco, A., Clark-Joseph, A. D., & Ye, M. (2019). Sparse signals in the cross-section of returns. *The Journal of Finance*, *74*(1), 449–492.

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

Daniel, K., Jagannathan, R., & Kim, S. (2012). Tail risk in momentum strategy returns. *working paper*.

DeMiguel, V., Martin-Utrera, A., Nogales, F. J., & Uppal, R. (2019). A transaction-cost perspective on the multitude of firm characteristics. *Available at SSRN 2912819*.

Dobrynskaya, V. (2017). Upside and downside risks in momentum returns. *Available at SSRN 2399359*.

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

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

Fan, J., & Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American statistical Association*, *96*(456), 1348–1360.

Freyberger, J., Neuhierl, A., & Weber, M. (2020). Dissecting characteristics nonparametrically. *Review of Financial Studies*.

Green, J., Hand, J. R. M., & Zhang, X. F. (2017). The Characteristics that Provide Independent Information about Average US Monthly Stock Returns. *Review of Financial Studies*, *30*(12), 4389–4436. doi: 10.1093/rfs/hhx019

Gu, S., Kelly, B., & Xiu, D. (2018). Empirical asset pricing via machine learning. *working paper*.

Haddad, V., Kozak, S., & Santosh, S. (2018). The Economics of Factor Timing. *WP*.

Hansen, L. P., & Jagannathan, R. (1991). Implications of security market data for models of dynamic economies. *Journal of political economy*, *99*(2), 225–262.

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

Hou, K., Mo, H., Xue, C., & Zhang, L. (2019). Which factors? *Review of Finance*, *23*(1), 1–35.

Huang, J., Horowitz, J. L., & Wei, F. (2010). Variable selection in nonparametric additive models. *Annals of statistics*, *38*(4), 2282.

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

Koijen, R. S., Moskowitz, T. J., Pedersen, L. H., & Vrugt, E. B. (2018). Carry. *Journal of Financial Economics*, *127*(2), 197–225.

Kozak, S., Nagel, S., & Santosh, S. (2019). Shrinking the cross-section. *Journal of Financial Economics*.

Lewellen, J. (2014). The cross section of expected stock returns. *Forthcoming in Critical Finance Review*.

Lintner, J. (1965). Security prices, risk, and maximal gains from diversification. *The journal of finance*, *20*(4), 587–615.

Markowitz, H. (1952). Portfolio selection. *The journal of finance*, *7*(1), 77–91.

McLean, R. D., & Pontiff, J. (2016). Does academic research destroy stock return predictability? *The Journal of Finance*, *71*(1), 5–32.

Mehra, R., & Prescott, E. C. (1985). The equity premium: A puzzle. *Journal of monetary Economics*, *15*(2), 145–161.

Merton, R. C. (1969). Lifetime portfolio selection under uncertainty: The continuous-time case. *The review of Economics and Statistics*, 247–257.

Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica: Journal of the econometric society*, 768–783.

Novy-Marx, R., & Velikov, M. (2016). A Taxonomy of Anomalies and Their Trading Costs. *Review of Financial Studies*, *29*(1), 104–147. doi: 10.1093/rfs/hhv063

Penasse, J. (2017). Understanding anomaly decay. *WP*.

Reinsch, C. H. (1967). Smoothing by spline functions. *Numerische mathematik*, *10*(3), 177–183.

Santos, T., & Veronesi, P. (2004). Conditional betas. *working paper*.

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

Shumway, T., & Warther, V. A. (1999). The delisting bias in crsp's nasdaq data and its implications for the size effect. *The Journal of Finance*, *54*(6), 2361–2379.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the*

Royal Statistical Society: Series B (Methodological), *58*(1), 267–288.
Varrette, S., Bouvry, P., Cartiaux, H., & Georgatos, F. (2014, July). Management of an academic hpc cluster: The ul experience. In *Proc. of the 2014 intl. conf. on high performance computing & simulation (hpcs 2014)* (pp. 959–967). Bologna, Italy: IEEE.
Vuolteenaho, J. Y. . C., & Tuomo. (2004). Bad Beta , Good Beta. *The American Economic Review*, *94*(5), 1249–1275.
Yuan, M., & Lin, Y. (2006). Model selection and estimation in regression with grouped variables. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *68*(1), 49–67.
Zou, H. (2006). The adaptive lasso and its oracle properties. *Journal of the American statistical association*, *101*(476), 1418–1429.
Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the royal statistical society: series B (statistical methodology)*, *67*(2), 301–320.

Appendix

A Tables

Table 1: **Summary Statistics**
This table reports summary statistics of characteristics and characteristic sorted portfolio returns. Panel A, presents average value-weighted monthly hedge returns in percentage. Stocks are sorted into extreme deciles based on their rank-normalized characteristic at the end of each month. Panel B, presents the mean and the standard deviation of the underlying characteristics.

Panel A: Long-Short Portfolio Returns

Anomalies	Mean	Median	Std. Dev.	Skewness	Kurtosis	Min	Max	Obs.
Full Sample	0.33	0.30	4.95	0.03	3.87	-24.43	24.42	57,978
Expansion	0.32	0.29	4.71	0.05	4.22	-23.20	23.47	50,539
Depression	0.41	0.45	6.30	-0.02	1.29	-17.55	18.12	7,439
Pre 2003	0.42	0.42	5.11	0.01	3.86	-23.56	23.66	40,790
Post 2003	0.12	0.06	4.44	0.06	2.14	-15.87	15.93	17,188
CRSP-Rf	0.70	1.02	5.24	-0.56	2.17	-27.54	20.48	1,417,151

Panel B: Descriptive Statistics of Characteristics

Anomalies	Mean	StDev	Anomalies	Mean	StDev	Anomalies	Mean	StDev
a2me	2.71	8.06	eps	1.16	2.53	pchgm_pchsale	-0.20	28.17
agr	0.14	0.51	ga_eps	-0.04	2.54	pchquick	0.11	2.29
at	4.14	40.82	gma	0.39	0.34	pchsale_pchrect	-0.09	32.09
bab	1.01	0.36	herf	0.08	0.07	pcm	-0.52	98.08
baspread	0.04	0.05	hire	0.11	1.63	pm	-0.93	101.22
bm	0.80	0.87	ill	0.00	0.00	pm_ia	2.55	101.22
bm_ia	-2.84	8.50	indmom6m	0.00	0.02	prc52	0.81	0.20
c2d	0.12	0.99	indmom6m_ia	0.10	0.17	prof	1.06	18.74
cashdebt	0.13	0.85	indrev1m	-0.00	0.14	ps	25.00	437.62
cashpr	-3.29	2152.32	invest	0.08	0.22	q	0.70	1.60
chaTO	-0.01	0.30	ipm	-0.97	99.62	quick	3.77	191.80
chatoia	-0.04	2.47	ldp	0.43	0.96	r62	0.06	0.40
chcsho	0.10	0.39	lev	1.75	5.18	retvol	0.03	0.03
chempia	-0.16	1.95	lgr	0.24	1.97	rev1m	-0.01	0.16
chinv	0.01	0.06	lprc	2.33	1.27	roa2	0.02	0.19
chlogso	0.03	0.14	maxret	0.07	0.08	roe	0.06	2.96
chmom	-0.00	0.62	mom11m	0.14	0.66	roic	0.01	1.34
chnwc	0.01	0.10	mom12m	0.16	0.70	salecash	0.12	2.24
chpm	-0.11	94.49	mom18m	-0.08	0.43	salerec	14.92	136.20
chpmia	-0.04	95.83	mom36m	0.35	1.22	sat	1.14	0.91
cTO	1.28	1.13	mom6m	0.07	0.44	sat_ia	-0.06	1.85
currat	5.43	461.58	mom6m_l6	0.08	0.44	sgr	0.32	25.94
daily_Beta	0.92	0.85	mve	5.40	2.21	sp	2.27	4.27
debt2p	0.69	2.12	mve_firm	2.81	15.60	std_dolvol	0.80	0.40
depr	0.31	3.96	mve_firm_ia	0.73	15.24	std_dolvol36m	1.41	8.19
dolvol	11.65	2.99	mve_ia	0.69	14.10	std_turn	4.06	14.61
dy	0.02	0.05	mve_jun_log	5.42	2.18	tan	0.53	0.14
ebitrev	-0.93	101.22	ni	0.05	0.93	tang	0.53	0.14
egr	0.24	8.00	operprof	0.32	17.56	turn	1.08	1.76
ep	-0.01	0.42	pchcurrat	0.08	2.07	zerotrade	1.00	2.85

Table 2: **Defining Market Anomalies**

This table provides information about the name, authors and sample period of the original study of the 85 anomalies constructed in this paper.

Id	Abbrev	Anomaly	Paper	Period
1	a2me	A2ME	Bhandari	1946-1981
2	age.m	Firm Age	Barry, Brown	1931-1982
3	agr	Asset growth	Cooper et al.	1963-2003
4	at	Total Assets	Gandhi, Lustig	1970-2013
5	bab	Beta	Frazzini, Pedersen	1984-2012
6	baspread	Bid-ask spread	Amihud, Mendelson	1961-1980
7	bm	Book-to-market	Rosenberg et al.	1963-1990
8	bm.ia	Ind.-adj. book-to-market	Asness et al.	1963-1998
9	c2d	C2D	Casey, Bartczak	1971-1982
10	cashdebt	Cash Flow to debt	Ou, Penman	1965-1983
11	cashpr	Cash productivity	Chandrashekar, Rao	1962-2003
12	chato	Asset Turnover	Soliman	1984-2002
13	chatoia	Ind.-adj. Chg. in asset turnover	Soliman	1984-2002
14	chcsho	Chg. in shares outstanding	Pontiff, Woodgate	1970-2003
15	chempia	Ind.-adj. Chg. in employees	Asness et al.	1963-1998
16	chinv	Chg. in inventory	Thomas, Zhang	1970-1997
17	chlogso	ΔSO	Fama, French	1963-2005
18	chmom	Chg. in 6-month momentum	Gettleman, Marks	1926-2003
19	chnwc	Chg. in Net Working Capital	Soliman	1984-2002
20	chpm	Chg. in Profit	Soliman	1984-2002
21	chpmia	Ind.-adj. Chg. in profit margin	Soliman	1984-2002
22	cto	CTO	Baker	1983-1993
23	currat	Current ratio	Ou, Penman	1965-1983
24	daily.Beta	Beta daily	Lewellen, Nagel	1964-2001
25	debt2p	Debt2P	Ramaswamy	1936-1977
26	depr	Depreciation divided PP&E	Holthausen, Larcker	1978-1988
27	dolvol	Dollar trading volume	Chordia et al.	1966-1995
28	dy	Dividend to price	Litzenberger et al.	1931-1951
29	ebitrev	Profit Margin	Soliman	1984-2002
30	egr	Growth in common equity	Richardson et al.	1962-2001
31	ep	Earnings to price ratio	Basu	1957-1971
32	eps	Earnings per share	Basu	1957-1971
33	ga.eps	Earnings Consistency	Alwathainani	1971-2002
34	gma	Gross profitability	Novy-Marx	1963-2010
35	herf	Industry sales concentration	Hou, Robinson	1963-2001
36	hire	Employee growth rate	Bazdresch, Belo, Lin	1965-2010
37	ill	Illiquidity	Amihud	1964-1997
89	indmom6m	Industry Momentum	Moskowitz, Grinblatt	1963-1995
38	indmom6m.ia	Industry Momentum	Grinblatt et al.	1963-1995
39	indrev1m	Industry Relative Reversals	Da, Liu, Schaumurg	1982-2009
40	invest	CAPEX and inventory	Chen, Zhang	1972-2006
41	ipm	IPM	NA	NA-NA
42	ldp	LDP	Litzenberger et al.	1936-1977

Id	Abbrev	Anomaly	Paper	Period
43	lev	Leverage	Bhandari	1946-1981
44	lgr	Growth in long-term debt	Richardson et al.	1962-2001
45	lprc	Price	Blume, Husic	1932-1971
46	maxret	Maximum daily return	Bali et al.	1962-2005
47	mom11m	r12-2	Fama, French	1963-1993
48	mom12m	Cash Flow volatility	Jegadeesh	1964-1989
49	mom18m	Momentum-Reversal	Jegadeesh, Titman	1964-1989
50	mom36m	36-month momentum	Jegadeesh, Titman	1926-1982
51	mom6m	Momentum-Volume	Jegadeesh, Titman	1964-1989
52	mom6m.l6	Lagged Momentum	Novy-Marx	1926-2010
53	mve	Size	Banz	1926-1975
54	mve.firm	LME	Fama, French	1963-1990
55	mve.firm.ia	LME Ind.-adj.	Asness et al.	1963-1998
56	mve.ia	Ind.-adj. size	Asness et al.	1963-1998
57	mve.jun.log	Size	Banz	1926-1975
58	ni	Net Issuance	Fama, French	1973-2002
59	operprof	Operating profitability	Fama, French	1977-2003
60	pchcurrat	Chg. in current ratio	Ou, Penman	1965-1983
61	pchgm.pchsale	Chg. in gross margin - Chg. in sales	Abarbanell, Bushee	1974-1988
62	pchquick	Chg. in quick ratio	Ou, Penman	1965-1983
63	pchsale.pchrect	Chg. in sales - Chg. in A/R	Abarbanell, Bushee	1974-1988
64	pcm	PCM	Gorodnichenko, Weber	1994-2009
65	pm	PM	Soliman	1984-2002
66	pm.ia	PM Ind.-adj.	Soliman	1984-2002
67	prc52	52-Week High	George, Hwang	1963-2001
68	prof	Prof	Ball et al.	1963-2013
69	q	Tobin's Q	Wernerfelt et al.	1960-1977
70	quick	Quick ratio	Ou, Penman	1965-1983
71	r62	r6-2	Jegadesh, Titman	1964-1989
72	retvol	Return volatility	Ang et al.	1986-2000
73	rev1m	Short-Term Reversal	Jegadeesh	1934-1987
74	roa2	ROA	Balakrishnan et al.	1976-2005
75	roe	Return-on-Equity	Haugen, Baker	1979-1993
76	roic	Return on invested capital	Brown, Rowe	1970-2005
77	salecash	Sales to cash	Ou, Penman	1965-1983
78	salerec	Sales to receivables	Ou, Penman	1965-1983
79	sat	SAT	Soliman	1984-2002
80	sat.ia	SAT Ind.-adj.	Soliman	1984-2002
81	sgr	Sales growth	Lakonishok et al.	1963-1990
82	sp	Sales/Price	Barbee et al	1979-1991
83	std.dolvol	Volatility of liquidity	Chordia et al.	1966-1995
84	std.dolvol36m	Volume Variance	Chordia et al.	1966-1995
85	std.turn	Volatility of liquidity	Chordia et al.	1966-1995
86	tan	Tan	Hahn, Lee	1973-2001
87	tang	Debt capacity/firm tangibility	Almeida, Campello	1973-2001
88	turn	Share turnover	Datar et al.	1962-1991
90	zerotrade	Zero trading days	Liu	1960-2003

Table 3: **Portfolio Returns per Market Anomaly**

This table reports monthly average returns for 10 value-weighted portfolios per market anomaly. Parentheses indicate standard errors and significant average returns at a confidence interval of 95% are printed in bold type. The sample period spans from January 1965 to December 2018.

Anomaly	P1	s.e.	P2	s.e.	P3	s.e.	P4	s.e.	P5	s.e.	P6	s.e.	P7	s.e.	P8	s.e.	P9	s.e.	P10	s.e.	P10-P1	s.e.
a2me	**0.86**	(0.21)	**0.88**	(0.18)	**0.87**	(0.18)	**0.98**	(0.17)	**0.97**	(0.18)	**1.01**	(0.17)	**1**	(0.18)	**1.12**	(0.2)	**1.04**	(0.22)	**1.13**	(0.26)	0.26	(0.21)
age.m	**1.02**	(0.3)	**0.94**	(0.27)	**1.06**	(0.26)	**1.04**	(0.26)	**1.21**	(0.26)	**1.12**	(0.25)	**1.02**	(0.24)	**1.02**	(0.21)	**1.05**	(0.21)	**0.99**	(0.18)	-0.03	(0.21)
agr	**0.73**	(0.24)	**0.9**	(0.22)	**0.89**	(0.19)	**0.93**	(0.17)	**0.92**	(0.17)	**0.97**	(0.16)	**0.95**	(0.16)	**1.08**	(0.18)	**1.11**	(0.19)	**1.12**	(0.23)	**0.38**	(0.15)
at	**0.59**	(0.32)	**0.89**	(0.3)	**1.09**	(0.28)	**1.1**	(0.27)	**1.1**	(0.25)	**1.13**	(0.24)	**1.03**	(0.22)	**1.08**	(0.21)	**0.92**	(0.19)	**0.86**	(0.16)	0.27	(0.26)
bab	**0.86**	(0.15)	**0.98**	(0.13)	**1.03**	(0.14)	**0.96**	(0.15)	**0.94**	(0.16)	**0.91**	(0.18)	**0.86**	(0.2)	**0.89**	(0.22)	**0.8**	(0.26)	**0.61**	(0.33)	-0.25	(0.3)
baspread	0.48	(0.43)	0.41	(0.36)	**0.71**	(0.31)	**0.83**	(0.28)	**1.14**	(0.25)	**1.05**	(0.21)	**1.04**	(0.19)	**0.9**	(0.17)	**0.96**	(0.15)	**0.83**	(0.14)	0.34	(0.4)
bm	**0.9**	(0.21)	**0.87**	(0.19)	**0.91**	(0.18)	**0.94**	(0.18)	**0.89**	(0.17)	**0.98**	(0.18)	**0.91**	(0.18)	**1.19**	(0.2)	**1.03**	(0.23)	**1.24**	(0.26)	0.34	(0.22)
bm.ia	**0.9**	(0.19)	**0.79**	(0.21)	**0.85**	(0.19)	**0.84**	(0.19)	**1.05**	(0.19)	**0.97**	(0.19)	**0.99**	(0.19)	**1.03**	(0.18)	**1**	(0.19)	**1.24**	(0.22)	**0.33**	(0.15)
c2d	**0.62**	(0.34)	**0.85**	(0.26)	**0.8**	(0.23)	**0.95**	(0.2)	**0.92**	(0.17)	**0.94**	(0.17)	**0.91**	(0.18)	**0.99**	(0.17)	**0.91**	(0.18)	**0.91**	(0.21)	0.29	(0.23)
cashdebt	0.55	(0.34)	**0.83**	(0.25)	**0.85**	(0.23)	**0.94**	(0.2)	**0.83**	(0.17)	**1**	(0.17)	**0.91**	(0.18)	**0.94**	(0.17)	**0.93**	(0.18)	**0.93**	(0.22)	0.37	(0.23)
cashpr	**0.75**	(0.19)	**0.93**	(0.18)	**0.92**	(0.2)	**1.06**	(0.2)	**1.23**	(0.2)	**1.07**	(0.19)	**1.08**	(0.19)	**0.99**	(0.2)	**1.01**	(0.19)	**1.1**	(0.18)	**0.35**	(0.13)
chato	**0.83**	(0.23)	**0.85**	(0.21)	**0.85**	(0.19)	**0.86**	(0.17)	**0.89**	(0.17)	**0.92**	(0.18)	**0.99**	(0.18)	**1.08**	(0.18)	**1**	(0.2)	**1**	(0.22)	0.17	(0.15)
chatoia	**0.9**	(0.23)	**0.79**	(0.2)	**0.84**	(0.18)	**0.83**	(0.18)	**0.97**	(0.17)	**0.93**	(0.18)	**0.99**	(0.17)	**1.01**	(0.18)	**0.97**	(0.2)	**1.06**	(0.22)	0.16	(0.13)
chcsho	**0.6**	(0.2)	**0.82**	(0.2)	**0.83**	(0.2)	**0.96**	(0.2)	**1.08**	(0.19)	**0.89**	(0.18)	**0.97**	(0.17)	**0.98**	(0.17)	**1.07**	(0.17)	**1.24**	(0.17)	**0.64**	(0.11)
chempia	**0.79**	(0.2)	**0.89**	(0.19)	**0.92**	(0.18)	**0.93**	(0.18)	**1**	(0.17)	**0.93**	(0.18)	**0.91**	(0.17)	**0.86**	(0.18)	**1.07**	(0.21)	**0.83**	(0.23)	0.04	(0.13)
chinv	**0.83**	(0.24)	**0.87**	(0.21)	**0.79**	(0.2)	**0.95**	(0.18)	**0.93**	(0.17)	**0.99**	(0.17)	**0.88**	(0.18)	**0.95**	(0.17)	**1.16**	(0.19)	**1.26**	(0.23)	**0.43**	(0.14)
chlogso	**0.48**	(0.21)	**0.72**	(0.2)	**0.83**	(0.21)	**1.04**	(0.2)	**0.99**	(0.19)	**0.91**	(0.18)	**0.9**	(0.17)	**0.98**	(0.17)	**1.05**	(0.17)	**1.29**	(0.18)	**0.81**	(0.12)
chmom	**0.87**	(0.24)	**0.83**	(0.2)	**0.86**	(0.18)	**0.76**	(0.18)	**0.96**	(0.18)	**0.97**	(0.17)	**1**	(0.18)	**1.09**	(0.2)	**1.11**	(0.23)	**1.09**	(0.28)	0.22	(0.21)
chnwc	**0.59**	(0.25)	**0.8**	(0.21)	**0.86**	(0.2)	**0.9**	(0.18)	**0.9**	(0.17)	**0.99**	(0.16)	**0.94**	(0.16)	**0.94**	(0.18)	**1.15**	(0.21)	**1.09**	(0.23)	**0.5**	(0.13)
chpm	**0.83**	(0.25)	**0.93**	(0.2)	**0.88**	(0.18)	**0.91**	(0.18)	**0.98**	(0.17)	**1.05**	(0.17)	**0.84**	(0.17)	**0.92**	(0.18)	**0.85**	(0.19)	**0.73**	(0.24)	-0.11	(0.14)
chpmia	**0.91**	(0.2)	**0.88**	(0.2)	**0.89**	(0.18)	**0.75**	(0.19)	**0.99**	(0.17)	**0.91**	(0.18)	**0.85**	(0.19)	**1.07**	(0.2)	**1.08**	(0.21)	**1.03**	(0.21)	0.11	(0.17)
cto	**0.85**	(0.21)	**0.79**	(0.2)	**0.81**	(0.18)	**0.92**	(0.17)	**0.92**	(0.19)	**1.05**	(0.18)	**0.84**	(0.19)	**0.97**	(0.2)	**1.04**	(0.2)	**1.07**	(0.21)	0.22	(0.17)
currat	**0.77**	(0.22)	**1.01**	(0.22)	**0.9**	(0.22)	**1.07**	(0.21)	**0.86**	(0.2)	**0.86**	(0.19)	**1**	(0.18)	**0.95**	(0.17)	**0.93**	(0.17)	**0.85**	(0.15)	0.08	(0.15)
daily.Beta	**0.67**	(0.17)	**0.76**	(0.14)	**0.85**	(0.15)	**0.88**	(0.15)	**0.97**	(0.16)	**1**	(0.17)	**0.99**	(0.19)	**1.01**	(0.22)	**0.99**	(0.26)	**0.9**	(0.32)	0.23	(0.28)
debt2p	**0.98**	(0.23)	**0.88**	(0.21)	**0.92**	(0.18)	**0.98**	(0.18)	**1.02**	(0.18)	**0.95**	(0.17)	**0.92**	(0.17)	**0.92**	(0.17)	**0.94**	(0.2)	**1.03**	(0.25)	0.05	(0.21)
depr	**0.8**	(0.15)	**0.84**	(0.16)	**0.88**	(0.17)	**0.95**	(0.18)	**0.92**	(0.18)	**0.93**	(0.19)	**1.02**	(0.21)	**1.01**	(0.22)	**1**	(0.23)	**0.91**	(0.24)	0.11	(0.18)
dolvol	**1.17**	(0.19)	**1.23**	(0.19)	**1.21**	(0.2)	**1.13**	(0.2)	**1.11**	(0.2)	**1.17**	(0.2)	**1.08**	(0.19)	**1.07**	(0.19)	**1.02**	(0.18)	**0.85**	(0.17)	**-0.32**	(0.17)
dy	**0.9**	(0.18)	**1.01**	(0.16)	**0.99**	(0.17)	**1.04**	(0.17)	**0.93**	(0.17)	**0.9**	(0.18)	**0.89**	(0.19)	**0.75**	(0.2)	**0.85**	(0.23)	**1.1**	(0.27)	0.2	(0.23)
ebitrev	0.53	(0.33)	**1.08**	(0.29)	**0.98**	(0.23)	**1.07**	(0.2)	**0.88**	(0.19)	**0.98**	(0.19)	**0.94**	(0.18)	**0.88**	(0.17)	**0.89**	(0.17)	**0.86**	(0.17)	0.33	(0.25)
egr	**0.72**	(0.24)	**0.83**	(0.21)	**0.95**	(0.19)	**0.96**	(0.17)	**0.9**	(0.17)	**0.92**	(0.16)	**0.94**	(0.17)	**1.05**	(0.17)	**1.07**	(0.19)	**1.11**	(0.22)	**0.39**	(0.15)
ep	**0.89**	(0.32)	**0.63**	(0.26)	**0.79**	(0.23)	**0.86**	(0.2)	**0.91**	(0.18)	**0.99**	(0.17)	**1**	(0.17)	**1.08**	(0.17)	**1.14**	(0.18)	**1.31**	(0.2)	**0.42**	(0.24)
eps	**0.97**	(0.29)	**0.84**	(0.27)	**0.89**	(0.26)	**0.92**	(0.24)	**0.79**	(0.23)	**0.88**	(0.21)	**0.83**	(0.19)	**0.92**	(0.18)	**0.89**	(0.17)	**0.96**	(0.16)	0	(0.21)
ga.eps	**0.84**	(0.24)	**0.82**	(0.2)	**0.76**	(0.2)	**0.8**	(0.18)	**0.99**	(0.16)	**0.95**	(0.16)	**0.94**	(0.17)	**0.97**	(0.19)	**0.91**	(0.21)	**1**	(0.22)	0.16	(0.14)
gma	**0.79**	(0.23)	**0.91**	(0.21)	**0.88**	(0.17)	**0.87**	(0.18)	**0.93**	(0.19)	**0.92**	(0.18)	**0.84**	(0.19)	**0.95**	(0.19)	**0.98**	(0.18)	**1.08**	(0.21)	**0.29**	(0.17)
herf	**0.82**	(0.16)	**1.01**	(0.23)	**0.94**	(0.21)	**0.85**	(0.22)	**0.97**	(0.21)	**0.91**	(0.2)	**1.02**	(0.18)	**0.94**	(0.21)	**0.9**	(0.22)	**0.92**	(0.17)	0.1	(0.14)
hire	**0.87**	(0.23)	**0.92**	(0.22)	**0.96**	(0.2)	**0.92**	(0.18)	**0.87**	(0.17)	**0.89**	(0.17)	**1**	(0.17)	**0.97**	(0.17)	**1.03**	(0.19)	**1.05**	(0.21)	0.18	(0.13)
ill	**0.85**	(0.17)	**1.01**	(0.19)	**1.04**	(0.2)	**1.1**	(0.21)	**1.11**	(0.21)	**1.18**	(0.22)	**1.27**	(0.22)	**1.3**	(0.22)	**1.15**	(0.23)	**1.24**	(0.26)	**0.39**	(0.23)
indmom6m	**0.75**	(0.19)	**0.85**	(0.2)	**0.89**	(0.19)	**1.01**	(0.19)	**1.09**	(0.19)	**1.14**	(0.19)	**1.19**	(0.18)	**1.04**	(0.18)	**0.96**	(0.17)	**0.84**	(0.17)	0.09	(0.13)
indmom6m.ia	**0.88**	(0.23)	**0.86**	(0.21)	**0.87**	(0.2)	**1.04**	(0.2)	**0.94**	(0.2)	**0.94**	(0.2)	**0.88**	(0.2)	**0.96**	(0.2)	**1.02**	(0.2)	**1.07**	(0.21)	0.19	(0.21)
indrev1m	**0.54**	(0.23)	**0.62**	(0.19)	**0.65**	(0.17)	**0.85**	(0.17)	**0.93**	(0.17)	**1.11**	(0.17)	**1.12**	(0.18)	**1.12**	(0.2)	**1.25**	(0.23)	**1.3**	(0.29)	**0.76**	(0.21)

Anomaly	P1	s.e.	P2	s.e.	P3	s.e.	P4	s.e.	P5	s.e.	P6	s.e.	P7	s.e.	P8	s.e.	P9	s.e.	P10	s.e.	P10-P1	s.e.
invest	**0.75**	(0.23)	**0.81**	(0.21)	**0.92**	(0.19)	**0.9**	(0.18)	**0.89**	(0.16)	**0.93**	(0.17)	**0.98**	(0.18)	**1.07**	(0.2)	**1.03**	(0.2)	**1.26**	(0.22)	**0.51**	(0.12)
ipm	**0.67**	(0.34)	**1.04**	(0.25)	**0.92**	(0.24)	**1.09**	(0.21)	**0.94**	(0.19)	**0.93**	(0.19)	**0.95**	(0.18)	**0.93**	(0.17)	**0.83**	(0.17)	**0.91**	(0.17)	0.24	(0.25)
ldp	**1.01**	(0.27)	**0.93**	(0.27)	**1.02**	(0.24)	**0.89**	(0.22)	**0.89**	(0.21)	**0.86**	(0.2)	**0.83**	(0.19)	**1**	(0.17)	**0.91**	(0.16)	**0.88**	(0.15)	-0.13	(0.19)
lev	**0.81**	(0.23)	**0.98**	(0.19)	**0.85**	(0.18)	**0.91**	(0.17)	**0.97**	(0.17)	**0.96**	(0.17)	**0.99**	(0.17)	**1.02**	(0.18)	**1.02**	(0.21)	**1.05**	(0.25)	0.24	(0.21)
lgr	**0.73**	(0.23)	**0.97**	(0.21)	**0.88**	(0.19)	**0.9**	(0.17)	**0.97**	(0.17)	**0.95**	(0.16)	**0.97**	(0.17)	**0.98**	(0.17)	**0.95**	(0.2)	**0.9**	(0.22)	0.18	(0.11)
lprc	**0.46**	(0.17)	**0.86**	(0.17)	**0.99**	(0.18)	**1.18**	(0.18)	**1.31**	(0.19)	**1.36**	(0.21)	**1.52**	(0.23)	**1.56**	(0.24)	**1.94**	(0.26)	**2.27**	(0.31)	**1.81**	(0.24)
maxret	0.41	(0.32)	**0.77**	(0.3)	**0.86**	(0.27)	**0.95**	(0.24)	**1.03**	(0.22)	**1.01**	(0.21)	**0.93**	(0.19)	**0.94**	(0.17)	**0.94**	(0.15)	**0.94**	(0.14)	**0.53**	(0.28)
mom11m	0.11	(0.36)	0.44	(0.27)	**0.64**	(0.23)	**0.85**	(0.2)	**0.83**	(0.18)	**0.84**	(0.18)	**0.94**	(0.17)	**1.09**	(0.18)	**1.19**	(0.2)	**1.56**	(0.26)	**1.45**	(0.32)
mom12m	0.11	(0.36)	**0.51**	(0.28)	**0.69**	(0.24)	**0.81**	(0.2)	**0.85**	(0.18)	**0.87**	(0.18)	**0.92**	(0.17)	**1.08**	(0.18)	**1.13**	(0.2)	**1.45**	(0.26)	**1.34**	(0.32)
mom18m	**0.71**	(0.25)	**0.91**	(0.21)	**0.94**	(0.18)	**0.9**	(0.17)	**0.99**	(0.17)	**1.05**	(0.18)	**1.05**	(0.18)	**0.98**	(0.2)	**1.08**	(0.23)	**1.3**	(0.28)	**0.58**	(0.22)
mom36m	**0.84**	(0.25)	**0.95**	(0.2)	**0.99**	(0.18)	**1.02**	(0.17)	**0.96**	(0.17)	**0.99**	(0.18)	**1.09**	(0.19)	**1.16**	(0.21)	**1.18**	(0.25)	**1.11**	(0.31)	0.27	(0.25)
mom6m	0.31	(0.35)	**0.81**	(0.26)	**0.87**	(0.22)	**0.93**	(0.2)	**0.9**	(0.18)	**0.97**	(0.17)	**0.91**	(0.17)	**0.95**	(0.18)	**0.95**	(0.2)	**1.38**	(0.25)	**1.08**	(0.31)
mom6m.l6	0.23	(0.29)	0.32	(0.23)	**0.65**	(0.2)	**0.8**	(0.18)	**0.81**	(0.18)	**0.87**	(0.17)	**0.96**	(0.17)	**1.01**	(0.19)	**1.3**	(0.21)	**1.56**	(0.27)	**1.33**	(0.25)
mve	**0.86**	(0.17)	**1.06**	(0.19)	**1.13**	(0.21)	**1.18**	(0.23)	**1.19**	(0.24)	**1.17**	(0.25)	**1.17**	(0.25)	**1.16**	(0.25)	**1.22**	(0.27)	**2.02**	(0.31)	**1.16**	(0.27)
mve.firm	**0.86**	(0.17)	**1.06**	(0.19)	**1.12**	(0.21)	**1.18**	(0.23)	**1.19**	(0.24)	**1.16**	(0.25)	**1.18**	(0.25)	**1.18**	(0.25)	**1.2**	(0.27)	**2.01**	(0.31)	**1.15**	(0.27)
mve.firm.ia	**0.84**	(0.17)	**1.04**	(0.2)	**1.03**	(0.21)	**1.15**	(0.21)	**1.16**	(0.22)	**1.13**	(0.21)	**1.19**	(0.2)	**1.2**	(0.2)	**1.24**	(0.22)	**1.4**	(0.24)	**0.56**	(0.14)
mve.ia	**0.84**	(0.17)	**1.04**	(0.2)	**1.15**	(0.21)	**1.11**	(0.22)	**1.2**	(0.22)	**1.11**	(0.21)	**1.11**	(0.19)	**0.99**	(0.19)	**1.05**	(0.21)	**1.2**	(0.23)	**0.35**	(0.14)
mve.jun.log	**0.85**	(0.17)	**1.07**	(0.19)	**1.09**	(0.21)	**1.14**	(0.22)	**1.15**	(0.23)	**1.23**	(0.25)	**1.24**	(0.25)	**1.29**	(0.26)	**1.19**	(0.26)	**1.36**	(0.28)	**0.51**	(0.24)
ni	**0.57**	(0.22)	**0.8**	(0.21)	**0.86**	(0.2)	**1.03**	(0.2)	**1**	(0.19)	**0.88**	(0.18)	**0.88**	(0.17)	**0.91**	(0.17)	**1.03**	(0.17)	**1.38**	(0.19)	**0.81**	(0.12)
operprof	**0.53**	(0.27)	**0.87**	(0.26)	**0.91**	(0.22)	**0.85**	(0.19)	**0.88**	(0.18)	**0.93**	(0.17)	**0.86**	(0.17)	**0.96**	(0.17)	**0.89**	(0.17)	**0.93**	(0.19)	**0.4**	(0.17)
pchcurrat	**0.7**	(0.21)	**0.89**	(0.18)	**0.87**	(0.18)	**0.92**	(0.18)	**1**	(0.17)	**0.94**	(0.18)	**0.92**	(0.17)	**0.96**	(0.18)	**0.9**	(0.19)	**0.81**	(0.2)	0.12	(0.09)
pchgm.pchsale	**0.68**	(0.22)	**0.79**	(0.2)	**0.96**	(0.19)	**0.86**	(0.18)	**0.86**	(0.17)	**0.97**	(0.17)	**0.96**	(0.17)	**1.02**	(0.18)	**0.93**	(0.2)	**0.91**	(0.22)	**0.23**	(0.12)
pchquick	**0.85**	(0.21)	**0.94**	(0.2)	**0.82**	(0.18)	**0.95**	(0.17)	**0.97**	(0.17)	**0.99**	(0.18)	**0.92**	(0.18)	**0.88**	(0.18)	**0.94**	(0.19)	**0.9**	(0.21)	0.05	(0.1)
pchsale.pchrect	**0.73**	(0.22)	**1.1**	(0.2)	**1.08**	(0.18)	**0.95**	(0.17)	**1**	(0.17)	**0.88**	(0.17)	**0.82**	(0.17)	**0.91**	(0.19)	**0.89**	(0.2)	**0.86**	(0.22)	0.13	(0.11)
pcm	**0.96**	(0.18)	**0.86**	(0.19)	**0.86**	(0.17)	**0.92**	(0.18)	**0.92**	(0.18)	**0.91**	(0.19)	**0.91**	(0.19)	**0.93**	(0.18)	**0.92**	(0.19)	**0.97**	(0.22)	0.01	(0.14)
pm	0.51	(0.32)	**1.1**	(0.29)	**0.99**	(0.23)	**1.02**	(0.2)	**0.94**	(0.19)	**0.96**	(0.19)	**0.89**	(0.18)	**0.9**	(0.17)	**0.92**	(0.17)	**0.85**	(0.17)	0.34	(0.25)
pm.ia	**1**	(0.19)	**0.88**	(0.18)	**0.93**	(0.18)	**0.97**	(0.2)	**0.93**	(0.2)	**0.93**	(0.2)	**1.02**	(0.2)	**0.8**	(0.19)	**0.88**	(0.2)	**1.01**	(0.19)	0.01	(0.13)
prc52	**0.86**	(0.16)	**0.86**	(0.16)	**0.86**	(0.17)	**0.94**	(0.18)	**1.01**	(0.2)	**0.96**	(0.22)	**0.97**	(0.24)	**0.78**	(0.29)	**0.8**	(0.33)	0.69	(0.42)	-0.17	(0.36)
prof	**0.72**	(0.21)	**0.8**	(0.18)	**0.73**	(0.17)	**0.86**	(0.18)	**0.85**	(0.19)	**0.92**	(0.19)	**0.95**	(0.19)	**1.01**	(0.19)	**1.08**	(0.19)	**1.14**	(0.19)	**0.43**	(0.14)
q	**0.88**	(0.21)	**0.86**	(0.18)	**0.83**	(0.18)	**0.94**	(0.17)	**0.98**	(0.18)	**0.89**	(0.18)	**1.09**	(0.18)	**1.12**	(0.2)	**1.11**	(0.21)	**1.21**	(0.21)	**0.32**	(0.18)
quick	**0.96**	(0.16)	**1**	(0.16)	**0.89**	(0.17)	**0.91**	(0.18)	**0.91**	(0.18)	**0.89**	(0.19)	**0.93**	(0.2)	**1.03**	(0.21)	**0.95**	(0.22)	**0.81**	(0.22)	-0.15	(0.13)
r62	0.24	(0.32)	**0.73**	(0.25)	**0.85**	(0.22)	**0.91**	(0.19)	**0.97**	(0.18)	**1.01**	(0.17)	**0.92**	(0.17)	**0.89**	(0.18)	**1**	(0.2)	**1.37**	(0.25)	**1.13**	(0.28)
retvol	-0.01	(0.38)	0.47	(0.33)	**0.76**	(0.3)	**1.02**	(0.27)	**1.13**	(0.24)	**1.02**	(0.21)	**0.97**	(0.19)	**0.95**	(0.17)	**0.97**	(0.15)	**0.87**	(0.13)	**0.89**	(0.34)
rev1m	**0.59**	(0.25)	**0.79**	(0.19)	**0.84**	(0.18)	**0.93**	(0.17)	**0.9**	(0.17)	**1**	(0.18)	**1.1**	(0.19)	**1.19**	(0.21)	**1.06**	(0.24)	**1.07**	(0.31)	**0.48**	(0.26)
roa2	**0.61**	(0.33)	**0.79**	(0.25)	**0.95**	(0.23)	**0.92**	(0.21)	**0.95**	(0.18)	**0.98**	(0.17)	**0.88**	(0.17)	**0.97**	(0.17)	**0.89**	(0.17)	**0.93**	(0.19)	0.32	(0.24)
roe	**0.68**	(0.32)	**0.88**	(0.25)	**0.99**	(0.21)	**0.9**	(0.19)	**0.94**	(0.18)	**0.88**	(0.17)	**0.9**	(0.17)	**0.91**	(0.18)	**0.91**	(0.18)	**0.93**	(0.2)	0.25	(0.23)
roic	**0.66**	(0.33)	**0.85**	(0.27)	**0.86**	(0.23)	**0.85**	(0.22)	**0.84**	(0.18)	**0.85**	(0.18)	**0.95**	(0.18)	**0.94**	(0.18)	**0.96**	(0.17)	**0.99**	(0.19)	0.34	(0.24)
salecash	**0.89**	(0.24)	**0.96**	(0.21)	**0.95**	(0.2)	**0.95**	(0.19)	**0.89**	(0.18)	**1.02**	(0.18)	**0.88**	(0.18)	**0.91**	(0.18)	**0.93**	(0.17)	**0.93**	(0.18)	0.05	(0.16)
salerec	**0.84**	(0.21)	**0.79**	(0.22)	**0.86**	(0.21)	**0.86**	(0.19)	**0.95**	(0.19)	**1.01**	(0.18)	**0.97**	(0.16)	**0.93**	(0.17)	**0.89**	(0.17)	**1.1**	(0.19)	**0.26**	(0.14)
sat	**0.79**	(0.23)	**0.78**	(0.2)	**0.84**	(0.18)	**0.86**	(0.18)	**0.92**	(0.19)	**0.97**	(0.18)	**1.01**	(0.18)	**0.95**	(0.19)	**1.12**	(0.19)	**1.15**	(0.2)	**0.37**	(0.17)
sat.ia	**0.85**	(0.2)	**0.93**	(0.21)	**0.88**	(0.2)	**0.77**	(0.18)	**0.97**	(0.17)	**0.92**	(0.19)	**0.9**	(0.19)	**0.9**	(0.17)	**1.02**	(0.18)	**1.13**	(0.19)	**0.27**	(0.12)
sgr	**0.81**	(0.23)	**0.93**	(0.22)	**0.95**	(0.2)	**0.91**	(0.18)	**0.93**	(0.17)	**0.98**	(0.16)	**0.91**	(0.17)	**0.95**	(0.17)	**1.05**	(0.2)	**0.86**	(0.22)	0.05	(0.14)
sp	**0.7**	(0.22)	**0.85**	(0.17)	**0.86**	(0.17)	**1.01**	(0.17)	**1.03**	(0.18)	**1.03**	(0.19)	**1.13**	(0.19)	**1.13**	(0.2)	**1.24**	(0.21)	**1.29**	(0.24)	**0.59**	(0.2)
std.dolvol	**1.15**	(0.18)	**1.15**	(0.19)	**1.1**	(0.19)	**1.02**	(0.19)	**0.98**	(0.19)	**1.02**	(0.19)	**0.98**	(0.19)	**0.96**	(0.19)	**0.89**	(0.18)	**0.85**	(0.17)	**-0.31**	(0.13)
std.dolvol36m	**0.78**	(0.17)	**0.99**	(0.17)	**0.97**	(0.17)	**1.07**	(0.18)	**1.05**	(0.18)	**1.06**	(0.18)	**1.09**	(0.18)	**1.13**	(0.18)	**1.13**	(0.18)	**1.24**	(0.17)	**0.46**	(0.13)
std.turn	**0.94**	(0.28)	**1.06**	(0.25)	**1.08**	(0.23)	**1.05**	(0.21)	**1.06**	(0.2)	**1.07**	(0.19)	**0.84**	(0.17)	**0.9**	(0.17)	**0.86**	(0.16)	**0.79**	(0.15)	-0.14	(0.23)
tan	**1.2**	(0.26)	**0.94**	(0.23)	**1.04**	(0.2)	**0.92**	(0.2)	**0.99**	(0.19)	**0.9**	(0.17)	**0.85**	(0.17)	**0.9**	(0.16)	**0.94**	(0.17)	**0.92**	(0.19)	-0.27	(0.18)
tang	**1.2**	(0.26)	**0.94**	(0.23)	**1.04**	(0.2)	**0.92**	(0.2)	**0.99**	(0.19)	**0.9**	(0.17)	**0.85**	(0.17)	**0.9**	(0.16)	**0.94**	(0.17)	**0.92**	(0.19)	-0.27	(0.18)
turn	**0.88**	(0.31)	**0.97**	(0.25)	**0.96**	(0.21)	**0.91**	(0.19)	**0.97**	(0.18)	**0.91**	(0.16)	**0.89**	(0.16)	**0.87**	(0.16)	**0.94**	(0.15)	**0.9**	(0.15)	0.01	(0.26)
zerotrade	**0.97**	(0.15)	**0.94**	(0.17)	**0.87**	(0.16)	**0.87**	(0.16)	**0.84**	(0.15)	**0.96**	(0.17)	**0.96**	(0.19)	**0.98**	(0.21)	**1.05**	(0.24)	**0.94**	(0.3)	-0.03	(0.24)

Table 4: **Jensen alpha's and Information ratios**

This table reports average monthly portfolio returns, Jensen alpha's and Information ratios in percentage. The statistics are reported on unadjusted returns and in excess of the CAPM or Fama and French three factor model. The null hypothesis suggests that the intercepts are less equal zero and is thus a one-sided t-test. Annualized Sharpe and information ratios are approximated by $\sqrt{12} \cdot \hat{\alpha}/\hat{\sigma_{\epsilon}}$. In the case of Sharpe ratios, $\hat{\alpha}$ is the average return of a long-short anomaly portfolio. Standard errors are reported in parentheses. Significant average returns at a single critical value of 5% are printed in bold type. The standard errors of the information ratios are estimated by moving block bootstrapping the data 10,000 times per month to retain the autocorrelation structure in the returns.

	Return				**CAPM**				**FF3**			
Anomaly	Avg	s.e.	SR	s.e.	$\hat{\alpha}$	s.e.	IR	s.e.	$\hat{\alpha}$	s.e.	IR	s.e.
a2me	0.26	(0.21)	0.17	(0.14)	0.19	(0.21)	0.11	(0.14)	-0.44	(0.14)	-0.39	(0.14)
age.m	-0.03	(0.21)	-0.02	(0.16)	0.27	(0.2)	0.21	(0.16)	0.1	(0.16)	0.09	(0.16)
agr	**0.38**	(0.15)	**0.35**	(0.14)	**0.44**	(0.15)	**0.4**	(0.14)	0.15	(0.13)	0.17	(0.14)
at	0.27	(0.26)	0.14	(0.14)	0.47	(0.25)	0.25	(0.14)	**0.5**	(0.14)	**0.45**	(0.14)
bab	-0.25	(0.3)	-0.11	(0.14)	-0.85	(0.22)	-0.55	(0.14)	-0.75	(0.21)	-0.5	(0.14)
baspread	0.34	(0.4)	0.12	(0.14)	**0.82**	(0.37)	**0.31**	(0.14)	**1.03**	(0.3)	**0.47**	(0.14)
bm	0.34	(0.22)	0.21	(0.14)	0.29	(0.22)	0.18	(0.14)	-0.41	(0.14)	-0.42	(0.14)
bm.ia	**0.33**	(0.15)	**0.31**	(0.14)	**0.32**	(0.15)	**0.28**	(0.14)	0.05	(0.13)	0.05	(0.14)
c2d	0.29	(0.23)	0.17	(0.14)	**0.49**	(0.22)	**0.29**	(0.14)	**0.64**	(0.19)	**0.48**	(0.14)
cashdebt	0.37	(0.23)	0.22	(0.14)	**0.57**	(0.22)	**0.33**	(0.14)	**0.69**	(0.19)	**0.53**	(0.14)
cashpr	**0.35**	(0.13)	**0.36**	(0.14)	**0.43**	(0.13)	**0.43**	(0.14)	0.05	(0.09)	0.06	(0.14)
chato	0.17	(0.15)	0.16	(0.14)	0.22	(0.15)	0.2	(0.14)	0.2	(0.15)	0.19	(0.14)
chatoia	0.16	(0.13)	0.17	(0.14)	0.22	(0.13)	0.22	(0.14)	0.21	(0.13)	0.23	(0.14)
chcsho	**0.64**	(0.11)	**0.76**	(0.14)	**0.74**	(0.11)	**0.95**	(0.14)	**0.53**	(0.1)	**0.76**	(0.14)
chempia	0.04	(0.13)	0.04	(0.14)	-0.06	(0.13)	-0.06	(0.14)	0.07	(0.12)	0.07	(0.14)
chinv	**0.43**	(0.14)	**0.41**	(0.14)	**0.48**	(0.14)	**0.48**	(0.14)	**0.27**	(0.13)	**0.28**	(0.14)
chlogso	**0.81**	(0.12)	**0.93**	(0.14)	**0.91**	(0.11)	**1.13**	(0.14)	**0.86**	(0.11)	**1.04**	(0.14)
chmom	0.22	(0.21)	0.14	(0.14)	0.09	(0.21)	0.07	(0.14)	0.12	(0.21)	0.09	(0.14)
chnwc	**0.5**	(0.13)	**0.51**	(0.14)	**0.56**	(0.13)	**0.6**	(0.14)	**0.58**	(0.13)	**0.61**	(0.14)
chpm	-0.11	(0.14)	-0.1	(0.14)	-0.08	(0.14)	-0.08	(0.14)	-0.18	(0.14)	-0.2	(0.14)
chpmia	0.11	(0.17)	0.09	(0.14)	0.12	(0.17)	0.09	(0.14)	0.07	(0.17)	0.05	(0.14)
cto	0.22	(0.17)	0.18	(0.14)	0.18	(0.17)	0.15	(0.14)	**0.47**	(0.14)	**0.45**	(0.14)
currat	0.08	(0.15)	0.07	(0.14)	**0.31**	(0.12)	**0.34**	(0.14)	**0.28**	(0.11)	**0.38**	(0.14)
daily.Beta	0.23	(0.28)	0.11	(0.14)	-0.26	(0.22)	-0.16	(0.14)	-0.07	(0.21)	-0.04	(0.14)
debt2p	0.05	(0.21)	0.03	(0.14)	0.03	(0.21)	0.02	(0.14)	-0.54	(0.15)	-0.48	(0.14)
depr	0.11	(0.18)	0.08	(0.14)	-0.12	(0.16)	-0.11	(0.14)	0.02	(0.14)	0.02	(0.14)
dolvol	-0.32	(0.17)	-0.25	(0.14)	-0.45	(0.17)	-0.39	(0.14)	-0.13	(0.12)	-0.14	(0.14)
dy	0.2	(0.23)	0.12	(0.14)	-0.14	(0.2)	-0.09	(0.14)	**0.25**	(0.13)	**0.28**	(0.14)
ebitrev	0.33	(0.25)	0.18	(0.14)	**0.6**	(0.23)	**0.39**	(0.14)	**0.65**	(0.18)	**0.54**	(0.14)
egr	**0.39**	(0.15)	**0.37**	(0.14)	**0.49**	(0.14)	**0.47**	(0.14)	0.25	(0.13)	0.26	(0.14)
ep	0.42	(0.24)	0.24	(0.14)	**0.64**	(0.23)	**0.41**	(0.14)	**0.53**	(0.21)	**0.38**	(0.14)
eps	0	(0.21)	0	(0.14)	0.27	(0.19)	0.18	(0.14)	**0.43**	(0.15)	**0.38**	(0.14)
ga.eps	0.16	(0.14)	0.16	(0.14)	0.19	(0.14)	0.18	(0.14)	0.24	(0.14)	0.24	(0.14)
gma	0.29	(0.17)	0.23	(0.14)	0.33	(0.17)	**0.27**	(0.14)	**0.75**	(0.13)	**0.76**	(0.14)
herf	0.1	(0.14)	0.1	(0.14)	-0.01	(0.14)	-0.01	(0.14)	0.08	(0.13)	0.08	(0.14)
hire	0.18	(0.13)	0.19	(0.14)	**0.26**	(0.13)	**0.27**	(0.14)	0.02	(0.11)	0.02	(0.14)
ill	0.39	(0.23)	0.23	(0.14)	0.42	(0.23)	**0.27**	(0.14)	0.06	(0.16)	0.05	(0.14)
indmom6m	0.09	(0.13)	0.09	(0.14)	0.14	(0.13)	0.14	(0.14)	0.21	(0.13)	0.24	(0.14)

	Return				CAPM				FF3			
Anomaly	Avg	s.e.	SR	s.e.	$\hat{\alpha}$	s.e.	IR	s.e.	$\hat{\alpha}$	s.e.	IR	s.e.
indmom6m.ia	0.19	(0.21)	0.12	(0.14)	0.21	(0.21)	0.14	(0.14)	0.25	(0.21)	0.17	(0.14)
indrev1m	**0.76**	(0.21)	**0.48**	(0.14)	**0.58**	(0.21)	**0.39**	(0.14)	**0.51**	(0.21)	**0.35**	(0.14)
invest	**0.51**	(0.12)	**0.57**	(0.14)	**0.56**	(0.12)	**0.64**	(0.14)	**0.35**	(0.11)	**0.46**	(0.14)
ipm	0.24	(0.25)	0.13	(0.14)	**0.55**	(0.23)	**0.35**	(0.14)	**0.68**	(0.18)	**0.47**	(0.14)
ldp	-0.13	(0.19)	-0.09	(0.14)	0.2	(0.16)	0.18	(0.14)	0.13	(0.1)	0.18	(0.14)
lev	0.24	(0.21)	0.16	(0.14)	0.23	(0.21)	0.15	(0.14)	-0.36	(0.14)	-0.36	(0.14)
lgr	0.18	(0.11)	0.21	(0.14)	0.21	(0.11)	0.25	(0.14)	0.05	(0.11)	0.07	(0.14)
lprc	**1.81**	(0.24)	**1.04**	(0.14)	**1.62**	(0.23)	**1.09**	(0.14)	**1.46**	(0.22)	**0.96**	(0.14)
maxret	0.53	(0.28)	0.26	(0.14)	**0.93**	(0.24)	**0.54**	(0.14)	**1.02**	(0.19)	**0.74**	(0.14)
mom11m	**1.45**	(0.32)	**0.61**	(0.14)	**1.62**	(0.32)	**0.7**	(0.14)	**1.89**	(0.32)	**0.8**	(0.14)
mom12m	**1.34**	(0.32)	**0.58**	(0.14)	**1.52**	(0.31)	**0.7**	(0.14)	**1.8**	(0.31)	**0.79**	(0.14)
mom18m	**0.58**	(0.22)	**0.37**	(0.14)	**0.56**	(0.22)	**0.32**	(0.14)	0.2	(0.2)	0.15	(0.14)
mom36m	0.27	(0.25)	0.15	(0.14)	0.23	(0.25)	0.13	(0.14)	-0.28	(0.21)	-0.17	(0.14)
mom6m	**1.08**	(0.31)	**0.47**	(0.14)	**1.28**	(0.31)	**0.57**	(0.14)	**1.44**	(0.31)	**0.71**	(0.14)
mom6m.l6	**1.33**	(0.25)	**0.74**	(0.14)	**1.35**	(0.25)	**0.77**	(0.14)	**1.57**	(0.24)	**0.91**	(0.14)
mve	**1.16**	(0.27)	**0.59**	(0.14)	**1.12**	(0.27)	**0.55**	(0.14)	**0.75**	(0.2)	**0.61**	(0.14)
mve.firm	**1.15**	(0.27)	**0.58**	(0.14)	**1.11**	(0.27)	**0.56**	(0.14)	**0.74**	(0.2)	**0.58**	(0.14)
mve.firm.ia	**0.56**	(0.14)	**0.56**	(0.14)	**0.43**	(0.13)	**0.47**	(0.14)	**0.37**	(0.09)	**0.56**	(0.14)
mve.ia	**0.35**	(0.14)	**0.35**	(0.14)	0.25	(0.13)	0.26	(0.14)	**0.22**	(0.09)	**0.37**	(0.14)
mve.jun.log	**0.51**	(0.24)	**0.28**	(0.14)	**0.49**	(0.25)	**0.29**	(0.14)	0.16	(0.16)	0.14	(0.14)
ni	**0.81**	(0.12)	**0.89**	(0.14)	**0.91**	(0.12)	**0.96**	(0.14)	**0.85**	(0.12)	**1.05**	(0.14)
operprof	**0.4**	(0.17)	**0.32**	(0.14)	**0.55**	(0.17)	**0.46**	(0.14)	**0.78**	(0.14)	**0.81**	(0.14)
pchcurrat	0.12	(0.09)	0.17	(0.14)	0.13	(0.09)	0.19	(0.14)	0.13	(0.1)	0.2	(0.14)
pchgm.pchsale	**0.23**	(0.12)	**0.27**	(0.14)	0.2	(0.12)	0.23	(0.14)	**0.33**	(0.11)	**0.42**	(0.14)
pchquick	0.05	(0.1)	0.07	(0.14)	0.04	(0.1)	0.05	(0.14)	0.04	(0.1)	0.05	(0.14)
pchsale.pchrect	0.13	(0.11)	0.15	(0.14)	0.12	(0.11)	0.13	(0.14)	0.06	(0.11)	0.08	(0.14)
pcm	0.01	(0.14)	0.01	(0.14)	-0.08	(0.13)	-0.07	(0.14)	-0.3	(0.11)	-0.36	(0.14)
pm	0.34	(0.25)	0.19	(0.14)	**0.62**	(0.23)	**0.38**	(0.14)	**0.64**	(0.17)	**0.49**	(0.14)
pm.ia	0.01	(0.13)	0.01	(0.14)	0.02	(0.14)	0.02	(0.14)	**0.29**	(0.12)	**0.33**	(0.14)
prc52	-0.17	(0.36)	-0.06	(0.14)	-0.68	(0.32)	-0.31	(0.14)	-0.9	(0.3)	-0.44	(0.14)
prof	**0.43**	(0.14)	**0.41**	(0.14)	**0.48**	(0.14)	**0.46**	(0.14)	**0.57**	(0.14)	**0.58**	(0.14)
q	0.32	(0.18)	0.24	(0.14)	**0.37**	(0.18)	**0.27**	(0.14)	-0.19	(0.12)	-0.22	(0.14)
quick	-0.15	(0.13)	-0.16	(0.14)	-0.32	(0.12)	-0.37	(0.14)	-0.27	(0.11)	-0.32	(0.14)
r62	**1.13**	(0.28)	**0.56**	(0.14)	**1.28**	(0.28)	**0.6**	(0.14)	**1.42**	(0.28)	**0.77**	(0.14)
retvol	**0.89**	(0.34)	**0.36**	(0.14)	**1.38**	(0.3)	**0.67**	(0.14)	**1.5**	(0.25)	**0.75**	(0.14)
rev1m	0.48	(0.26)	0.26	(0.14)	0.31	(0.25)	0.15	(0.14)	0.23	(0.25)	0.13	(0.14)
roa2	0.32	(0.24)	0.18	(0.14)	**0.55**	(0.23)	**0.37**	(0.14)	**0.71**	(0.2)	**0.53**	(0.14)
roe	0.25	(0.23)	0.15	(0.14)	**0.46**	(0.22)	**0.28**	(0.14)	**0.68**	(0.18)	**0.51**	(0.14)
roic	0.34	(0.24)	0.19	(0.14)	**0.56**	(0.23)	**0.4**	(0.14)	**0.71**	(0.19)	**0.51**	(0.14)
salecash	0.05	(0.16)	0.04	(0.14)	0.18	(0.16)	0.16	(0.14)	0.18	(0.16)	0.17	(0.14)
salerec	0.26	(0.14)	0.25	(0.14)	0.27	(0.14)	0.26	(0.14)	**0.46**	(0.13)	**0.48**	(0.14)
sat	**0.37**	(0.17)	**0.29**	(0.14)	**0.38**	(0.17)	**0.33**	(0.14)	**0.62**	(0.16)	**0.58**	(0.14)
sat.ia	**0.27**	(0.12)	**0.31**	(0.14)	**0.33**	(0.12)	**0.4**	(0.14)	**0.36**	(0.12)	**0.39**	(0.14)
sgr	0.05	(0.14)	0.05	(0.14)	0.11	(0.14)	0.11	(0.14)	-0.15	(0.13)	-0.17	(0.14)
sp	**0.59**	(0.2)	**0.41**	(0.14)	**0.58**	(0.2)	**0.42**	(0.14)	-0.01	(0.13)	-0.01	(0.14)
std.dolvol	-0.31	(0.13)	-0.31	(0.14)	-0.36	(0.13)	-0.4	(0.14)	-0.17	(0.09)	-0.27	(0.14)
std.dolvol36m	**0.46**	(0.13)	**0.49**	(0.14)	**0.55**	(0.13)	**0.62**	(0.14)	**0.25**	(0.08)	**0.46**	(0.14)
std.turn	-0.14	(0.23)	-0.08	(0.14)	0.22	(0.2)	0.15	(0.14)	0.17	(0.15)	0.15	(0.14)
tan	-0.27	(0.18)	-0.21	(0.14)	-0.16	(0.18)	-0.13	(0.14)	-0.43	(0.16)	-0.38	(0.14)
tang	-0.27	(0.18)	-0.21	(0.14)	-0.16	(0.18)	-0.13	(0.14)	-0.43	(0.16)	-0.38	(0.14)
turn	0.01	(0.26)	0.01	(0.14)	**0.45**	(0.22)	**0.27**	(0.14)	0.26	(0.18)	0.21	(0.14)
zerotrade	-0.03	(0.24)	-0.02	(0.14)	-0.48	(0.19)	-0.36	(0.14)	-0.23	(0.17)	-0.19	(0.14)

Table 5: **Linear Model In-Sample Results**

This table reports selected characteristics obtained from penalized regressions from a total universe of 90 market anomalies. The columns report model selection results for different samples, regularization models and number of steps. The in-sample Sharpe ratios summarize the performance of equally-weighted hedge portfolios going long the decile of stocks with highest predicted returns and shorting the decile of stocks with lowest predicted returns. The sample period spans from January 1965 to December 2018 unless otherwise specified.

Firms	All	All	All	NYSE 10	NYSE 10	NYSE 10	NYSE 20	NYSE 20	NYSE 20
Period	1965-2018	1965-2018	1965-2018	1965-2018	1965-2018	1965-2018	1965-2018	1965-2018	1965-2018
Sample Size	1,417,151	1,417,151	1,417,151	861,982	861,982	861,982	690,545	690,545	690,545
Regularization	SCAD	Elastic Net	Adaptive LASSO	SCAD	Elastic Net	Adaptive LASSO	SCAD	Elastic Net	Adaptive LASSO
# Steps	1	1	2	1	1	2	1	1	2
# Selected	65	86	59	42	75	45	37	62	38
IS Sharpe Ratio	2.95	2.91	2.88	3.47	3.47	3.48	3.27	3.22	3.29
Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
a2me	a2me	a2me	a2me	a2me	a2me	a2me	a2me	a2me	a2me
agr	agr	agr	agr		agr				
at		at	at		at			at	
bab	bab	bab	bab		bab			bab	
baspread		baspread		baspread	baspread	baspread		baspread	
bm	bm	bm	bm		bm			bm	
bm_ia	bm_ia	bm_ia		bm_ia	bm_ia	bm_ia	bm_ia	bm_ia	bm_ia
c2d	c2d	c2d	c2d	c2d	c2d	c2d		c2d	
cashdebt		cashdebt			cashdebt			cashdebt	
cashpr	cashpr	cashpr	cashpr						
chaTO		chaTO		chaTO	chaTO	chaTO	chaTO	chaTO	chato
chatoia	chatoia	chatoia	chatoia		chatoia				
chempia		chempia		chempia	chempia	chempia	chempia	chempia	chempia
chinv	chinv	chinv	chinv	chinv	chinv	chinv	chinv	chinv	
chmom	chmom	chmom	chmom	chmom	chmom	chmom	chmom	chmom	chmom
chnwc	chnwc	chnwc	chnwc	chnwc	chnwc	chnwc	chnwc	chnwc	chnwc
chpm	chpm	chpm			chpm			chpm	
chpmia		chpmia			chpmia			chpmia	
cTO	cTO	cTO	cTO		cTO	cTO		cTO	
currat	currat	currat	currat		currat				
daily_Beta	daily_Beta	daily_Beta	daily_Beta	daily_Beta	daily_Beta	daily_Beta	daily_Beta	daily_Beta	daily_Beta
debt2p	debt2p	debt2p						debt2p	
depr	depr	depr		depr	depr	depr	depr	depr	depr
dolvol	dolvol	dolvol	dolvol	dolvol	dolvol		dolvol	dolvol	dolvol
dy	dy	dy	dy	dy	dy	dy	dy	dy	dy
ebitrev	ebitrev	ebitrev	ebitrev		ebitrev				
egr	egr	egr	egr	egr	egr	egr	egr	egr	egr
ep	ep	ep	ep		ep			ep	
eps	eps	eps	eps	eps	eps	eps	eps	eps	eps
ga_eps		ga_eps			ga_eps				
gma	gma	gma	gma						
herf	herf	herf	herf	herf	herf	herf	herf	herf	herf
hire		hire							
ill	ill	ill	ill	ill	ill	ill		ill	
indmom6m	indmom6m	indmom6m	indmom6m	indmom6m	indmom6m	indmom6m	indmom6m	indmom6m	indmom6m

indmom6m_ia	indmom6m_ia	indmom6m_ia	indmom6m_ia						
indrev1m	indrev1m	indrev1m	indrev1m	indrev1m	indrev1m	indrev1m	indrev1m	indrev1m	indrev1m
invest		invest							
ipm		ipm							
ldp	ldp	ldp	ldp	ldp	ldp	ldp	ldp	ldp	ldp
lev	lev	lev	lev		lev	lev		lev	
lgr		lgr		lgr	lgr				
lprc	lprc	lprc	lprc	lprc	lprc	lprc	lprc	lprc	lprc
maxret	maxret	maxret	maxret						
mom11m	mom11m	mom11m	mom11m		mom11m	mom11m		mom11m	mom11m
mom12m	mom12m	mom12m	mom12m		mom12m	mom12m		mom12m	
mom18m	mom18m	mom18m	mom18m	mom18m	mom18m	mom18m	mom18m	mom18m	mom18m
mom6m	mom6m	mom6m	mom6m		mom6m			mom6m	
mom6m_l6	mom6m_l6	mom6m_l6	mom6m_l6	mom6m_l6	mom6m_l6	mom6m_l6	mom6m_l6	mom6m_l6	mom6m_l6
mve		mve							
mve_firm		mve_firm							
mve_firm_ia	mve_firm_ia	mve_firm_ia	mve_firm_ia	mve_firm_ia	mve_firm_ia	mve_firm_ia	mve_firm_ia	mve_firm_ia	mve_firm_ia
mve_ia	mve_ia	mve_ia	mve_ia	mve_ia	mve_ia	mve_ia	mve_ia	mve_ia	mve_ia
mve_jun_log	mve_jun_log	mve_jun_log	mve_jun_log	mve_jun_log	mve_jun_log	mve_jun_log		mve_jun_log	mve_jun_log
ni	ni	ni	ni	ni	ni	ni	ni	ni	ni
operprof	operprof	operprof	operprof		operprof				
pchcurrat	pchcurrat	pchcurrat	pchcurrat		pchcurrat			pchcurrat	
pchgm_pchsale	pchgm_pchsale	pchgm_pchsale			pchgm_pchsale				
pchquick		pchquick							
pchsale_pchrect	pchsale_pchrect	pchsale_pchrect			pchsale_pchrect				
pcm		pcm							
pm	pm	pm	pm		pm				
pm_ia	pm_ia	pm_ia	pm_ia	pm_ia	pm_ia	pm_ia	pm_ia	pm_ia	pm_ia
prc52	prc52	prc52	prc52	prc52	prc52	prc52	prc52	prc52	prc52
prof	prof	prof		prof	prof	prof	prof	prof	prof
ps	ps	ps	ps		ps	ps		ps	ps
q	q	q			q			q	
quick	quick	quick	quick	quick	quick		quick	quick	
r62	r62	r62	r62		r62			r62	
retvol	retvol	retvol	retvol	retvol	retvol	retvol	retvol	retvol	retvol
rev1m	rev1m	rev1m	rev1m	rev1m	rev1m	rev1m	rev1m	rev1m	rev1m
roa2		roa2							
roe		roe	roe						
roic	roic	roic	roic		roic	roic	roic	roic	roic
salecash		salecash	salecash	salecash	salecash	salecash	salecash	salecash	salecash
salerec	salerec	salerec		salerec	salerec		salerec	salerec	
sat		sat	sat		sat				
sat_ia	sat_ia	sat_ia	sat_ia	sat_ia	sat_ia	sat_ia		sat_ia	
sgr		sgr			sgr				
sp		sp	sp		sp	sp		sp	sp
std_dolvol	std_dolvol	std_dolvol	std_dolvol	std_dolvol	std_dolvol	std_dolvol	std_dolvol	std_dolvol	std_dolvol
std_turn	std_turn	std_turn	std_turn	std_turn	std_turn	std_turn	std_turn	std_turn	std_turn
tan	tan	tan			tan				
tang	tang	tang			tang				
turn	turn	turn	turn	turn	turn	turn	turn	turn	turn
zerotrade	zerotrade	zerotrade	zerotrade	zerotrade	zerotrade	zerotrade	zerotrade	zerotrade	zerotrade
chcsho					chcsho			chcsho	
mom36m				mom36m	mom36m	mom36m	mom36m	mom36m	mom36m
std_dolvol36m					std_dolvol36m				
chlogso								chlogso	

Table 6: **Non-Linear Model In-Sample Results**

This table reports selected characteristics obtained from penalized regressions from a total universe of 90 market anomalies. The columns report model selection results for different samples, regularization models and number of steps. The in-sample Sharpe ratios summarize the performance of equally-weighted hedge portfolios going long the decile of stocks with highest predicted returns and shorting the decile of stocks with lowest predicted returns. The sample period spans from January 1965 to December 2018 unless otherwise specified.

Firms	All	All	NYSE 20	NYSE 20	NYSE 20	NYSE 20
Period	1965-2018	1965-2018	1965-2018	1965-2018	1965-2018	1965-2018
Sample Size	1,417,151	1,417,151	690,545	690,545	690,545	690,545
Regularization	SCAD	Adaptive LASSO	LASSO	Elastic Net	Adaptive LASSO	SCAD
# Steps	1	2	1	1	2	1
Knots	19	19	19	19	19	19
# Selected	3	6	9	27	8	2
IS Sharpe Ratio	2.35	2.57	2.77	2.94	2.82	1.89
Models	(1)	(2)	(3)	(4)	(5)	(6)
a2me		a2me		a2me		
bm				bm		
bm_ia				bm_ia		
chaTO				chaTO		
chcsho				chcsho		
chlogso				chlogso		
chnwc				chnwc		
daily_Beta				daily_Beta		
depr				depr		
eps			eps	eps	eps	
herf				herf		
indmom6m				indmom6m		
indmom6m_ia		indmom6m_ia				
indrev1m	indrev1m	indrev1m	indrev1m	indrev1m	indrev1m	indrev1m
ldp			ldp	ldp	ldp	
lprc	lprc	lprc	lprc	lprc	lprc	lprc
mom11m			mom11m	mom11m	mom11m	
mom18m		mom18m	mom18m	mom18m	mom18m	
mom6m_l6			mom6m_l6	mom6m_l6	mom6m_l6	
ni				ni		
pm_ia				pm_ia		
prc52	prc52	prc52	prc52	prc52	prc52	
prof				prof		
retvol				retvol		
sp				sp		
std_turn				std_turn		
ep			ep	ep		
roic				roic		

Table 7: **Linear Model Out-of-Sample Prediction**

This table reports out-of-sample, value-weighted hedge portfolio returns, going long the stocks in the highest predicted return decile and shorting the stocks in the lowest predicted return decile for different estimation periods and regularization terms. The table reports the first four moments of monthly percentage returns, turnover, and predictive slopes and R^2 for the hedge portfolios in Panel A, as well as for the long legs in Panel B and the short legs in Panel C. The model selection and estimation period spans from January 1965 to December of the year before start of the 10 year out-of-sample period.

Firms	NYSE 20	NYSE 20	NYSE 20	NYSE 20	NYSE 20	NYSE 20	NYSE 20
Period	1965-1982	1965-1982	1965-1982	1965-1990	1965-1990	1983-2000	1983-2000
Sample Size	138,733	138,733	138,733	227,802	227,802	258,044	258,044
Regularization	Adaptive LASSO	Elastic Net	SCAD	Adaptive LASSO	SCAD	Adaptive LASSO	SCAD
# Steps	2	1	1	2	1	2	1
Knots	9	9	9	9	9	9	9
# Selected	14	36	19	14	22	18	30
IS Sharpe Ratio	3.84	3.86	3.77	3.97	4.08	4.06	4.34
Panel A: Long-Short Portfolio							
Mean	2.72	2.24	2.9	2.01	2.09	0.85	1.11
Std. Dev.	4.94	4.04	3.99	5.51	5.49	6.26	6.12
Sharpe Ratio	1.91	1.92	2.52	1.27	1.32	0.47	0.63
Skewness	1.1	0.39	0.37	0.17	-0.09	-0.22	-0.24
Kurtosis	3.08	-0.01	0.98	2.1	2.01	3.84	4.46
β	0.51	0.5	0.53	0.86	0.81	0.62	0.61
R^2	0.03	0.03	0.03	0.04	0.04	0.04	0.04
TO	0.42	0.36	0.4	0.39	0.41	0.33	0.34
TO_R	0.42	0.37	0.4	0.4	0.41	0.33	0.34
Panel B: Long Leg							
Mean	2.65	2.3	2.39	2.18	2.27	0.87	1.08
Std. Dev.	7.47	6.16	5.92	5.39	5.78	7.96	8.07
Sharpe Ratio	1.23	1.29	1.4	1.4	1.36	0.38	0.46
Skewness	-0.84	-0.4	-0.67	0.09	-0.06	0.1	0.13
Kurtosis	5.07	1.87	4.01	1.14	0.68	1.49	1.43
β	0.48	0.44	0.45	0.99	0.92	0.71	0.66
R^2	0.06	0.07	0.06	0.08	0.08	0.1	0.09
TO	0.2	0.2	0.2	0.2	0.21	0.21	0.21
TO_R	0.21	0.21	0.21	0.2	0.21	0.22	0.22
Panel C: Short Leg							
Mean	-0.07	0.06	-0.51	0.16	0.18	0.02	-0.03
Std. Dev.	5.46	4.45	5.44	5.2	5.08	4.42	4.42
Sharpe Ratio	-0.04	0.05	-0.32	0.11	0.12	0.01	-0.02
Skewness	-0.89	-0.74	-1.1	-0.22	-0.33	-0.22	-0.28
Kurtosis	5.88	2.8	4.83	0.02	0.55	1.05	1.03
β	0.38	0.33	0.41	0.08	0.09	0.24	0.24
R^2	0.05	0.06	0.05	0.06	0.06	0.08	0.08
TO	0.21	0.16	0.19	0.19	0.19	0.11	0.12
TO_R	0.21	0.16	0.19	0.19	0.19	0.11	0.11

Table 8: **Non-Linear Model Out-of-Sample Prediction**
This table reports out-of-sample, value-weighted hedge portfolio returns, going long the stocks in the highest predicted return decile and shorting the stocks in the lowest predicted return decile for different estimation periods and regularization terms. The table reports the first four moments of monthly percentage returns, turnover, and predictive slopes and R^2 for the hedge portfolios in Panel A, as well as for the long legs in Panel B and the short legs in Panel C. The model selection and estimation period spans from January 1965 to December of the year before start of the 10 year out-of-sample period.

Firms	NYSE 20	NYSE 20	NYSE 20	NYSE 20	NYSE 20	NYSE 20	NYSE 20
Period	1965-1982	1965-1982	1965-1982	1965-1990	1965-1990	1983-2000	1983-2000
Sample Size	138,733	138,733	138,733	227,802	227,802	258,044	258,044
Regularization	Adaptive LASSO	Elastic Net	SCAD	Adaptive LASSO	SCAD	Adaptive LASSO	SCAD
# Steps	2	1	1	2	1	2	1
Knots	9	9	9	9	9	9	9
# Selected	2	16	1	3	4	3	2
IS Sharpe Ratio	2.65	3.57	2.12	3.12	3.07	2.83	2.36
Panel A: Long-Short Portfolio							
Mean	1.18	1.93	0.67	2.21	2.8	1.47	1.5
Std. Dev.	4.04	3.69	4.2	6.56	6.02	6.5	7.49
Sharpe Ratio	1.01	1.81	0.56	1.17	1.61	0.78	0.69
Skewness	0.34	0.78	0.59	0.26	0.14	0.14	-0.28
Kurtosis	0.93	2.56	0.99	2.11	1.77	0.53	0.91
β	0.75	0.83	0.75	0.95	0.97	0.65	0.68
R^2	0.05	0.03	0.05	0.05	0.05	0.05	0.05
TO	0.52	0.4	0.52	0.44	0.43	0.43	0.44
TO_R	0.52	0.41	0.52	0.45	0.43	0.43	0.44
Panel B: Long Leg							
Mean	1.55	2.22	1.3	2.63	3.11	1.32	1.39
Std. Dev.	6.68	6.41	6.79	8.05	7.96	9.9	10.38
Sharpe Ratio	0.8	1.2	0.66	1.13	1.35	0.46	0.47
Skewness	-0.85	-0.63	-0.78	-0.09	0.07	-0.23	-0.46
Kurtosis	3.1	2.86	2.89	1.11	1.63	0.49	0.44
β	0.77	0.9	0.76	1.01	1.06	0.74	0.75
R^2	0.08	0.08	0.08	0.09	0.09	0.1	0.1
TO	0.24	0.21	0.24	0.22	0.22	0.22	0.22
TO_R	0.25	0.22	0.25	0.23	0.23	0.23	0.23
Panel C: Short Leg							
Mean	0.37	0.29	0.63	0.42	0.31	-0.14	-0.1
Std. Dev.	5.54	5.13	5.46	3.98	4.46	5.21	4.51
Sharpe Ratio	0.23	0.2	0.4	0.37	0.24	-0.1	-0.08
Skewness	-1.5	-1.15	-1.45	-0.65	-0.3	-0.83	-0.69
Kurtosis	6.88	4.44	6.09	1.75	0.16	1.38	0.78
β	0.87	0.17	-9.2	0.12	0.28	0.17	0.02
R^2	0.03	0.08	0.04	0.06	0.06	0.11	0.1
TO	0.24	0.19	0.24	0.2	0.19	0.18	0.19
TO_R	0.24	0.19	0.24	0.19	0.18	0.18	0.19

Table 9: **Non-Linear Model Out-of-Sample Rolling Prediction**

This table reports out-of-sample, value-weighted hedge portfolio returns, going long the stocks in the highest predicted return decile and shorting the stocks in the lowest predicted return decile for different out-of-sample periods and regularization terms. The rolling estimation window is 20 years long and starts 1970. The table reports the first four moments of monthly percentage returns and predictive slopes and R^2 for the hedge portfolios in Panel A, as well as for the long legs in Panel B and the short legs in Panel C. The model selection and estimation period spans from January 1965 to December of the year before start of the 10 year out-of-sample period.

OoS Period	1991-2018	1991-2018	1991-2018	1991-2018	1991-2018	1991-2014	1991-1999
Firms	ALL	ALL	ALL	ALL	ALL	ALL	ALL
Sample Size	1,010,876	1,010,876	1,010,876	1,010,876	1,010,876	866,105	347,456
Regularization	MCP	SCAD	Adaptive LASSO	MCP	SCAD	SCAD	SCAD
# Steps	1	1	2	2	2	1	1
Avg. # Selected	11	13	10	10	10	13	13
IS Sharpe Ratio	3	3.01	2.91	2.96	2.95	3.01	3.01
Panel A: Long-Short Portfolio							
Mean	2.63	2.92	2.53	2.53	2.53	2.96	3.87
Std. Dev.	5.02	5.45	5.08	5.08	5.08	5.22	4.37
Sharpe Ratio	1.82	1.86	1.73	1.73	1.73	1.96	3.07
Skewness	1.86	2.35	1.84	1.84	1.84	1.82	1.67
Kurtosis	6.06	9.75	6.21	6.21	6.21	5.88	3.9
β	0.47	0.5	0.47	0.47	0.47	0.51	0.57
R^2	0.04	0.04	0.04	0.04	0.04	0.05	0.08
Panel B: Long Leg							
Mean	3.04	3.29	3.03	3.03	3.03	3.42	4.46
Std. Dev.	8.48	8.53	8.53	8.53	8.53	8.9	7.12
Sharpe Ratio	1.24	1.34	1.23	1.23	1.23	1.33	2.17
Skewness	0.69	0.88	0.68	0.68	0.68	0.6	0.2
Kurtosis	3.29	3.81	3.28	3.28	3.28	2.98	1.84
β	0.59	0.61	0.59	0.59	0.59	0.66	0.8
R^2	0.1	0.09	0.1	0.1	0.1	0.1	0.08
Panel C: Short Leg							
Mean	0.4	0.37	0.5	0.5	0.5	0.47	0.6
Std. Dev.	5.96	5.81	5.91	5.91	5.91	6.11	5.22
Sharpe Ratio	0.23	0.22	0.29	0.29	0.29	0.27	0.4
Skewness	-0.13	-0.1	-0.09	-0.09	-0.09	-0.05	-0.87
Kurtosis	2.25	2.28	2.06	2.06	2.06	2	2.68
β	0.06	-0.02	0.04	0.04	0.04	0.06	-0.21
R^2	0.08	0.08	0.08	0.08	0.08	0.08	0.06

B Figures

Figure 1: **Correlation between Firm Characteristics, 1970-2018**
This heat map presents pairwise Pearson correlations between rank-normalized firm characteristics. Dark blue (red) indicates positively (negatively) correlated features and uncorrelated variables are transparent. Features are ordered by complete hierarchical clustering.

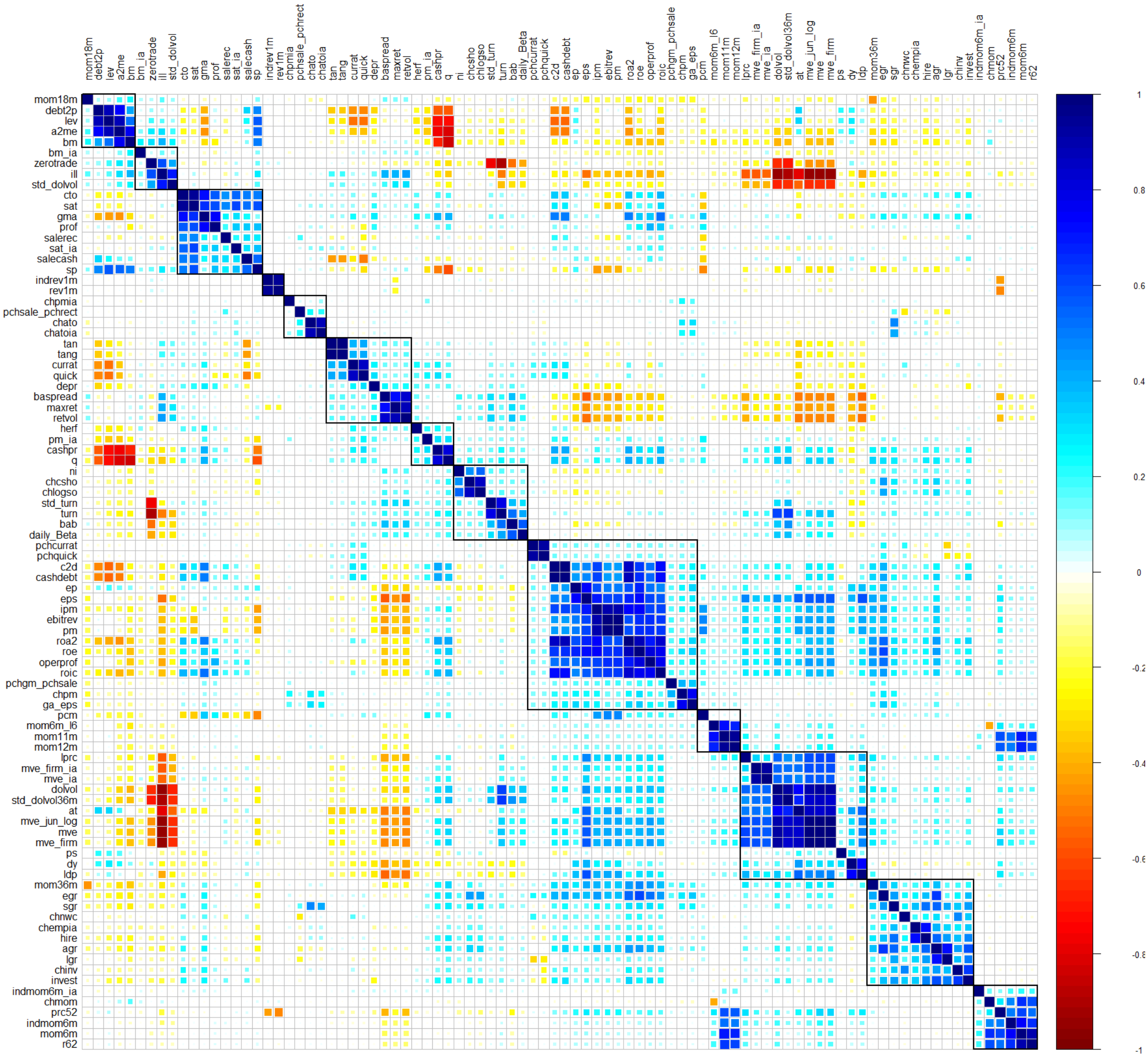


Figure 2: **Correlation between Firm Characteristics NYSE 10%, 1970-2018**
This heat map presents pairwise Pearson correlations between rank-normalized firm characteristics of all stocks that are larger than the first decile of NYSE listed stocks. Dark blue (red) indicates positively (negatively) correlated features and uncorrelated variables are transparent. Features are ordered by complete hierarchical clustering.

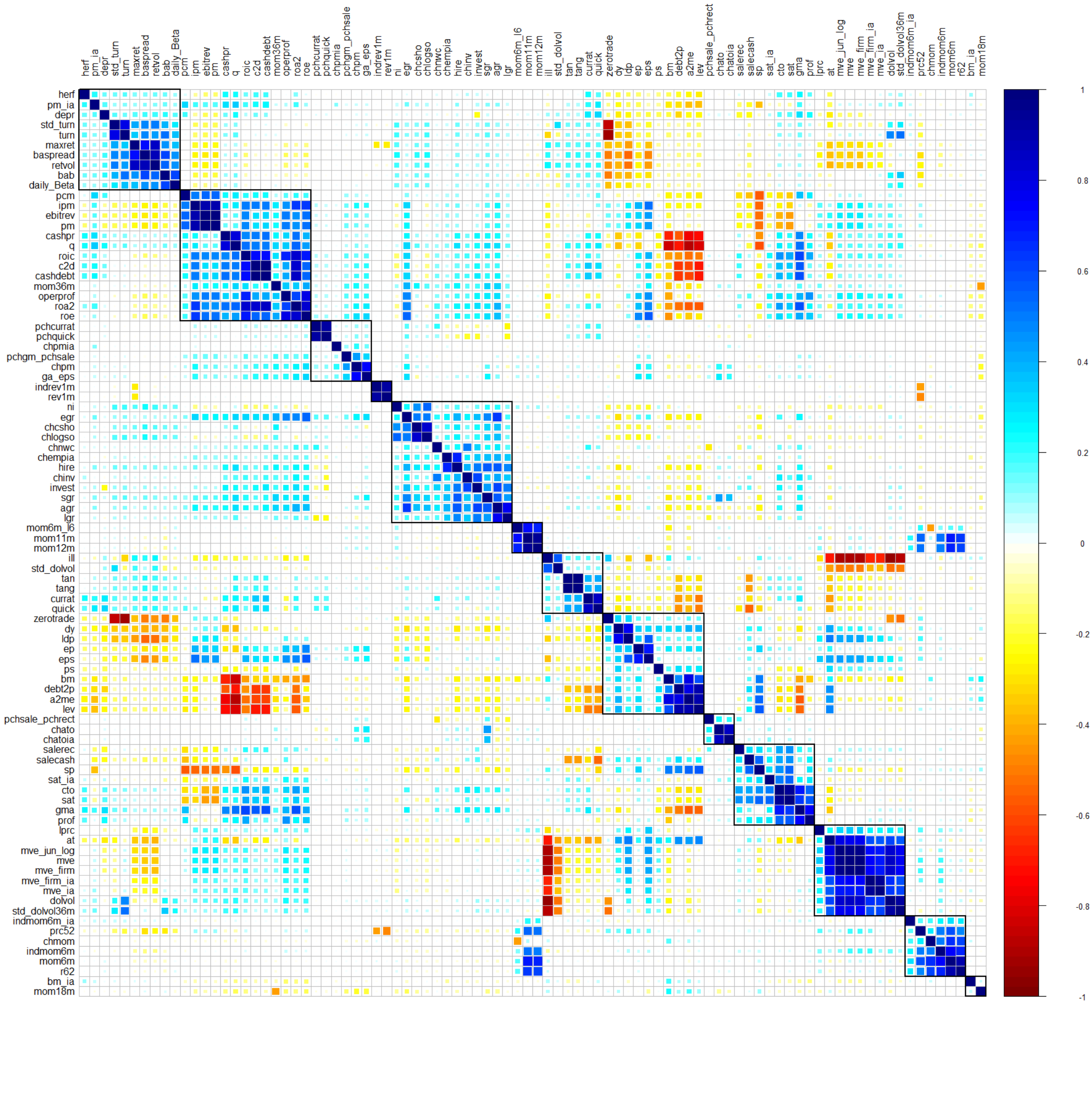


Figure 3: **Correlation between Firm Characteristics NYSE 20%, 1970-2018**
This heat map presents pairwise Pearson correlations between rank-normalized firm characteristics of all stocks that are larger than the second decile of NYSE listed stocks. Dark blue (red) indicates positively (negatively) correlated features and uncorrelated variables are transparent. Features are ordered by complete hierarchical clustering.

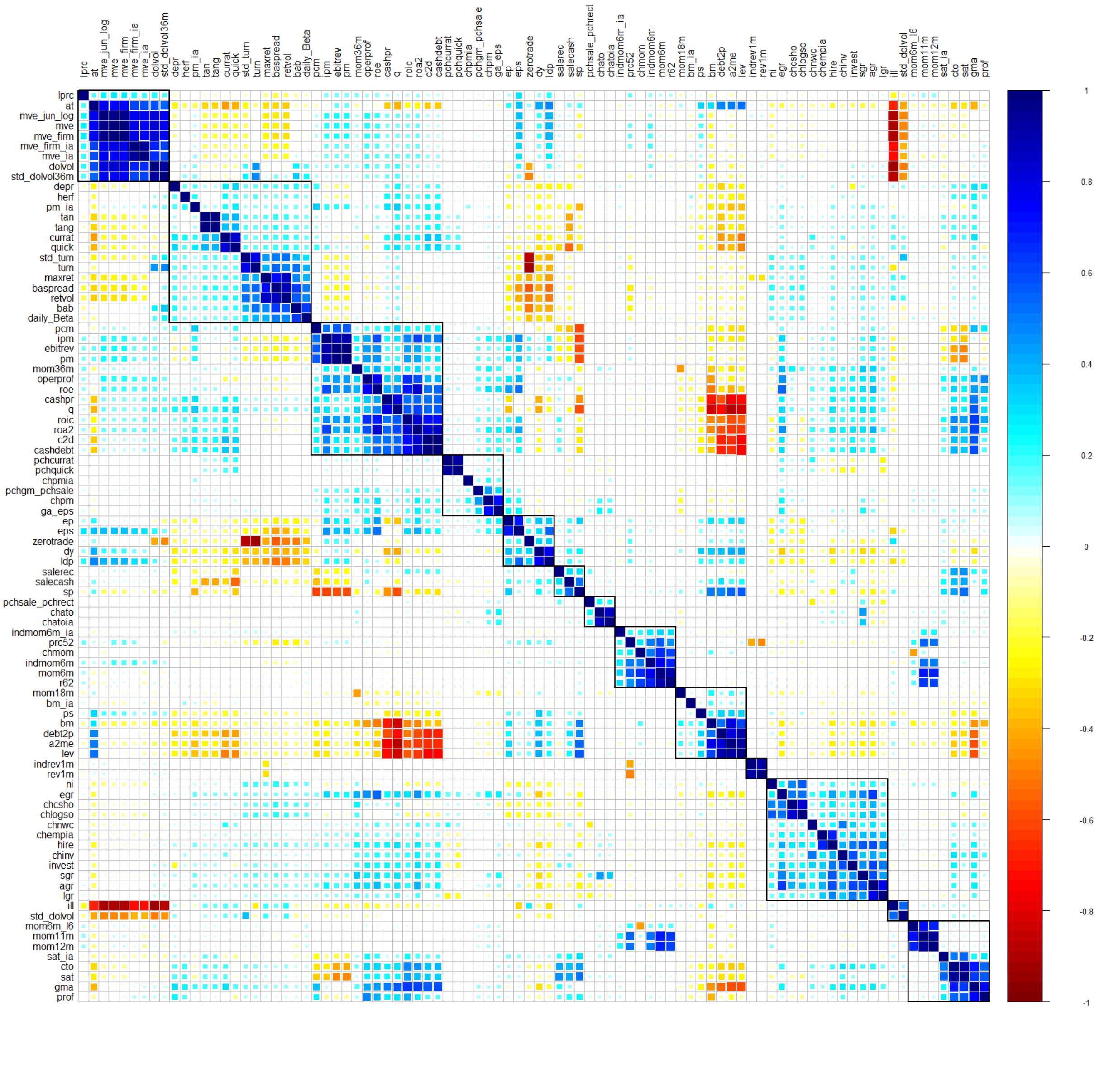


Figure 4: **Anomaly Covariances on CAPM and Fama French 3 Factors, 1965-2018**

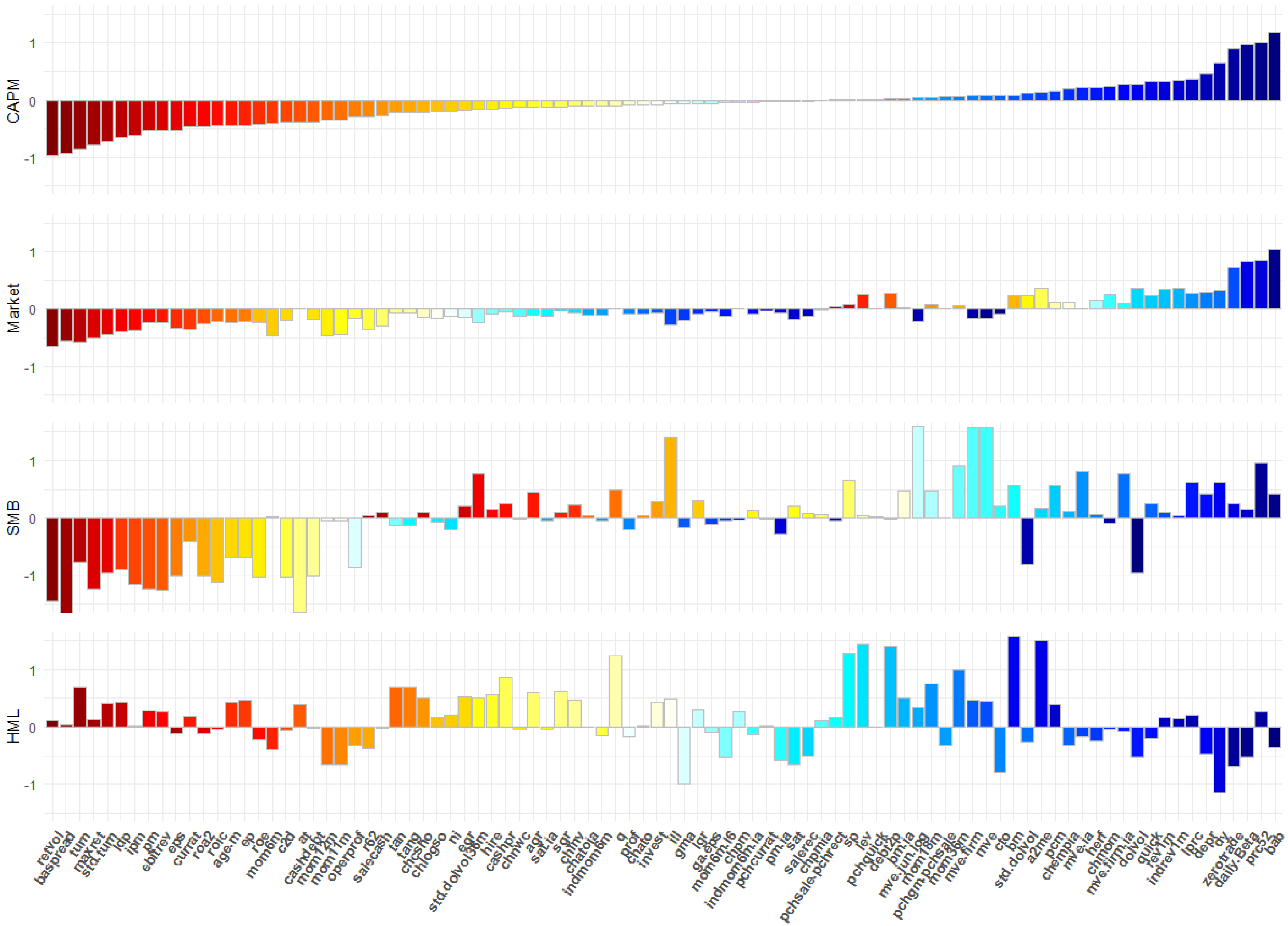
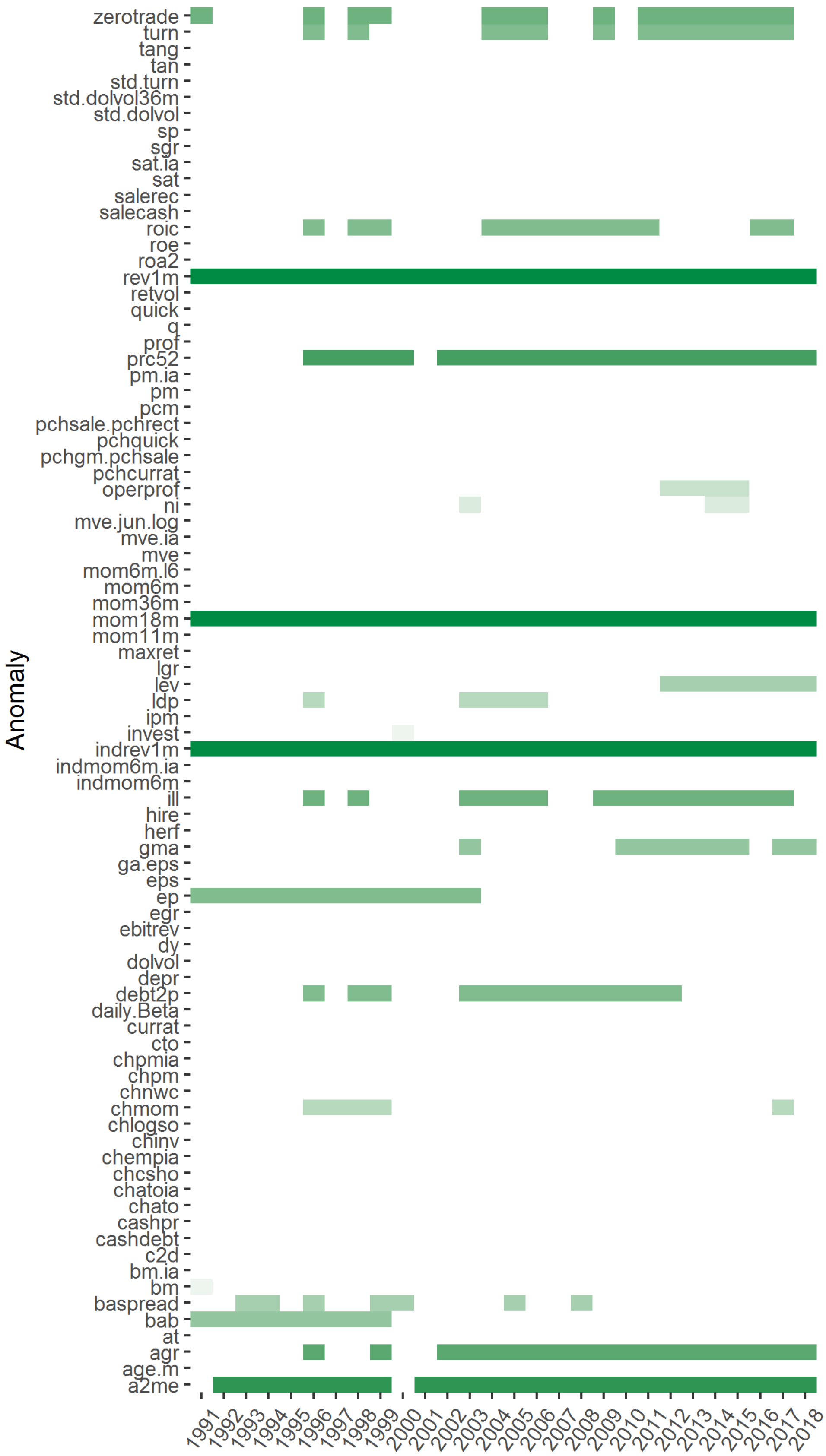


Figure 5: **Time-Varying Feature Subset Selection (Rolling Selection), 1991-2018**



C Anomaly Definition

1. **A2ME:** We follow Bhandari (1988) and define assets-to-market cap as total assets (AT) over market capitalization as of December t-1. Market capitalization is the product of shares outstanding (SHROUT) and price (PRC).

2. **Firm Age:** The number of months that a firm has been listed in the CRSP database. Updated monthly. I exclude stocks with a price below $5.

3. **Asset growth:** Annual percent change in total assets (at).

4. **Total Assets:** Total assets (AT) as in Gandhi and Lustig (2015).

5. **Beta:** We follow Frazzini and Pedersen (2014) and define the CAPM beta as product of correlations between the excess return of stock i and the market excess return and the ratio of volatilities. We calculate volatilities from the standard devi ations of daily log excess returns over a one-year horizon requiring at least 120 observations. We estimate correlations using overlapping three-day log excess returns over a five-year period requiring at least 750 non-missing observations.

6. **Bid-ask spread:** Monthly average of daily bid-ask spread divided by average of daily spread.

7. **Book-to-market:** Book value of equity (ceq) divided by end of fiscal-year-end market capitalization.

8. **Ind.-adj. book-to-market:** Industry adjusted book-to-market ratio.

9. **C2D:** Cash flow to price is the ratio of income and extraordinary items (IB) and depreciation and amortization (dp) to total liabilities (LT).

10. **Cash Flow to debt:** Earnings before depreciation and extraordinary items (ib+dp) divided by avg. total liabilities (lt).

11. **Cash productivity:** Fiscal year end market capitalization plus long term debt (dltt) minus total assets (at) divided by cash and equivalents (che).

12. **Asset Turnover:** Asset Turnover (t) = Sales (t)/ ((Net Operating Assets (t) + Net Operating Assets (t-1)) / 2) Net Operating Assets = Receivables + Total Inventory + Other Current Assets + PP&E + Intangibles - Payables - Other Current Liabilities - Other Liabilities. Updated annually. I exclude stocks with a price below $5.

13. **Ind.-adj. Chg. in asset turnover:** 2-digit SIC - fiscal-year mean adjusted change in sales (sale) divided by average total assets (at).

14. **Chg. in shares outstanding:** Annual percent change in shares outstanding (csho).

15. **Ind.-adj. Chg. in employees:** Industry-adjusted change in Number of employees.

16. **Chg. in inventory:** Change in inventory (inv) scaled by average total assets (at).

17. **ΔSO:** Log change in the split adjusted shares outstanding as in Fama and French (2008). Split adjusted shares outstanding are the product of Compustat shares outstanding (CSHO) and the adjustment factor (AJEX).

18. **Chg. in 6-month momentum:** Cumulative returns from months t-6 to t-1 minus months t-12 to t-7.

19. **Chg. in Net Working Capital:** Yearly change in net working capital scaled by total assets. Net working capital is measured as current assets minus current liabilities. Current assets are measured as total current assets minus cash and cash equivalents. Current liabilities are measured as total current liabilities minus debt in current liabilities.

20. **Chg. in Profit:** Profit Margin (t) - Profit Margin (t-1). Updated annually. Data from year t are used to forecast returns for 12 months beginning in April of year t+1. I exclude stocks with a price below $5.

21. **Ind.-adj. Chg. in profit margin:** 2-digit SIC - fiscal-year mean adjusted change in income before extraordinary items (ib) divided by sales (sale).

22. **CTO:** We follow Haugen and Baker (1996) and define capital turnover as ratio of net sales (SALE) to lagged total assets (AT).

23. **Current ratio:** Current assets divided by current liabilities.

24. **Beta daily:** Beta daily is the sum of the regression coefficients of daily excess returns on the market excess return and one lag of the market excess return as in Lewellen and Nagel (2006).

25. **Debt2P:** Debt to price is the ratio of long-term debt (DLTT) and debt in current liabilities (DLC) to the market capitalization as of December t-1 as in Litzenberger and Ramaswamy (1979). Market capitalization is the product of shares outstanding (SHROUT) and price (PRC).

26. **Depreciation divided PP&E:** Depreciation divided by PP&E.

27. **Dollar trading volume:** Natural log of trading volume times price per share from month t-2.

28. **Dividend to price:** Total dividends (dvt) divided by market capitalization at fiscal year-end.

29. **Profit Margin:** EBIT / Revenues. Updated annually. I exclude stocks with a price below $5.

30. **Growth in common equity:** Annual percent change in book value of equity (ceq).

31. **Earnings to price ratio:** Annual income before extraordinary items (ib) divided by end of fiscal year market cap.

32. **Earnings per share:** We follow Basu (1977) and define earnings per share as the ratio of income before extraordinary items (IB) to shares outstanding (SHROUT) as of December t-1..

33. **Earnings Consistency:** Geometric average of Earnings Growth from t-1 to t-5. Earnings growth is: Earnings Per Share (t) – Earnings Per Share (t-1) / ((Absolute Value of Earnings Per Share (t-1) + Absolute Value of Earnings Per Share (t-2) / 2). I exclude stocks with a price below $5 and an absolute value of growth greater than 6 if growth is positive this year, but negative last year, or vice versa. Updated annually.

34. **Gross profitability:** Revenues (revt) minus cost of goods sold (cogs) divided by lagged total assets (at).

35. **Industry sales concentration:** 2-digit SIC - fiscal-year sales concentration (sum of squared percent of sales in industry for each company).

36. **Employee growth rate:** Percent change in Number of employees (emp).

37. **Illiquidity:** Average of daily (absolute return / dollar volume).

38. **Industry Momentum:** Value-weighted return from t-6 to t-1 within each industry. Industry is measured with two-digit SIC code. Updated monthly.

39. **Industry Relative Reversals:** In each month, firms are sorted based on the difference between their prior month's return and the prior month's return of their industry based on the Fama and French 49 industries.

40. **CAPEX and inventory:** Annual change in gross property, plant, and equipment (ppegt) + annual change in inventories (invt) all scaled by lagged total assets (at).

41. **IPM:** We define pre-tax profit margin as ratio of pre-tax income (PI) to sales (SALE).

42. **LDP:** We follow Litzenberger and Ramaswamy (1979) and define the dividend-price ratio as annual dividends over last months price (PRC). We measure annual dividends as the sum of monthly dividends over the last 12 months. Monthly dividends are the scaled difference between returns including dividends (RET) and returns excluding dividends (RETX).

43. **Leverage:** Total liabilities (lt) divided by fiscal year end market capitalization.

44. **Growth in long-term debt:** Annual percent change in total liabilities (lt).

45. **Maximum daily return:** Maximum daily return from returns during calendar month t-1.

46. **r12-2 :** We define momentum as cumulative return from 12 months before the return predifftion to two months before as in Fama and French (1996).

47. **Momentum-Reversal:** Buy and hold returns from t-18 to t-13. Updated monthly.

48. **36-month momentum:** Cumulative returns from months t-36 to t-13.

49. **Momentum-Volume:** Buy- and- hold returns from t-6 through t-1. We limit the sample to high trading volume stocks, i.e., stocks in the highest quintile of average monthly trading volume measured over the past six months. NYSE and AMEX only. Updated monthly.

50. **Lagged Momentum:** Buy-and-hold returns from t-13 through t-8. Updated monthly.

51. **Size:** Natural log of market capitalization at end of month t-1.

52. **Ind.-adj. size:** 2-digit SIC industry-adjusted fiscal year-end market capitalization.

53. **Size:** Natural log of market capitalization at end of month t-1.

54. **LME:** Size is the total market capitalization of the previous month defined as price (PRC) times shares outstanding (SHROUT) as in Fama and French (1992).

55. **LME Ind.-adj.:** Industry-adjusted-size is the total market capitalization of the previous month defined as price (PRC) times shares outstanding (SHROUT) minus the average industry market capitalization at the Fama-French 48 industry level as in Asness et al. (2000).

56. **Ind.-adj. size:** 2-digit SIC industry-adjusted fiscal year-end market capitalization.

57. **Size:** The log of market value of equity of June. Updated monthly.

58. **Net Issuance:** Net issuance is the year-over-year percent change in adjusted shares outstanding, CFACSHR SHROUT, where CFACSHR is the monthly CRSP split adjustment factor and SHROUT is common shares outstanding. Rebalanced annually, uses the recent period.

59. **Operating profitability:** Revenue minus cost of goods sold - SG&A expense - interest expense divided by lagged common shareholders equity.

60. **Chg. in current ratio:** Percent change in currat.

61. **Chg. in gross margin - Chg. in sales:** Percent change in gross margin (sale-cogs) minus percent change in sales (sale).

62. **Chg. in quick ratio:** Percent change in quick.

63. **Chg. in sales - Chg. in A/R:** Annual percent change in sales (sale) minus annual percent change in receivables (rect).

64. **PCM:** The price-to-cost margin is the difference between net sales (SALE) and, costs of goods sold (COGS) diffidef by net sales (SALE) as in Gorodnichenko and Weber, (2016) and D'Acunto, Liu, P?ueger, and Weber (2017)

65. **PM:** The profit margin is operating income after defreciation (OIADP) over sales, _x000D_ (SALE) as in Soliman (2008)

66. **PM Ind.-adj.:** The adjusted profit margin is operating income after defreciation, (OIADP) over net sales (SALE) minus the average profit margin at the Fama-French 48, industry level as in Soliman (2008)

67. **52-Week High:** Price scaled by the highest price or bid/ask average during the last 12 months. Updated monthly.

68. **Prof:** We follow Ball, Gerakos, Linnainmaa, and Nikolaev (2015) and define, profitability as gross profitability (GP) diffidef by the book value of equity as defined, above

69. **Tobin's Q:** Tobin's Q is total assets (AT), the market value of equity (SHROUT times PRC), minus cash and short-term investments (CEQ), minus deferred taxes (TXDB) scaled by, _x000D_ total assets (AT)

70. **Quick ratio:** (current assets - inventory) / current liabilities.

71. **r6-2 :** We define r6-2 as cumulative return from 6 months before the return prediфtion, to two months before as in Jegadefsh and Titman (1993)

72. **Return volatility:** Standard deviation of daily returns from month t-1.

73. **Short-Term Reversal:** Return in month t. Updated monthly. I exclude stocks with a price below $5.

74. **ROA:** Return-on-assets is income before extraordiffary items (IB) to lagged total, assets (AT) following Balakrishnan, Bartov, and Faurel (2010)

75. **Return-on-Equity:** Net income scaled by book value of equity. Exclude if price<\$5. Updated annually.

76. **Return on invested capital:** Annual earnings before interest and taxes (ebit) minus non-operating income (nopi) divided by non-cash enterprise value (ceq+lt-che).

77. **Sales to cash:** Annual sales divided by cash and cash equivalents.

78. **Sales to receivables:** Annual sales divided by accounts receivable.

79. **SAT:** We follow Soliman (2008) and define asset turnover as the ratio of sales, _x000D_ (SALE) to total assets (AT)

80. **SAT Ind.-adj.:** We follow Soliman (2008) and define adjusted asset turnover as the ratio, of sales (SALE) to total assets (AT) minus the average asset turnover at the Fama-French, 48 industry level

81. **Sales growth:** Annual percent change in sales (sale).

82. **Sales/Price:** Total revenues divided by stock price. Updated annually.

83. **Volatility of liquidity:** Monthly standard deviation of daily dollar trading volume.

84. **Volume Variance:** Standard deviation of monthly trading volume over the last 36 months. NYSE only. Updated monthly.

85. **Volatility of liquidity:** Monthly standard deviation of daily share turnover.

86. **Tan:** We follow Hahn and Lee (2009) and define tangibility as (0.715 x total, receivables (RECT) + 0.547 x inventories (INVT) + 0.535 x property, plant and, equipment (PPENT) + cash and short-term investments (CHE)) / total assets (AT)

87. **Debt capacity/firm tangibility:** Cash holdings + 0.715 receivables + 0.547 in-

ventory + 0.535 PPE/total assets.

88. **Share turnover:** Average monthly trading volume for most recent 3 months scaled by Number of shares outstanding in current month.

89. **Industry Momentum:** In each month, the Fama and French 49 industries are sorted on their value-weighted firms in decile 10 (from the 5 winner industries) form the value-weighted portfolio. Rebalanced monthly.

90. **Zero trading days:** Turnover weighted Number of zero trading days for most recent 1 month.