QUANTITATIVE STYLE INVESTING, PORTFOLIO OPTIMIZATION, AND FACTOR MODELS

by

Robert Michael Dickson Jr

A dissertation submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in
Business Administration

Charlotte

2015

Approved by:

____

Dr. Christopher Kirby

____

Dr. Steven Clark

____

Dr. I-Hsuan Ethan Chiang

____

Dr. Craig Depken

ABSTRACT

ROBERT MICHAEL DICKSON JR. Quantitative style investing, portfolio optimization, and factor models. (Under the direction of DR. CHRISTOPHER KIRBY)

This dissertation consists of three related chapters in the field of empirical asset pricing. Broadly speaking the chapters investigate issues related to active portfolio management, stock-picking, portfolio optimization, and asset pricing model performance. Chapter 1 introduces a systematic portfolio choice solution that advances contemporary models of return predictability by implementing multivariable cross-sectional regressions of key stock characteristics. These models generate tradeable portfolios which significantly outperform common benchmarks. Chapter 2 is the first study to conduct a comprehensive portfolio analysis using individual stock data. Results show that naive diversification consistently outperforms active timing strategies and parametric portfolio choice solutions. These results add to the mounting evidence that practical implementation of portfolio theory often performs poorly out-of-sample. Chapter 3 investigates whether signals from conditional asset pricing models can be used to construct tradeable portfolios and also revisits the characteristics vs. factors debate using a larger set of factors. I conclude that factors do provide reasonable proxies for the observable characteristics. To extend my current body of work, I intend on investigating the appropriate functional form of return predictability regressions and using these results for portfolio construction. Extensions of these studies offer several avenues for future contributions to the field of empirical asset pricing.

ACKNOWLEDGMENTS

I would like to dedicate this dissertation to my family. To my mother Marianne Warren for always being a source of encouragement and optimism. To my father Mike Dickson Sr. for coining the phrase,“details make the world go round.” This quote embodies his teachings and tutelage and was repeated by me many times while completing this dissertation. To my brother Mitchell Dickson for being my closest friend and confidant. Only through their love and support was my success possible.

I would like to thank my Committee, Dr. Christopher Kirby my Committee Chair, Dr. Steven Clark, Dr. I-Hsuan Ethan Chiang, and Dr. Craig Depken. Thank you all for having patience with my endless questioning during this process and also thank you all for serving on my Dissertation Committee.

I would also like to thank all my friends, colleagues, previous Professors, and the entire Belk College of Business staff for their support over the last few years. Finally, I'd like to thank God for giving me the abilities and talents to complete my PhD.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

The three chapters of my dissertation extend the literature on the cross-sectional determinants of expected stock returns and their implications for optimal portfolios and asset pricing models.

Chapter 1 introduces a systematic portfolio choice solution that significantly beats a benchmark market portfolio by an average of 34.2% per year after transaction costs. The corresponding annual Sharpe ratio is 1.97 per year compared to 0.42, over 4.7 times the size of the benchmark for a dataset consisting of all major exchange traded stocks over the last 50 years. This portfolio solution is constructed by applying multivariable cross-sectional regressions of six key stock characteristics, to aggregate forecasting signals from multiple sources. I apply simple filtering techniques to reduce estimation and sampling error, use only information known at time t, and predict expected returns. These hypothetical portfolios are implementable in real time and complete with conservative measures of trading costs. By sorting stocks by expected returns into more extreme portfolios, i.e. 25 and 50 portfolios, I am able to further enhance performance gains over existing works. The large return spreads generated by this proposed procedure provides implications of future advances for candidate asset pricing models.

In Chapter 2 I conduct a horse race of 15 portfolio construction techniques over 8 datasets comprised of individual stocks. To create the datasets I use the sequential cross-sectional regression methodology described in Chapter 1 to predict top performing stocks. This is the first study to conduct a comprehensive portfolio analysis using individual stock data and not portfolios of stocks. Using individual stock data has a practical benefit over portfolios; that is, it more accurately replicates a fund manager's portfolio construction dilemma and therefore provides excellent guidelines for practical portfolio construction. I also conduct a robust Monte Carlo analysis that confirms that recent extensions of mean-variance optimization due to Kirby and Ostdiek (2012) are successful in curbing estimation risk and turnover. Despite these facts, my results indicate that no strategy consistently outperforms naive diversification in terms of mean excess return, Sharpe ratio, and turnover. I introduce a statistic, the time series average of the cross-sectional mean absolute deviation of risk and return, to explain why I observe these results. Data limitations and dataset characteristics contribute the most to the performance of a candidate strategy. I also propose several extensions to active timing strategies and include new characteristics in a parametric portfolio choice framework. Naive diversification continues to prevail, suggesting practical optimization techniques are inferior to naive diversification when forming portfolios of individual stocks.

In Chapter 3 I conduct a comprehensive analysis that indicates characteristic-based factor loadings provide reasonable proxies for equity fundamentals. My methodology relies on the equilibrium relationships between cross-sectional expected return regressions and time-series empirical asset pricing models. I compare the performance of portfolios formed by observed equity fundamentals and pre-formation factor loadings. To generate the factors I use new methods developed by Kirby and Cordis (2015) and for the characteristics I use those described by Fama and French (2015). My analysis shows that the portfolios formed by book-to-market factor loadings closely matches the empirical return distribution of portfolios formed by the book-to-market characteristic. However, for market equity, gross-profitability, and investment, only the top performing portfolios are well approximated by factor loadings. My results are enhanced when conducting the same analyses on portfolios of equities, meant to reduce estimation risk. Finally, a portfolio formed from aggregating the signals from multiple pre-formation factor loadings beat an equally-weighed benchmark by nearly 400 basis points a year, and earned a Sharpe ratio that was 20% larger. The implications of this work suggest that in some cases, conditional factor loadings may be useful in forecasting the cross-section of expected stock returns.

CHAPTER 2: QUANTITATIVE STYLE INVESTING

Introduction

I introduce a systematic portfolio choice solution that focuses on two key elements: (1) exploiting the cross-section of expected stock returns and (2) creating equally-weighted portfolios from a subset of the entire universe of stocks. Drawing from the recent works of Novy-Marx (2013) and Fama and French (2015), I use month by month multivariable cross-sectional regressions of key stock characteristics to estimate conditional coefficient estimates. This procedure allows me to aggregate forecasting signals from multiple sources. I then apply simple filtering techniques to reduce estimation and sampling error, use only information known at time t, and predict expected returns. I form implementable portfolios by sorting stocks by these expected returns. My results indicate economically meaningful, and statistically significant returns, above and beyond a variety of benchmarks. My main contribution hinges on the result that combining an effective stock picking rule derived from predictive regressions, with naive diversification, does a lot better than a simple naive portfolio. This result is exactly what active fund managers seek to accomplish by using their "stock picking" skill to purchase undervalued stocks. Even though prediction error may be large for individual stocks, aggregating these estimates into portfolios yields large average return spreads, a ubiquitous benchmark for successful asset pricing models.

After accounting for transaction costs, the best performing equally-weighted portfolio beats a benchmark market portfolio of all stocks by 34.2% per year, with only a modest increase in volatility. The corresponding annual Sharpe ratio is 1.97 per year compared to 0.42, over 4.7 times the size of the benchmark. Looking at a more conservative sample that excludes micro cap stocks still yields annual returns after transaction costs that are 11.46% larger than the benchmark, and a Sharpe ratio 2.18 times the benchmark. I examine extreme values and tail risks and find that my portfolios outperform the market even in the worst of times. I also examine the incremental predictive power of the explanatory variables and find that the past return variables have the largest incremental impact, while size actually makes the expected return forecasts worse. At face value these results almost appear too good to be true. To appease the skeptical reader I provide comparisons with similar studies and show my procedure yields nearly identical results. I am able to increase performance gains over prior studies by adding a short-term reversal variable, and looking at more extreme portfolios, the top and bottom 2% and 4%; instead of the standard decile and quantile sorts commonly found in the literature. While it is quite common in the literature to create trading strategies from just one or two characteristics, *e.g.* size and value, there is a gap in the literature that combines multiple styles and characteristics into a composite trading strategy in the presence of transaction costs. I fill this gap and show that a composite trading strategy, using only an information set available at time t, is feasible and performs very well.

This study is most closely related to Haugen and Nardin (1996), Hanna and Ready (2005), Lewellen (2014), and Han, Zhou, and Zhu (2015). These authors all combine forecasting regressions with portfolio sorts as I do. Novy-Marx and Velikov (2014) studies many anomalies and their associated turnover using the Fama and French sorting procedure. While my methodology is much different than the Novy-Marx and Velikov (2014) study, I too consider trading costs as large limits to arbitrage. This paper differs from these works in key ways and thus my contributions are clear. First, Haugen and Nardin (1996) and Hanna and Ready (2005) consider portfolio turnover but only measure this indirectly through simulated data: I measure turnover in the portfolios directly, assessing performance ex-post. Lewellen (2014) comments that his strategies are of low turnover but fails to provide any measurements or estimates. Han et al. (2015) also ignore transaction costs associated with implementing their trend factor portfolios. Second, all five of these works only consider zero-cost hedged decile or quintile portfolios, and fail to examine the profitability of more extreme quantile sorts. My results suggest that expected return increases even further, and volatility declines in the extremes: making these portfolios more optimal investment choices. I also consider the performance of portfolios for short-sale constrained investors, providing a more complete picture of the profitability and feasibility of my strategies. Third, my model incorporates only theoretically motivated explanatory variables that have proven success in asset pricing applications, mitigating data snooping concerns. Fourth, I provide a more complete set of performance statistics by incorporating tail risk, drawdown measures, and statistical inference on the portfolio's performance relative to various benchmarks. Finally, I provide a complete analysis on the incremental contribution of each of the explanatory variables consistent with Fama and French (2014). This is a necessary piece of the analysis as my results indicate that size actually makes my forecasts worse. Armed with this new information I question the reliability of these studies that contain large numbers of explanatory variables.

Literature Review

I rely on the large literature of firm characteristics that are used to predict the cross-section of expected stock returns. Haugen and Nardin (1996) was one of the first studies to aggregate the large anomaly literature into a feasible predictive model for the cross-section of expected stock returns. As their model was largely successful, beating their benchmark by 15% in the top deciles, their results provided a direct violation of semi-strong form market efficiency and was of large interest to the investment community. Some authors such as Fama and French (1992, 1993) argued that return spreads from stock characteristics are expected and required by investors, while others argue that return spreads are a surprise and thus unexpected, e.g. the short-term reversal of the Jegadessh (1990). The literature on anomalies is large and the predictive model employed by Haugen and Nardin (1996) consisted of over fifty variables, all well documented and studied. Thanks to the size and value factors of Fama and French, there appeared to be some order introduced into the huge anomaly literature around this same time. Cochrane (2011) stated in his AFA presidential address that following this work of Fama and French, the anomaly literature was once again "descending into chaos." More recently, Subrahmanyam (2010) documented over 50 new anomalies, Hou, Xue, and Zhang (2012) tested over 80 new anomalies, and Harvey, Liu, and Zhu (2013) document over 314 different anomalous predictive variables. The methodologies from these early works were also quite different. Of course Fama and French (1993) popularized the univariate sorting procedure while Haugen and Nardin (1996) constructed multivariable cross-sectional forecasting regressions. Fama and French (2008) discuss advantages and disadvantages to both methodologies but it is quite obvious that they preferred the sorting procedures in much of their earlier work. However, even Fama and French have begun to entertain the parsimony of cross-sectional forecasting more recently.

Fama and French (2006) renewed this procedure to test for profitability and investment effects in expected returns. In Fama and French (2008) they again used this procedure to jointly analyze multiple anomalies. Clarke (2014a), Cochrane (2011), and Fama and French (2008) all discuss the awkwardness of sorts using multiple anomaly variables and point towards the rather obvious simplicity of using cross-sectional regressions. Most importantly the regression slopes provide direct estimates of marginal effects and as Cochrane (2011) states, they are “really the same thing” as univariate cross-sectional regressions. He goes on to suggest that in the “zoo of new variables” everyone will end up running multivariable regressions. Well, looking closely at the literature this is exactly what has happened. These regressions in the spirit of Fama and Macbeth (1973) have been used recently by the following authors: Fama and French (2008, 2006) with various anomaly variables, Clarke (2014a) with the same anomaly variables as Fama and French (2008), Fama and French (2014) with size, value, and momentum, Lewellen (2014) with over fifteen firm characteristics, and Han et al. (2015) synthesizing the information of short-, intermediate-, and long-term price trends.

Fama and French (2014) make an astute observation picking on the work of Lewellen (2014); a variable’s incremental contribution in the average return spread is really what matters. They point to the results in Lewellen (2014) that a model of size, book-to-market, and momentum creates a return spread about as large as their model with over fifteen forecasting variables. Their explanation goes back to the basics, that is, a new explanatory variable often attenuates the slopes of variables already in the regression. Their observation is obvious when we think about the two-step alternative to estimate a regression slope from a multivariable regression. First extract the orthogonal components of your explanatory variable relative to all other explanatory variables, and then run a single variable regression using this orthogonal construct. This same principal is applied when estimating partial autocorrelation coefficients, i.e. controlling for the other variables. In this context we are just removing variation in the explanatory variables and running univariate regressions. In the presence of many explanatory variables, especially if collinearity is an issue, attenuation will follow. So when do have we have enough variables? Should a correctly specified forecasting equation contain three variables, fifteen variables, or even fifty variables as in Haugen and Nardin (1996)? To answer these questions I use the cross-sectional counterparts of the most recent innovations in the asset pricing literature and pull from Novy-Marx (2013) and Fama and French (2015).

Novy-Marx (2013) found that profitability, measured by gross profits-to-assets, has roughly the same explanatory power in predicting the cross-section of expected stock returns as the conventional book-to-market ratio. Additionally, controlling for profitability drastically improves the performance of value strategies especially among the largest size quantiles. By combining this new measure with traditional value strategies means investors can screen quality stocks at undervalued prices. Adding this variable to the Fama and French (1993) benchmark explains many profitable trading strategies. He also tests his model in the presence of short-term reversal and momentum measures. Additionally Han et al. (2015) and Schwet (2003) state that these anomalies are some of the most "robust and persistent," and for these reasons I also include these variables in the forecasting regressions. Fama and French (2015) use a slight variant of the Novy-Marx (2013) profitability measure, in addition to their newly introduced investment factor, to create a five-factor model. The main difference between their new model and their three-factor model lies in the theoretical starting point. The three-factor model is an empirical asset pricing model seen as an application of the Ross (1976) APT model. However in Fama and French (2015) they begin instead with the dividend discount model of Modigliani and Merton (1958), and relate their five factors to state variables of expected stock returns.

While I borrow from the aforementioned work to support my data and methodologies, my contributions are most similar to Haugen and Nardin (1996), Hanna and Ready (2005), Lewellen (2014), and and Han et al. (2015). Specifically each of these authors combine variables thought to explain the cross-section of expected stock returns into forecasting regressions, and then form portfolios based on sorts of the expected returns. Novy-Marx and Velikov (2014) study many anomalies and their associated turnover using the Fama and French sorting procedure, and as expected, find transaction costs significantly reduce the strategies' profitability. He further provides evidence that the statistical significance of the results are also reduced, increasing concerns related to data snooping. This result is troublesome for the forecasting regressions of Haugen and Nardin (1996) and Lewellen (2014) which consists of a large number explanatory variables. While Haugen and Nardin (1996) and Hanna and Ready (2005) consider turnover they only do so with simulated data. Lewellen (2014) comments that his strategies are of low turnover but fails to provide any measurements or estimates. All five of these related works only consider decile or quintile sorted portfolios and fail to consider the profitability of more extreme sorts such as the top and bottom 2% and 4%.

Empirical Application

Data

My sample spans July 1963 to Dec 2013 with monthly holding period returns obtained from the Center for Research in Security Prices (CRSP) and accounting data from Compustat. The sample includes common equity securities (share codes 10 and 11) for all firms traded on the NYSE, NASDAQ or AMEX (exchange codes 1, 2, and 3). I consider the following stock characteristics consistent with Novy-Marx (2013) and Fama and French (2015): size (log(ME)), book-to-market (log(BE/ME))[1], profitability (ratio of gross profits to assets), past performance measured at horizons of one month ($r_{1,0}$) to capture short-term reversals, and 12 to two months ($r_{12,2}$), to capture momentum, and investment (growth of total assets from previous fiscal year). I provide exact definition for these variables and their construction in the appendix. As in Novy-Marx (2013), to reduce the effect of outliers, I trim all independent variables at the 1% and 99% levels. My analysis consists of two sets of data denoted the "Full Sample" and the "No Micro Sample". Fama and French (2008) define microcaps as stocks with a market value of equity below the 20th percentile of the NYSE market capitalization distribution. Microcaps make up about one half of the stocks on NYSE, AMEX, and NASDAQ, but account for only about 3% of the total market cap. As they note, these small stocks may be less liquid than the representative sample and thus result in above average transaction fees. I estimate all models using the full sample without the microcap stocks to examine the extent to which the performance gains are driven by microcap stocks.

[1]Taking logs makes the cross-sectional distribution of market equity and book-to-market more symmetric, reducing the impact of outliers

Figure 1 describes these six firm characteristics. The first column plots the cross-sectional means of these characteristics and the second column plots the cross-sectional standard deviations. The solid blue line shows the statistics for the full sample and the dotted green line shows these statistics for the no microcap sample. For the level variables log market equity (me), log book-to-market equity (btm), and gross profitability(Prof), the plots show distinct differences in the firm level variables between the two samples. However for the level variable investment (Inv), and the flow variables short-term reversal (Str) and momentum (Mom), there is little distinction. For a few of the characteristics I can draw some conclusive observations. Specifically the market equity of the average firm has been trending upwards over time in both samples. I also notice that the average cross-sectional profitability has been falling over time while the volatility of profitability has been increasing over time.

Figure 2 plots the number of firms in the best performing reported portfolios, per month, over the out-of-sample testing period July 1968 to December 2013. These portfolios consists of a zero-cost hedge portfolio with a long position in the top portfolio as sorted by expected return, and a short position in the bottom portfolio as sorted by expected return. This figure plots the number of firms when expected returns are sorted into 50 and 25 portfolios (top and bottom 2% and 4%). These firm counts are just scaled down versions of the number of firms and I can see an obvious upward trend in both samples. The average annual growth rates are 0.24% and 0.18% for the full sample and no micro cap sample respectively. For the full sample the average number of firms is 3350, with the fewest firms in November 1962 (805) and the most firms in July 1997 (5657). In the no micro cap sample the average number of firms is 1483, with the fewest firms in September 1962 (564) and the most firms in November 2010 (2258). The average number of firms for the no micro cap sample is a little less than half (44.2%) of the number of firms in the full sample. This number shows the disproportionate number of micro cap firms. As noted by Fama and French (2008), on average microcaps are 60% of all sample stocks while only accounting for about 3% of the market cap.

Generating the Expected Returns

To generate the expected returns I begin with the following specification for the month by month cross sectional regressions:

$$\begin{aligned} r_{i,t+1} = \alpha + \beta_1 ln(BE/ME)_{i,t} + \beta_2 ln(ME)_{i,t} + \beta_3 GPdat_{i,t} + \beta_4 R1to0_{i,t} \\ + \beta_5 R12to2_{i,t} + \beta_6 INV_{i,t} + \epsilon_{i,t} \end{aligned} \tag{1}$$

I estimate this model for each of the 606 months from July 1963 through December 2013. Since the conditional coefficient estimates from month to month are quite noisy, I apply a simple rolling average to the coefficients to filter out the signal. As long as the regression coefficients are relatively stable over time, this method should significantly reduce estimation and sampling error. Rolling estimators are quite common in the literature on portfolio selection and a number of studies use a fixed-width rolling data window to estimate the mean vector and covariance matrix of asset returns in a mean-variance framework. For example Demiguel, Garlappi, and Uppal (2009), Kirby and Ostdiek (2012, 2015), and Tu and Zhou (2011) all apply this approach. While it does seem natural to use a fixed-width window, this is typically less efficient than methods that exploit the full history of available asset returns. R. Merton (1980) and Foster and Nelson (1996) provide evidence to this effect and promote the use of the full history of available returns. I apply this same logic to the rolling average estimates of the cross-sectional regression coefficients and use the full available history when computing these averages.[2] I initialize the rolling averages with 60 months of data and use the estimated mean parameter estimates to generate the expected returns. Specifically the smoothed coefficient estimates are generated as:

$$\hat{\beta}_t = \frac{1}{t-1}\sum_{i=1}^{t-1}\beta_i$$

Since I begin with a sixty month burn-in period, the first β estimate occurs at $t = 61$. Note that the mean parameter estimates for month t are based on estimates before month t, which ensures that I have an implementable trading strategy. This means that I make the investment decisions at time t based solely on information available and derived from information before time t. Thus I use the parameter estimates for July 1963 - June 1968 to generate the first out-of-sample return in July 1968. Once I have these smoothed coefficient estimates, I generate the expected returns at time t for each stock, by applying these coefficients to each firm's observed characteristic at time t. These are the expected returns I use in my sorts. To curb excessive turnover and transaction costs, I update the smoothed coefficients annually, specifically every June.[3] Denote these estimates as $\hat{\beta}_s$ for "smoothed," and the estimated expected returns are computed as follows:

$$\hat{r}_{i,t+1} = \hat{\alpha}_s + \hat{\beta}_{1,s} ln(BE/ME)_{i,t} + \hat{\beta}_{2,s} ln(ME)_{i,t} + \hat{\beta}_{3,s} GPdat_{i,t} + \hat{\beta}_{4,s} R1to0_{i,t}$$
$$+\hat{\beta}_{5,s} R12to2_{i,t} + \hat{\beta}_{6,s} INV_{i,t} \qquad \text{(Model 1)}$$

I refer to this specification as All Variables which represents the forecasting regression

[2] I experimented with various rolling estimates such 12 month, 60 month, and 120 month and found all to be inferior to using the entire sample.

[3] I found this method to be superior to updating on a monthly basis

using all variables. The motivation for this model is derived from the recent asset pricing works of Novy-Marx (2013) and Fama and French (2015) and consists of both slow moving level variables (Me, BeMe, Prof, and Inv) and short-lived predictors such as short-term reversal and momentum. These short-lived predictors vary considerably, and my analyses shows that they contribute to a higher level of portfolio turnover. I present results net of transaction costs, which allows direct comparisons between portfolios of varying levels of turnover. However, I also construct a separate model, Model 2, that consists only of the slow-moving level variables to appeal to turnover sensitive investors. To construct expected returns I estimate the cross-sectional regression of Eq. 1. As in All Variables I apply the smoothing filter in Eq. 2 for each of the βs, but only use $R1to0_{i,t}$ and $R12to2_{i,t}$ as controls and thus β_4 and β_5 were not used to generate expected returns. Therefore the expected returns are computed as:

$$\begin{aligned}\hat{r}_{i,t+1} = \hat{\alpha}_s + \hat{\beta}_{1,s} ln(BE/ME)_{i,t} + \hat{\beta}_{2,s} ln(ME)_{i,t} + \hat{\beta}_{3,s} GPdat_{i,t} \\ + \hat{\beta}_{6,s} INV_{i,t}\end{aligned} \tag{Model 2}$$

I also use this model to facilitate comparisons with Lewellen (2014) who uses similar variables. To facilitate comparisons with Clarke (2014a) I create Model 3. For this model $R1to0_{i,t}$ was used only as a control and thus β_4 was not used to generate expected returns:

$$\begin{aligned}\hat{r}_{i,t+1} = \hat{\alpha}_s + \hat{\beta}_{1,s} ln(BE/ME)_{i,t} + \hat{\beta}_{2,s} ln(ME)_{i,t} + \hat{\beta}_{3,s} GPdat_{i,t} \\ + \hat{\beta}_{5,s} R12to2_{i,t} + \hat{\beta}_{6,s} INV_{i,t}\end{aligned} \tag{Model 3}$$

Measuring a Variable's Incremental Contribution

In-sample incremental contribution

To measure the in-sample incremental contribution of each explanatory variable I perform R^2 decomposition in Fama and Macbeth (1973) regressions. Lewellen (2014) makes some informative comments regarding the interpretation of the R^2 from Fama-Macbeth regressions. He says that it would be wrong to interpret the R^2 as informative about the predictive power of that variable. The R^2 only provides information about the fraction of contemporaneous volatility, and nothing about the predictive ability. Therefore, using an R^2 decomposition tells me the fraction of the model's contemporaneous volatility, that is explained by each of explanatory variables in the presence of all the explanatory variables. A variable with a high (low) $\%$ of R^2 indicates that variable explains a large (small) fraction of the model's contemporaneous volatility. The Appendix provides the details and proof of the R^2 decomposition.

Out-of-sample incremental contribution

Referring back to the comments of Fama and French (2014), a variable's incremental contribution in the average return spread, in the presence of other variables, is really what matters. All the variables used in my analysis have been shown to have predictive content in Fama and Macbeth (1973) regressions from previous studies. I also present Fama and Macbeth (1973) regression statistics in Table 5: but how much out-of-sample predictive power do each of the variables contain? To study each variable's incremental contribution I use an approach similar to Clarke (2014a). I start with the full estimation model, Eq. 1, and the full expected return model, All Variables, and systematically drop each explanatory variable from both. If an explanatory variable has a large incremental contribution to the out-of-sample prediction accuracy, then the return spread and other performance statistics should drop substantially. If an explanatory variable has little incremental contribution, these statistics should remain relatively unchanged. I present the economic significance of the explanatory variables in Tables 8 and 9.

Portfolio Turnover and Trading Costs

Portfolio turnover is an often overlooked but very real cost to investors. Transactional brokerage fee costs are typically not included in the calculation of a fund's operating expense ratio and thus the true operating expense of high turnover funds can be significant. As long as transaction costs are greater than zero, anything that increases turnover directly reduces the true performance of a fund. To examine the amount of trading required to implement each strategy I follow Kirby and Ostdiek (2015). Turnover is simply the fraction of invested wealth traded each period needed to re-balance the portfolio to the desired weights. At any time t I calculate turnover as:

$$Turnover_t = \sum_{i=1}^{N} \frac{1}{2} |\hat{w}_{i,t+1} - \hat{w}_{i,t^+}| \tag{2}$$

This definition of turnover is consistent with what is used in the mutual fund industry, i.e. the lesser of the value of purchases or sales in the period divided by the net asset value. (Kirby & Ostdiek, 2015) Since there are no fund inflows or outflows these must be equal. I define $\hat{w}_{i,t}$ as the portfolio weight in asset i at time t; $\hat{w}_{i,t^+}$ is the portfolio weight before re-balancing at time $t+1$; and $\hat{w}_{i,t+1}$ is the desired portfolio weight at time $t+1$, after re-balancing. To compute $\hat{w}_{i,t^+}$ I must consider the mechanical changes that occur within the portfolio. Assets that have done well over the time period will make up more than their starting share of weight at the end of the period, and assets that have done poorly will make up less than there starting share. I compute $\hat{w}_{i,t^+}$ as:

$$\hat{w}_{i,t^+} = \frac{\hat{w}_{i,t}(1 + r_{i,t})}{1 + \sum_{i=1}^{N} \hat{w}_{i,t} r_{i,t}} \tag{3}$$

Starting from the beginning of my sample, the first weights occur in month 61, therefore the first turnover calculation occurs in month 62. Studies such as Kirby and Ostdiek (2015, 2012) and Demiguel et al. (2009) do not ignore these mechanical weights while others such as Brandt, Santa-Clara, and Valkanov (2009) do ignore these mechanical changes. I have found that ignoring these mechanical changes is innocuous in this setting but do include them to capture the most conservative view of the trading costs. Now the return of the portfolio net of the proportional transactions costs becomes:

$$r_{p,t+1} = \sum_{i=1}^{N} \hat{w}_{i,t} r_{i,t+1} - 2 \times c_{i,t} |\hat{w}_{i,t} - \hat{w}_{i,t-1^+}|, \tag{4}$$

where $c_{i,t}$ reflects the proportional transaction cost for stock i and time t. Since turnover is the value of assets both purchased and sold as a fraction of total wealth, and both purchases and sales incur transaction costs, I multiply the turnover in Eq. 4 by 2. Novy-Marx and Velikov (2014) find that a momentum based trading strategy had one of the largest time-series average costs of trading in his rigorous analysis of the trading costs of over twenty common anomalies. These costs were estimated at 48.39 basis points per month. While it has been noted by Domowitz, Glen, and Madhavan (2001) and Hasbrouck (2009) that the cost of trading U.S. equities has declined over time, I want to be as conservative as possible in accounting for these limits to arbitrage. To do so I set $c = 50$ basis points consistent with the conservative measures used by Brandt et al. (2009), Demiguel et al. (2009), Kirby and Ostdiek (2012, 2015), and the even more recent estimates by Novy-Marx and Velikov (2014).

Finally, letting L reference the burn-in-period of 60 months, and T represent the total number of months in my study, numerically the average turnover I report in the tables is:

$$Turnover = \frac{1}{T-L-1} \sum_{t=L+1}^{T-1} \left(\frac{1}{2} \sum_{i=1}^{N} |\hat{w}_{i,t+1} - \hat{w}_{i,t^+}| \right) \tag{5}$$

Statistical Inference

To conduct statistical inferences about the relative performance of the various strategies using the Sharpe ratio, I follow Kirby and Ostdiek (2012) and use large sample t and $chi-squared$ statistics. I consistently compute these statistics using the generalized method of moments (GMM). For details of the proof of the general results see Hansen (1982). As Hansen (1982) shows, the Delta method, Slutky's theorem and LLN are all used to derive the asymptotic distribution of the GMM estimators. Recent asymptotic distribution derivations for Sharpe ratios are also provided by Opdyke (2007) and Bailey and de Prado (2011) who also use these theorems in their derivations. However I use GMM standard errors to appeal to these more recent derivations while still applicable in a more general context. In the analysis I begin with a set of moment conditions of the form $E(g(R_t, \theta)) = 0$, where $g(R_t, \theta)$ is a $J \times 1$ vector of moments, analogous to disturbances, R_t is a vector or returns, and θ is $J \times 1$ vector of parameters. I use the fundamental result from Hansen (1982) that, subject to general conditions, the limiting distribution of $\hat{\theta}$ is given by:

$$\sqrt{T}(\hat{\theta} - \theta) \xrightarrow{d} N(0, V) \tag{6}$$

I have the following:

$$V = D^{-1}SD^{-1'},\ D = E(\partial g(R_t, \theta)/\theta'),\ S = \sum_{-\infty}^{\infty} E(g(Y_t, \theta)g(Y_{t-j}, \theta)')$$

The moment conditions are specified as follows:

$$g(R_t, \theta) = \begin{pmatrix} R_{bench,t} - \sigma_{bench} \times SR_{bench} \\ R_{test,t} - \sigma_{test} \times SR_{test} \\ (R_{bench,t} - \sigma_{bench} \times SR_{bench})^2 - (\sigma_{bench})^2 \\ (R_{test,t} - \sigma_{test} \times SR_{test})^2 - (\sigma_{test})^2 \end{pmatrix} \tag{7}$$

Using Eq. 6 I have the asymptotic standard errors of the Sharpe ratios and can also easily conduct a Wald test of linear restrictions to determine if the differences between the Sharpe ratios of the benchmark and test portfolios are statistically different from zero. To do so I consider the following test statistic:

$$(\hat{SR}_{test} - \hat{SR}_{bench})(R_{SR}VR'_{SR})^{-1}(\hat{SR}_{test} - \hat{SR}_{bench}) \sim \mathcal{X}(1) \tag{8}$$

In Eq. 8 the discrepancy vector $R_{SR} = (-1, 1, 0, 0)$ and V is the asymptotic covariance matrix described in Eq. 6. It can be shown that the square root of this statistic is equivalent to the following limiting distribution:

$$\sqrt{T}((\hat{SR}_{test} - \hat{SR}_{bench}) - (SR_{test} - SR_{bench}) \xrightarrow{d} N(0, R_{SR}VR'_{SR}) \tag{9}$$

So in the case where the population Sharpe ratios are equal in the benchmark and test portfolios, I have the following large-sample test statistic:

$$\sqrt{T}\left(\frac{\hat{SR}_{test} - \hat{SR}_{bench}}{(R_{SR}\hat{V}R'_{SR})^{1/2}}\right) \overset{\text{asy}}{\sim} N(0,1) \tag{10}$$

I also use this asymptotic covariance matrix to conduct a simple $t - test$ on the individual significance of the Sharpe ratio. Of course I no longer need a discrepancy vector and can simply take the square root of the appropriate diagonal element of the matrix V to compute the test statistic.

Results

Return Spreads

To generate a profitable stock-picking strategy, what I am really after is a spread in expected returns. This is the same idea popularized by Fama and French: sort assets into portfolios based on characteristics, look at high minus low mean returns, and then test if this spread in means corresponds to a spread in betas. As Cochrane (2011) points out, in the "zoo of new variables," multivariate regressions provide an excellent alternative to the awkwardness of portfolio sorts across multiple characteristics. One of my contributions to the literature is the investigation of more extreme sorts in an implementable trading strategy. Haugen and Nardin (1996), Hanna and Ready (2005), Lewellen (2014), Novy-Marx and Velikov (2014), and Han et al. (2015) all consider only decile or quintile portfolios in their analyses.

Table 1 shows the mean return spread and Sharpe ratio of my three models for groups of 2, 5, 10, 25, 50, and 100 portfolios[4]. This table serves a guide to picking the optimal

[4]For hedge portfolios I focus on the top and bottom portfolios; 2 portfolios correspond to 50%, 5 to 20%, 10 to 10%, 25 to 4%, 50 to 2%, and 100 to 1%.

size of the portfolios. When sorting into only two portfolios, each portfolio will have a larger number of securities, it will be highly diversified, and should have low levels of idiosyncratic volatility. When sorting into 100 portfolios, each portfolio will have a much smaller number of securities, will be less diversified, and thus will have more idiosyncratic volatility. By analyzing both the return spreads and the Sharpe ratios I can get an idea on the appropriate size of a portfolio such that the marginal return to accepting more idiosyncratic volatility is optimal. For all models in Table 1 the return spread is monotonically increasing with the number of portfolios. This means my models are effective in generating a large and consistent return spread. For my main model, Model 1, the spread in returns is as high as 4.83% per month for 100 portfolios and 1.09% for 2 portfolios for the full sample. I also observe that the return spreads for the no micro cap sample are smaller in all cases. For 100 portfolios using Model 1 the return spread is 2.53%, nearly half of the full sample. This shows that there is significant return spread in the micro cap stock universe. While 100 portfolios produces a large spread in return, all models do not show that 100 portfolios produce the largest Sharpe ratio. In fact, the largest Sharpe ratio for the full sample is 0.64 with 50 portfolios using Model 1, and 0.35 for the no micro cap sample with 50 portfolios using Model 1. I also observe, in nearly all cases, that 25 portfolios produce a larger Sharpe ratio than 100 portfolios. Additionally both 25 and 50 portfolios always yield larger spreads and Sharpe ratios than decile portfolios, providing evidence that these more extreme sorts provide more desirable investing opportunities. Based on these results, I focus on 25 and 50 portfolios when presenting the detailed performance statistics in Tables 2, 3, and 4.

Figure 3 plots the mean and volatility of actual returns for each portfolio using Model 1 for the full sample and no micro cap sample, and Model 2 with the full sample. For brevity I present results only for 25 portfolios. The figure shows an almost perfectly monotonically decreasing mean actual return from the top portfolio as sorted by expected return to the bottom portfolio. These results support the predictive quality of my model. While each of these portfolios was created by a sort on expected returns, the actual returns match very closely. The volatility plots show an interesting U-shaped pattern. While I do expect actual returns and volatility to be positively related based on the fundamental risk requires return relationship, I do not necessarily expect the lowest expected return portfolio to have volatility nearly as high as the top portfolio. However, I can reconcile this with an argument based on asymmetric volatility. That is, as prices for these securities decreased, their volatility increased, giving way to the U-shape in the figure.

To facilitate comparisons with prior literature and validate the results of my procedure, I look at Model 3, which does not contain the short-term reversal variable.[5] I compare the results to Clarke (2014b), Table VI, whose model and procedure closely match my Model 3.[6] For 100 portfolios Clarke's model generates a spread of 2.62% vs. my spread of 2.77% and a Sharpe ratio of 0.33 vs my Sharpe ratio 0.34. For 10 portfolios Clarke's model generates a spread of 1.65% vs my spread of 1.84% and a Sharpe ratio of 0.38 vs. my 0.28. Lewellen (2014) presents results for a model of 15 explanatory variables and finds an equal-weighted spread of 2.36% for decile portfolios with a monthly Sharpe ratio of 0.47. These results a more similar to my Model 1 generating a spread of 2.89% and a Sharpe ratio of 0.54. Most importantly, my model generates commensurate results as these existing studies, passing the litmus test for validity.

---

[5]The STR variable is shown to have significant incremental contribution to the return spread and has not been included in prior works.

[6]He also includes net stock issues and accruals

Traditional Performance Statistics for Implementable Portfolios

Table 2 presents traditional performance statistics and highlights my main results. That is, even in the presence of transaction costs, my model yields returns and Sharpe ratios that far exceed a benchmark market portfolio.[7] An equal-weighted hedge portfolio constructed from 50 portfolios using Model 1, yields a monthly return of 3.98% after transaction costs, compared to 1.13% for the equal-weighted benchmark portfolio that includes all stocks. This equates to 34.2 % per year in excess of the market. The volatility of this portfolio is 6.29% per month compared to the benchmark 5.84%. The Sharpe ratio after transaction costs is 0.57 per month or 1.97 per year compared to 0.42 for the benchmark: over 4.7 the size of the benchmark. While the turnover for this portfolio, 0.45 per month or 5.34 per year, is quite high, the performance of the strategy more than makes up for the additional costs needed to implement the strategy. For the equal-weighted hedge portfolio constructed from 25 portfolios I see comparable performance in Sharpe ratios and about 0.5% drop per month in mean return relative to 50 portfolios. Using Model 2 my results have the desired effect of reducing turnover from 0.445 per month to 0.108 per month, which puts the strategy at just over 100% turnover for the year. This would be considered a reasonable amount of turnover in the industry. At the cost of a reduction in turnover the mean returns and Sharpe ratios also fall. After transaction costs this model still returns 2.38% per month, 15.06% per year in excess of the benchmark, and has an annualized Sharpe ratio of 1.03. I see a similar relationship as before with 25 portfolios; 0.4% per month less mean return and comparable Sharpe ratios relative to 50 portfolios. All estimated Sharpe ratios are

[7]For sake of brevity I only present results for the best performing models but have an on-line appendix with many additional robustness checks.

statistically different than zero and are statistically larger than the benchmark. One concern when viewing these results is the size of the average firm. For the hedge portfolio using 50 portfolios, the average firm size is only $371 million, and using 25 portfolios it is only $454 million. Model 2 picks larger firms with $691 million for 50 portfolios and $854 million for 25 portfolios.

The results for the no micro cap stocks are uniformly lower, but still represent substantial gains over the benchmark. In the presence of transaction costs, an equal-weighted hedge portfolio constructed from 50 portfolios yields a monthly return of 2.08% and an annual Sharpe ratio of 0.93. This corresponds to 11.46% per year in excess return over the benchmark and a Sharpe ratio 2.18 times the value of the benchmark. For 25 portfolios the mean return drops about 0.3% per month but I again return comparable Sharpe ratios as the 50 portfolio case. The average firm size for no micro cap universe is much larger, $1.13 billion and $1.28 billion for 50 and 25 portfolios respectively.

Even for a short-sale constrained investor who is sensitive to turnover, my stock selection model yields impressive results. Using Model 2 and a long only portfolio constructed as the top portfolio from 25 portfolios, the mean return is 2.08% per year with an annual Sharpe ratio 0.76 after transaction costs. The turnover is estimated at 0.10 per month. For the no micro cap sample the mean return is 1.52% per month with roughly the same turnover as the benchmark portfolio. Even in this extremely conservative example with multiple constraints, my model beats a benchmark by over 4% per year.

For the most risk-averse of investors I present results that I call a “Low Volatility” portfolio. This is constructed as going long the top 50% of stocks and short the bottom 50% of stocks estimated by Model 1. The annualized volatility of this portfolio is only 7.36% per year while it returns 13.1% per year. This portfolio essentially matches the market return with less than half the volatility. The turnover is also the same as the market as a whole which makes this an attractive alternative for risk-averse investors.

Figure 7 presents 60-month rolling Reward-to-Risk ratios and figure 6 presents 60-month rolling mean excess returns, accounting for transaction costs. These figures allow me to view the performance of the best performing strategy, over time, in both samples. The Reward-to-Risk ratio is always larger in both samples than the benchmark portfolio, but the gap does seem to be shrinking over the past 20 years. For the full sample the mean excess return is always larger than the benchmark, although shrinking as well over the last 10 years. For the no micro cap sample, I notice from about 2007 until the end of the sample, the mean excess return was slightly less than the benchmark. While outside of the scope of this paper these results do beg the question: Is anomaly arbitrage disappearing? I do not attempt to rigorously answer this question but I do offer a plausible answer. Large quantitative trading funds such as AQR and DFA actively create and trade on strategies based on published factor models and well known anomalies. It is likely that their presence in the market and the funds that they manage are responsible for the recent decrease in performance of characteristic based strategies.

How does this performance compare with existing literature?

Since prior authors conducting similar procedures have not considered the transaction fees encountered to form their portfolios, I must reference the raw performance statistics to facilitate comparisons. As previously noted, the best model presented by Clarke (2014b) had a spread of 2.62% and a Sharpe ratio of 0.33, both monthly. The best model presented by Lewellen (2014) had a spread of 2.36% for decile portfolios and a monthly Sharpe ratio of 0.47. My best model (in terms of Sharpe ratio) generates a spread of 4.43% and a Sharpe ratio of 0.64, both monthly. Lewellen (2014) also presents a table for all-but-tiny stocks which corresponds to my no micro cap sample. His best model returns a spread of 2.24% and a Sharpe ratio of 0.85 compared to a 2.52% mean return and a Sharpe ratio of 1.20 for my best model. However without statistics for portfolio turnover I have no way of knowing if these author's strategies are profitable to implement.

Factor Model Performance Evaluation

To further examine the performance of my selected portfolios I present statistics for risk-adjusted realized returns relative to the Fama-French 4-factor model in Table 3. I also present the information ratio (IRα). This statistic is computed as the alpha from the factor regression divided by the standard deviation of the residuals, also known as the tracking error or idiosynchratic volatility (IVOL). This is an idiosynchratic reward to idiosynchratic risk measure, the higher the better. Similar in spirit and interpretation as a Sharpe ratio except that the expected return from the Sharpe ratio is replaced with the expected return from a factor model, i.e. alpha, and the standard deviation of the stock price is replaced by the IVOL, i.e. the standard deviation of the residuals from the factor model. I find that these risk-adjusted measures match up quite well to the univariate statistics. My best performing model, after transaction costs, returns 36.9% in excess α over the benchmark compared to 34.2 % per year mean return in excess of the market. Similarly the IRα for my best model is 0.60 per month compared to the Sharpe ratio of 0.64 per month. The univariate performance statistics and risk-adjusted performance statistics follow a similar pattern for all other portfolios in the table.

I also present an information ratio relative to the S&P 500, ($\text{IR}_{S\&P}$). This statistic is computed as the difference between the portfolio return and the return of the S&P 500, divided by the standard deviation of this difference. It can be thought of as the Sharpe ratio of returns in excess of the S&P 500. The higher the better indicating a higher average return that is consistently larger than the S&P 500. My best performing model yields 0.51 per month and even the short-sale constrained, low turnover model, yields 0.194 per month.

Analyzing the factor loadings provides insight into the style of the stocks represented by my portfolios. The most striking observations from the table are the differences in the loadings comparing Models 1 and 2. For both samples, the exposure to the market factor decreases by about 0.5 when switching from Model 1 to Model 2. Further Comparing Model 1 to Model 2 I also see an increase on the SMB factor and HML factors. This indicates that when I remove the past performance variables, even though my model actually selects larger stocks on average, these stocks actually have more exposure to the size factor. Furthermore, Model 2 also selects more value stocks. Combining these results Model 2 tilts toward small cap value stocks and away from the market.

For the full sample hedge portfolios all momentum βs are near zero and statistically insignificant and the R^2 values are very low. This evidence points to the fact that the Fama-French 4-factor model has a difficult time reliably pricing these portfolios.

Extreme Values and Tail Risk of Implementable Portfolios

I compare the maximum drawdown, Calmar ratio, conditional value-at-risk (also known as the Expected Shortfall), and the frequency of large losses for both the benchmark and my best performing portfolios in Table 4. The maximum drawdown (MDD) is defined as the largest percentage drop in price from a peak to bottom. It measures the absolute worst case scenario of an investor and is a popular metric in the mutual funds industry. An investor would earn this return if they invested in the candidate portfolios at the worst possible time and subsequently sold the portfolios at the worst possible time. My most aggressive and best performing portfolio has an MDD of 47.05% compared to the market's 44.65%. Both my Model 2 and the no micro cap sample have MDD measures smaller than the market. My low-volatility portfolio has the most favorable MDD of 17.39%. The Calmar ratio is closely related and is defined as the annualized rate of return divided by the MDD. This is also a popular mutual fund statistic and measure return versus downside risk. The higher the better and in every model I find that all portfolios have larger Calmar ratios than the benchmark. The largest Calmar ratio is 1.13 while the smallest is 0.37 compared to compared to 0.32 for the benchmark. The conditional value-at-risk (Cvar) is often described as a more robust statistic than a simple value-at-risk measure. It defined as the average of the worst $q\%$ of returns. I choose $q = 5\%$ and find again that all of my hedge portfolios have more favorable measures of Cvar while the long-only portfolios are only slightly worse. The Cvar for my most aggressive portfolio is -7.80% compared to the market's -12.48%. These results strongly indicate that my portfolios outperform the market even during the worst of times.

Fama-Macbeth Regressions

My estimation procedure is a variant of Fama-Macbeth regressions in that I do not use contemporaneous measures of β to generate expected returns. Using predicted values from a Fama-Macbeth regression is not an implementable trading strategy because the information set includes variables that occur at time $t + 1$, specifically the expected return on the left-hand side. Nonetheless, Fama-Macbeth regressions provide a viable benchmark for the significance of each of the variables and a comparison with other studies. Table 5 reports the average slopes, t-statistics, and R^2s for 606 monthly cross-sectional regressions. I present results for the full model, and then systematically drop one of the explanatory variables to asses this dropped variables impact on the attenuation and significance of the estimated risk premiums. Several interpretations of Fama-Macbeth regressions are worth noting. First, Fama (1976) describes how Fama-Macbeth slopes can be interpreted as returns on characteristic-based portfolios. Second, Fama (1976) also describes how the Fama-Macbeth R^2 reflects how much ex post volatility is explained, and is not indicator of the predictive ability of the characteristics. Finally, if I can interpret the slopes as returns on characteristic-based portfolios, then I can directly interpret the size of the t-stat as the size of the Sharpe ratios of characteristic-based portfolios. Consider the definition of the t-statistic and the Sharpe ratio:

$$t = \frac{E(R)}{\sigma(R)/\sqrt{T}} \text{ and } SR = \frac{E(R)}{\sigma(R)} \\ SR = \frac{t}{\sqrt{T}} \tag{11}$$

As long as T is the same for all regressions, then a higher t-stat indicates a higher Sharpe ratio for the return on a characteristic-based portfolio.

My results are consistent with prior research. In both samples the slopes on book-to-market, momentum, and gross profitability are significant and positive, while the slopes on investment, short-term reversal and size are significantly negative. In fact, in my full model, the risk premium for size is less than 2 standard errors from zero, -1.72, indicating that this is not statistically different than zero at the 5% level. Examining the size of the t-statistics I obtain similar interpretations as Novy-Marx (2013) regarding the relationship between book-to-market and gross profitability. For my full model the t-stat is 6.04 for gross-profitability and 6.52 for book-to-market. This t-statistic means that the Sharpe ratios of characteristic-based portfolios sorted on these variables is roughly the same. Therefore I confirm Novy-Marx's conclusions that gross profitability has roughly the same power as book-to-market. Examining column 4 I further corroborate his claims that value and profitability complement each other. Note that when book-to-market is removed from the regressions, the risk premium for gross profitability falls from 0.68% to 0.41% and the t-stat falls from 6.04 to 3.60 in the full model. In the no micro cap sample the risk premium falls from 0.65% to 0.32% while the t-stat falls from 4.75 to 2.26. This shows that both the risk-premiums and Sharpe ratios of characteristic-based portfolios are negatively impacted when value is not controlled for. A similar result holds for the risk premium and Sharpe ratios for value portfolios, i.e book-to-market. My results also indicate a complementary but reversed relationship between value and investment. When investment is included, the risk premium for value falls in both the full and no micro sample. When value is included, the risk premium for investment falls over 0.20% in both samples as well. For all other variables, the risk-premiums and t-statistics remain relatively unchanged when variables are added or removed.

Finally I compare the estimates for the short-term reversal variable with that of Novy-Marx (2013). Both my estimates and that of Novy-Marx are consistent; risk-premiums fall between -5% and -6% with t-statistics between -12 and -14. These estimates indicate a large overreaction to last month's return that has much higher explanatory power in the cross-section than the other variables analyzed. This is the only variable in my sample in which the value is determined month by month, and thus I would expect that including this variable in my model would increase turnover. While momentum also changes monthly, it is computed as a combination of past returns and thus evolves more slowly than short-term reversal.

A key observation from these results is that all of my predictive variables have explanatory power in the cross-section.[8] This is not true for the predictive regressions of Clarke (2014a, 2014b), Lewellen (2014), Hanna and Ready (2005), and Han et al. (2015). Referring back to my earlier question: when have we included enough anomaly variables? By focusing on the recent developments in the asset pricing literature, I have selected a subset of the anomaly variables such that each has significant explanatory power in the cross-section, in the presence of all other variables. This is a claim that cannot be made by the aforementioned studies.

R^2 Decomposition: Does it mean anything?

The Fama-Macbeth R^2s are 0.046 in the full sample and 0.067 in the no micro cap sample at their largest. However, these measures are commensurate with prior studies e.g. Lewellen (2014). Table 6 presents the results from R^2 decomposition for Model 1 and Model 2 using both the full and no micro cap sample. The results are thought provoking. Can we relate the % of R^2 to the t-statistic in Fama-Macbeth regressions? At first glance it does not appear so. In column 1 both gross profitability and investment explain only

[8]Market equity is the only exception in that the t-stat is marginal in the full sample, closer to 10% significance. However this variable has always been included in asset pricing models and so I too include it here.

a fraction, 5% each, of the model's R^2 while market equity explains over 25%. Looking at Model 2 the results are even more puzzling with market equity explaining 51% of the R^2. These peculiar results regarding market equity largely disappear in the no micro cap sample as the % of R^2 falls to 7% and 21% in Models 1 and 2 respectively. In columns 1 and 3, short-term reversal does account for the largest % of R^2 which matches the t-statistic results. However, no other meaningful comparisons exist. While it is difficult to reconcile these results with Table 5, it does serve as an analysis from another angle. While market equity has low t-statistics in Fama-Macbeth regressions, it explains a large portion of the contemporaneous volatility. In spite of the low t-statistics, market equity remains a staple in asset pricing models in the words of Cochrane, "A statistically insignificant elephant is worth looking at." Low t-statistics could be a result of many things, most notably, a high correlation with the other variables in the model. I hypothesize that a complex cross-sectional correlation structure causes the lack of transparency between these two tables.

Beta Summary Statistics

To provide evidence for the question, "Are βs stable over time?," Figure 4 plots rolling 60-month averages of Fama-Macbeth βs. I see that with the exception of the short-term reversal β, all estimates do seem relatively stable over time. The short-term reversal β is trending upward which may be evidence of either an attenuation effect over time, or a persistent change in the characteristic over time. Table 7 provides summary statistics of my estimated βs. The mean column matches that of my full sample from the Fama-Macbeth regressions. Due to the frequency of measurement of the past return characteristics, the volatility of the short-term reversal and momentum βs are much larger than the variables updated annually. Since I smooth the βs by taking a full sample averages based on prior estimates, I also present the worst case standard errors (i.e. 60 months) of the β estimates. Once again I see considerable variability in the 95% confidence interval for my past return characteristics. This indicates the likelihood that these variables have large impacts on portfolio turnover.

The Accuracy of Predicted Expected Returns

Figure 5 presents plots of predicted vs. actual returns, to examine the accuracy of expected return models. For brevity I only show results for 25 portfolios. Data points plotting on the 45 degree line indicate an exact match in predicted vs. actual returns. Model 1 tends to overestimate poor returns and underestimate good returns, but generally does a good job in predicting portfolio expected returns. Model 2 also generally does a good job in predicting portfolio expected returns but tends to persistently underestimate all returns more often than Model 1. The Mean Absolute Error (MAE) is 0.284 for Model 1 for the full sample, 0.4071 for Model 1 for the no micro cap sample, and 0.209 for Model 2 for the full sample. This figure validates my claim that even though prediction error for individual stocks may be large, by aggregating these estimates into portfolios, many individual errors cancel out, resulting in a fairly accurate estimate at the portfolio level. This accuracy in the prediction of the portfolio's expected return is how my models fared so well in out-of-sample tests.

Incremental Explanatory Power

To examine the incremental contribution of each predictor, Tables 8 and 9 present selected performance statistics after each predictor was removed from the model. If an explanatory variable has a large incremental contribution to the out-of-sample prediction accuracy, then the return spread and other performance statistics should drop substantially. If an explanatory variable has little incremental contribution, these statistics should remain relatively unchanged. For the full sample the mean return drops by 1.50% per month when the short-term reversal is removed from model. The Sharpe ratio also falls by 0.28 per month and the CAPM α and β also fall. This is a huge drop in performance and therefore I can attribute most of the model's out-performance compared to existing studies to the inclusion of this variable. While this variable does contain considerable volatility and causes excessive portfolio turnover, its predictive content more than makes up for the additional transaction costs. In the no micro cap sample the short-term reversal results are not as large, only a 0.36% decrease in monthly performance. In fact, in the no micro cap sample it is momentum that has the largest incremental impact, a decrease in performance of 0.53%. In both samples, the size variable actually improves performance in every category when dropped from the model. All return measures increase and all risk measures decrease, so size actually hurts my model's performance in spite of the large contribution to R^2 in Fama-Macbeth regressions. This evidence confirms the claim by Lewellen (2014) that R^2 says nothing about the predictive power of that variable. Analyzing the other variables, in the full sample, book-to-market has the largest incremental contribution and investment has about twice the impact as profitability. In the no micro cap sample it is investment with the largest impact while profitability and book-to-market share equal contributions.

Robustness Checks

Available on the web-appendix found at the following link [9], I present complete performance statistics for over 20 different specifications not presented in the main tables, Tables 2, 3, and 4. I also show complete performance statistics for value weighted portfolios, including duplicate tables of Tables 2, 3, and 4. In all cases the value-weighted portfolios perform worse than the equal-weighted portfolios. This result is common in empirical studies since value-weighted portfolios tilt away from small firms and thus away from the small-firm effect. After transaction costs, the best performing value-weighted portfolio beats the benchmark value-weighted portfolio by 13.25% per year with a Sharpe ratio 1.8 times the size of the benchmark. While this is much lower than the equal-weighted case, the performance gains are still substantial. The additional specifications I present show complete performance statistics for hedged and short-sale constrained portfolios formed by groups of 100, 50, 25, 10, 5, and 2 portfolios. I also present alternate measures of turnover similar to Brandt et al. (2009), and the results show inconsequential differences. Given that I construct equal-weight portfolios, the individual weight on each asset is nearly identical in adjacent periods. Therefore, the turnover of a portfolio each time period is dominated by the percentage of stocks that are simply added and removed each period; with the mechanical changes playing a much smaller role.

Asset Pricing Model Implications

Clarke (2014a, 2014b) introduces the concept of a level, slope and curve factor model for stock returns with the familiar level, slope, and curve pattern that can be extracted from

[9] https://belkcollegeofbusiness.uncc.edu/rdickso6/

bond portfolios sorted by maturity. He finds that it performs better than leading factor models. He does this by extracting only priced factors using principal component analysis on portfolios formed by sorted expected returns. Figure 8 presents these same results, i.e. a perfect level, slope, and curve, using portfolios formed from my Model 1. My figure shows a much less ambiguous interpretation for a level, slope, and curve model for stock returns than that of Clarke (2014a, 2014b). This indicates that Clarke's model is not unique; furthermore, it also indicates that his procedure may not be optimal. If his level, slope, and curve model can beat leading factor models what about my model? I will investigate this question in another paper but it is an interesting result that a level, slope, and curve interpretation can be so unambiguously generated. This observations poses an interesting research question: Can I sort stocks by any characteristic and get a level, slope, and curve pattern? If so, which one is optimal and have I really simplified the factor structure of stock returns if this relationship is not unique? I am currently pursuing research down this path to determine how unique level, slope, and curve factors really are.

Conclusions

The cross-sectional anomaly literature has become large and unwieldy, with authors recently reporting between 50 and 300 known anomalous variables. The univariate characteristic sorting procedure popularized by Fama and French becomes awkward and infeasible in the presence of multiple characteristics, as the intersection of portfolios grows exponentially. Cochrane (2011) asks the question, "Which characteristics really provide *independent* information about average returns? Which are subsumed by others?" Drawing from the recent works of Novy-Marx (2013) and Fama and French (2015), I use only theoretically-motivated firm characteristics from these most recent innovations in the asset pricing literature. I use month by month multivariable cross-sectional regressions of these key characteristics to estimate conditional coefficient estimates, avoid the awkward sorting procedure, and aggregate the signals from multiple stock characteristics. By applying simple filtering techniques, I reduce estimation and sampling error, use only information known at time t, and predict expected returns. Finally, I sort stocks by their expected returns to form implementable portfolios. Unlike similar studies such as Haugen and Nardin (1996), Hanna and Ready (2005), and Lewellen (2014), who look at decile sorts only, I look at more extreme sorts of expected returns into 25 and 50 portfolios (top and bottom 2% and 4%). To appease the skeptical reader, note that my model's results are nearly identical to these prior studies when using decile sorts. Additionally I account for conservative measures of portfolio turnover and trading costs to measure the true return of the strategies. I also use a manageable set of explanatory variables which have statistically significant nonzero prices of risk in Fama-Macbeth regressions containing all variables, a claim that cannot be made by these studies. These related works use upwards of 50 variables in mutlivariable regressions, increasing concerns of data snooping.

My best performing equal-weighted portfolio beats a benchmark market portfolio of all stocks by 34.2% per year with only a modest increase in volatility, even after accounting for transaction costs. The corresponding annual Sharpe ratio is 1.97 per year compared to 0.42 per year, over 4.7 times the size of the benchmark. Looking at a more conservative sample that excludes micro cap stocks still yields annual returns after transaction costs that are 11.46% larger than the benchmark, and a Sharpe ratio 2.18 times the benchmark. I examine extreme values and tail risks and find that my portfolios outperform the market even in the worst of times. I examine the incremental predictive power of variables and find that the past return variables have the largest incremental impact while size actually makes expected return forecasts worse. My analysis provides detailed evidence that by exploiting the cross-section of expected returns combined with Naive diversification, an investor can earn large risk-adjusted returns in excess of a market portfolio, even after accounting for trading costs.

For future research, my observations regarding the incremental contributions of the past performance variables, measured monthly, offers an interesting avenue. If I can find variables that proxy for the level variables size, book-to-market, gross profitability, and investment, that are also measured monthly, perhaps this increased frequency of measurement would increase their incremental contributions. I am currently working on estimators to accomplish this goal.

Appendix

Data Description

For all accounting variables I employ the standard fiscal year matching popularized by Fama and French (1992). The accounting variables for fiscal years that end in calendar year t are matched with stock returns for July of year $t+1$ to June of year $t+2$. So there is at least a six month lag for the accounting variables in each monthly cross-sectional regression.

1. log(ME): Market equity is defined as price per share times shares outstanding from CRSP. To get ME for the firm, I aggregate values of all equity for a given permno and date. This aggregate value is assigned to the permno with the largest ME. Slightly deviating the Novy-Marx (2013) definition I update using the June market equity to

compute this variable rather than the previous December value since his increases its explanatory power in the cross section.

2. log(BE/ME): Book-to-market is book equity scaled by market equity. Book equity is shareholder equity, plus deferred taxes, minus preferred stock, when available. The shareholder equity components follow the tiered definitions consistent with those used in Fama and French (1993) to construct the HML factor. Stockholder equity is defined in Compustat as (SEQ) if available, or else common equity plus the carrying value of preferred stock is available (CEQ + PSTX) if available, otherwise total assets minus total liabilities (AT - LT) is used. Deferred taxes is deferred taxes and investment tax credits (TXDITC) if available, or else deferred taxes and/or investment tax credit (TXDB and/or ITCB). Preferred stock is redemption value (PSTKR) if available, or else liquidating value (PSTKRL) if available, or else carrying value (PSTK).

3. GPdat: Gross profits and earnings before extraordinary items are Compustat data items GP and IB, respectively. For free cash flow I employ net income plus depreciation and amortization minus changes in working capital minus capital expenditures (NI + DP - WCAPCH - CAPX). Gross profits are also defined as total revenue (REVT) minus cost of goods sold (COGS).

4. Inv: Investment for firms in year t is the growth of total assets for the fiscal year ending in year $t-1$ divided by total assets at the end of year $t-2$. This matches the definition used by (Fama & French, 2015). In their valuation equation, the investment variable is actually defined as the expected growth of book equity, not assets.

However, as they state, sorts on asset growth result in larger spreads of average return and using growth in book equity produces similar results.

5. R1to0: The short-term reversal measure is simply the return at time t, lagged by one period.

6. R12to2: The momentum measure is the previous year's 11 month return, skipping the previous month to prevent capturing short-term reversal in the momentum measure.

R^2 Decomposition

I begin with the following regression equation:

$$y = \mathbf{X}\gamma + \mu \tag{12}$$

I begin with the basic definition of R^2 for Eq. 12 found in Greene (1997).

$$R^2 = (\rho(y, \hat{y}))^2 \tag{13}$$

$$R^2 = (\rho(y, \hat{y}))^2 = \hat{\beta}$$

from the regression of (Lemma 1)

$$\hat{y} = \alpha + \beta y$$

Proof:

$$R^2 = (\rho(y, \hat{y}))^2 = \frac{cov(\hat{y}, y)^2}{var(\hat{y})var(y)} \tag{14}$$

$$\hat{y} = \alpha + \beta y$$

$$\hat{\beta} = \frac{cov(\hat{y}, y)}{var(y)} \tag{15}$$

$$cov(X\beta, X\beta + u) = cov(X\beta, X\beta) = cov(\hat{y}, \hat{y}) = var(\hat{y}) \tag{16}$$

$$R^2 = (\rho(y,\hat{y}))^2 = \frac{cov(\hat{y},y)^2}{var(\hat{y})var(y)} = \hat{\beta} = \frac{cov(\hat{y},y)}{var(y)} \tag{17}$$

Therefore to decompose R^2 I simply have:

$$R^2 = \sum_{\forall k} \frac{cov(y, \hat{\gamma_k} X_k)}{var(y)} \tag{18}$$

I construct this decomposition each month in my sample. Since the cross-sectional size differs from month to month (i.e. the number of firms), I multiply the covariance for k, each month, by the appropriate degrees of freedom, $N-1$, to yield the sum of squares. To get the R^2 decomposition presented in Table 6, I add up the marginal sum of squares for all time periods, divided by the total sum of squares for all time periods. I then normalize these and produce a statistic that sums to one.

$$\begin{aligned} SS_{k,t} &= (N-1)_t \times cov(y_t, \hat{\gamma_k} X_{k,t}) \\ SS_{y,t} &= (N-1)_t \times var(y_t) \\ \%R^2_k &= \frac{\sum_{\forall t} \frac{SS_{k,t}}{SS_{y,t}}}{\sum_{\forall k} \sum_{\forall t} \frac{SS_{k,t}}{SS_{y,t}}} \end{aligned} \tag{19}$$

Table 1: Return spreads

Spreads generated from the top portfolio as sorted by expected return to the bottom portfolio as sorted by expected return. To generate the expected returns a model of the form $r_{i,t+1} = \alpha + \beta_1 ln(BE/ME)_{i,t} + \beta_2 ln(ME)_{i,t} + \beta_3 GPdat_{i,t} + \beta_4 R1to0_{i,t} + \beta_5 R12to2_{i,t} + \beta_6 INV_{i,t}$ was estimated and the estimates were averaged over an estimation window initialized with 60 months of data. Each subsequent time period the model was re-estimated by enlarging the sample and the average estimates were updated every June. These average estimates were used generate the expected returns. All averaged estimates at time period t were generated using estimates prior to time period t. For Model 1 all estimates were used, for Model 2 $R1to0_{i,t}$ and $R12to2_{i,t}$ were used only as controls and thus β_4 and β_5 were not used to generate expected returns, and for Model 3 $R1to0_{i,t}$ was used only as a control and thus β_4 was not used to generate expected returns.

Portfolios	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
	Return Spreads			Sharpe Ratios		
Panel A: Full Sample						
2 Portfolios	1.0889	0.5571	0.7551	0.3130	0.0567	0.1394
5 Portfolios	2.1256	1.1143	1.3959	0.4598	0.1715	0.2386
10 Portfolios	2.8880	1.5767	1.8431	0.5348	0.2347	0.2836
25 Portfolios	3.9086	2.0832	2.3618	0.6247	0.2872	0.3498
50 Portfolios	4.4275	2.4841	2.6305	0.6364	0.3148	0.3619
100 Portfolios	4.8262	2.5981	2.7744	0.5821	0.2912	0.3387
	Return Spreads			Sharpe Ratios		
Panel B: No MicroCaps						
2 Portfolios	0.7634	0.3834	0.5390	0.1784	-0.0182	0.0411
5 Portfolios	1.3813	0.7216	1.0534	0.2850	0.0754	0.1276
10 Portfolios	1.7737	0.9143	1.4514	0.3230	0.1002	0.1727
25 Portfolios	2.2046	1.2145	1.8015	0.3367	0.1323	0.2065
50 Portfolios	2.5200	1.3158	2.1031	0.3449	0.1324	0.2217
100 Portfolios	2.5265	1.5176	2.0704	0.2829	0.1349	0.1918

All models are equally-weighted, i.e. Naive diversification. Value weighted statistics are available on the web appendix.

Table 2: Monthly traditional performance statistics

Performance statistics for out of sample returns for portfolios formed by stocks sorted by expected return. To generate the expected returns the models from the text estimated and the estimates were averaged over an estimation window initialized with 60 months of data. Each subsequent time period the model was re-estimated by enlarging the sample and the average estimates were updated every June. All averaged estimates at time period t were generated using estimates prior to time period t. For M1 all estimates were used, for M2 $R1to0_{i,t}$ and $R12to2_{i,t}$ were used only as controls. The number of portfolios were either 50, 25, or 2 as denoted in the Model column. ‘Turn’ is the monthly turnover, and TC denotes performance after transaction costs. All models are equally-weighted, i.e. Naive diversification. Value weighted statistics are available on the web appendix.

Models	Mean	Std	SR	Turn	MeanTC	SRTC	P-SR>0	P-SR$_{test}$ >SR$_{ben}$	AvgSize	AvgNum
Panel A: Full Sample, Hedge Portfolios										
Hedge50M1	4.427	6.291	0.636	0.445	3.982	0.566	0.000	0.000	370.936	144.143
Hedge25M1	3.909	5.578	0.625	0.423	3.485	0.549	0.000	0.000	454.279	288.285
Hedge50M2	2.484	6.543	0.315	0.108	2.377	0.298	0.000	0.016	691.070	144.143
Hedge25M2	2.083	5.777	0.287	0.099	1.984	0.270	0.000	0.048	853.580	288.285
Panel B: No MicroCaps, Hedge Portfolios										
Hedge50M1	2.520	6.078	0.345	0.437	2.083	0.273	0.000	0.000	1126.797	62.958
Hedge25M1	2.205	5.288	0.337	0.418	1.786	0.258	0.000	0.000	1279.738	125.915
Panel C: Shortsale Constrained Portfolios										
FullSamp25M2	2.182	7.644	0.230	0.100	2.083	0.217	0.000	0.000	18.715	144.143
NoMicroCaps25M2	1.524	6.952	0.158	0.073	1.451	0.148	0.000	0.105	472.531	62.958
Panel D: Low Volatility, Full Sample, Hedge Portfolios										
Hedge2M1	1.089	2.125	0.313	0.056	1.033	0.287	0.000	0.001	899.775	3603.566
Panel E: Benchmark Portfolio, All Stocks										
Benchmark	1.181	5.835	0.130	0.056	1.125	0.120	0.003	1.000	899.720	3603.566

Table 3: Fama and French 4-factor model performance

Monthly estimates and performance statistics for out of sample returns for portfolios formed by stocks sorted by expected return regressed against the Fama and French 4-factor model. To generate the expected returns the models from the text estimated and the estimates were averaged over an estimation window initialized with 60 months of data. All averaged estimates at time period t were generated using estimates prior to time period t. For M1 all estimates were used, for M2 $R1to0_{i,t}$ and $R12to2_{i,t}$ were used only as controls and thus β_4 and β_5 were not used to generate expected returns. The number of portfolios were either 50, 25, or 2 as denoted in the Model column. 'IRα' is the information ratio using the α estimate and the $\sigma_{residuals}$. 'IR$_{S\&P}$' is the information ratio relative the S&P500 index, commonly known as the tracking error. 'α_{TC}' is the α adjusted by transactions costs of 50 basis points. All models are equally-weighted, i.e. Naive diversification. Value weighted statistics are available on the web appendix.

Models	α	t-stat	β_{Mkt}	t-stat	β_{Smb}	t-stat	β_{Hml}	t-stat	β_{Mom}	t-stat	R^2	IR$_\alpha$	IR$_{S\&P}$	α_{TC}
Panel A: Full Sample, Hedge Portfolios														
Hedge50M1	3.648	11.517	0.166	1.971	0.312	1.800	0.462	2.433	0.070	0.513	0.062	0.599	0.514	3.202
Hedge25M1	3.145	11.120	0.156	2.047	0.331	1.893	0.447	2.631	0.056	0.434	0.078	0.588	0.481	2.722
Hedge50M2	1.664	6.825	-0.324	-5.529	0.578	5.639	1.099	10.247	0.049	0.601	0.340	0.312	0.206	1.556
Hedge25M2	1.322	6.438	-0.299	-6.082	0.561	6.537	1.048	11.381	-0.024	-0.380	0.398	0.294	0.173	1.223
Panel B: No MicroCaps, Hedge Portfolios														
Hedge50M1	1.488	5.570	0.260	3.504	0.292	1.978	0.614	4.188	0.296	2.677	0.123	0.262	0.265	1.051
Hedge25M1	1.154	5.003	0.262	3.760	0.246	1.855	0.638	4.724	0.320	3.519	0.168	0.239	0.241	0.735
Panel C: Shortsale Constrained Portfolios														
Full25M2	1.201	5.820	0.820	15.963	1.308	15.911	0.444	4.499	-0.331	-4.857	0.699	0.285	0.245	1.101
NoMicro25M2	0.365	3.355	1.131	42.080	0.868	18.452	0.721	14.401	-0.336	-6.045	0.899	0.164	0.194	0.292
Panel D: Low Volatility, Full Sample, Hedge Portfolios														
Hedge2M1	0.510	4.927	0.044	1.606	0.232	3.407	0.183	2.881	0.037	0.693	0.149	0.260	0.095	0.454
Panel E: Benchmark Portfolio, All Stocks														
Benchmark	0.182	2.615	0.959	49.157	0.830	21.625	0.235	6.406	-0.165	-5.287	0.954	0.146	0.160	0.127

Table 4: Extreme values and tail statistics

Statistics for monthly out of sample returns for portfolios formed by stocks sorted by expected return. To generate the expected returns a model of the form $r_{i,t+1} = \alpha + \beta_1 ln(BE/ME)_{i,t} + \beta_2 ln(ME)_{i,t} + \beta_3 GPdat_{i,t} + \beta_4 R1to0_{i,t} + \beta_5 R12to2_{i,t} + \beta_6 INV_{i,t}$ was estimated and the estimates were averaged over an estimation window initialized with 60 months of data. Each subsequent time period the model was re-estimated by enlarging the sample and the average estimates were updated every June. These average estimates were used generate the expected returns. All averaged estimates at time period t were generated using estimates prior to time period t. For M1 all estimates were used, for M2 $R1to0_{i,t}$ and $R12to2_{i,t}$ were used only as controls and thus β_4 and β_5 were not used to generate expected returns. 'Hedge' denotes a zero cost hedged portfolio formed by shorting the bottom portfolio and buying the top portfolio. The number of portfolios were either 50, 25, or 2 as denoted in the Model column. 'R<X%' counts the number of months where the portfolio return exceeded the stated threshold. There were 546 out-of-sample monthly returns.

Models	MaxDD	Calmar	CVaR (5%)	R<0%	R<-5%	R<-10%
Panel A: Full Sample, Hedge Portfolios						
Hedge50M1	47.054	1.129	-7.799	98.000	4.000	1.000
Hedge25M1	41.390	1.133	-6.886	109.000	4.000	1.000
Hedge50M2	36.962	0.806	-10.401	192.000	13.000	0.000
Hedge25M2	33.425	0.748	-9.817	186.000	10.000	0.000
Panel B: No MicroCaps, Hedge Portfolios						
Hedge50M1	40.940	0.739	-9.348	173.000	9.000	0.000
Hedge25M1	37.553	0.704	-8.542	165.000	8.000	0.000
Panel C: Shortsale Constrained Portfolios						
FullSamp25M2	45.914	0.570	-14.153	195.000	26.000	3.000
NoMicroCaps25M2	49.170	0.372	-15.424	214.000	24.000	5.000
Panel D: Low Volatility, No MicroCaps, Hedge Portfolios						
Hedge2M1	17.387	0.752	-3.386	136.000	0.000	0.000
Panel E: Benchmark Portfolio, All Stocks						
Benchmark	44.647	0.317	-12.476	211.000	16.000	2.000

All models are equally-weighted, i.e. Naive diversification. Value weighted statistics are available on the web appendix.

Table 5: Statistical significance of the average estimated marginal effects

Panel A: All firms

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Constant	1.39	1.66	1.31	1.53	1.10	1.51	1.54
	(4.65)	(5.55)	(4.31)	(5.05)	(5.81)	(4.89)	(4.81)
GPdAT	0.68	.	0.76	0.41	0.80	0.66	0.68
	(6.04)	(.)	(6.90)	(3.60)	(6.96)	(5.87)	(5.83)
INV	−0.56	−0.57	.	−0.77	−0.52	−0.52	−0.63
	(−7.59)	(−7.84)	(.)	(−8.55)	(−6.87)	(−6.79)	(−7.92)
log(BE/ME)	0.36	0.30	0.44	.	0.46	0.35	0.28
	(6.52)	(5.39)	(7.47)	(.)	(8.32)	(6.10)	(4.47)
log(ME)	−0.07	−0.08	−0.07	−0.11	.	−0.10	−0.06
	(−1.72)	(−2.09)	(−1.72)	(−2.90)	(.)	(−2.36)	(−1.50)
$R_{1:0}$	−5.72	−5.60	−5.57	−5.47	−5.30	.	−5.51
	(−14.17)	(−13.82)	(−13.65)	(−13.05)	(−12.45)	(.)	(−12.57)
$R_{12:2}$	0.75	0.75	0.83	0.68	0.83	0.73	.
	(3.98)	(3.96)	(4.34)	(3.57)	(4.19)	(3.72)	(.)
R^2	0.046	0.043	0.044	0.040	0.034	0.038	0.036
	(23.46)	(22.21)	(22.74)	(22.17)	(21.02)	(22.35)	(22.29)

Panel B: Excluding MicroCap Firms

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Constant	1.36	1.57	1.22	1.49	0.94	1.24	1.63
	(4.01)	(4.55)	(3.48)	(4.30)	(5.12)	(3.59)	(4.56)
GPdAT	0.65	.	0.78	0.32	0.69	0.66	0.57
	(4.75)	(.)	(5.77)	(2.26)	(4.91)	(4.83)	(4.07)
INV	−0.26	−0.29	.	−0.55	−0.22	−0.25	−0.34
	(−2.83)	(−3.19)	(.)	(−4.78)	(−2.22)	(−2.64)	(−3.43)
log(BE/ME)	0.29	0.20	0.36	.	0.32	0.29	0.17
	(4.64)	(3.17)	(5.16)	(.)	(5.16)	(4.51)	(2.42)
log(ME)	−0.07	−0.08	−0.06	−0.09	.	−0.06	−0.08
	(−1.89)	(−2.10)	(−1.51)	(−2.54)	(.)	(−1.66)	(−2.08)
$R_{1:0}$	−4.20	−4.04	−4.04	−3.97	−4.08	.	−3.87
	(−9.55)	(−9.13)	(−9.06)	(−8.75)	(−9.13)	(.)	(−8.05)
$R_{12:2}$	0.83	0.83	0.88	0.75	0.86	0.88	.
	(3.78)	(3.79)	(3.98)	(3.40)	(3.90)	(3.93)	(.)
R^2	0.067	0.062	0.063	0.059	0.059	0.057	0.050
	(28.40)	(26.16)	(27.49)	(27.39)	(26.24)	(27.17)	(26.61)

The table summarizes the results of the cross-sectional regressions for NYSE, AMEX, and NASDAQ stocks. I fit the regressions for every month from July 1963 to December 2013, and report the average value of the time series of estimated slope coefficients for each explanatory variable. Fama-Macbeth t-statistics are shown below the average slopes in parentheses.

Table 6: Percentage of r-squared decomposition for cross-sectional regressions from July 1963 to Dec 2013.

Predictor	M1	M2	M3	M4
logBEME	12.1221	26.7761	19.4277	45.2777
logME	25.4207	51.4146	7.2388	20.9481
GPdaT	4.5065	8.3561	5.5373	11.5742
INV	5.1980	13.4532	7.8589	22.2001
R1to0	24.4558	–	23.8045	–
R12to2	28.2970	–	36.1327	–

M1: Full Model all Variables

M2: Full Model no R1to0 and no R12to2

M3: No Micro Sample all Variables

M4: No Micro Sample no R1to0 and no R12to2

Table 7: Beta summary statistics of the full sample

July 1963 to Dec 2013. The columns 'StdErr - Min T' is the standard error of the mean of the beta estimate for the minimum T case, i.e. 60 months. Each estimation period I use 12 more months of data to smooth the beta estimates so these 95% confidence limits represent the instability of the beta estimates in the worst case. Each subsequent time period this interval gets smaller at the rate of $\sqrt{(T)}$.

Statistic	Mean	StDev	StdErr - Min T	Lower 95% - Min T	Upper 95% - Min T
Intercept	1.3874	7.3515	0.9491	-0.5107	3.2856
logBEME	0.3603	1.3608	0.1757	0.0089	0.7116
logME	-0.0666	0.9517	0.1229	-0.3123	0.1792
GrossProf	0.6757	2.7536	0.3555	-0.0353	1.3867
Inv	-0.5594	1.8141	0.2342	-1.0277	-0.0910
STR	-5.7194	9.9372	1.2829	-8.2852	-3.1536
Mom	0.7500	4.6356	0.5985	-0.4469	1.9469

Table 8: Performance of hedged top 4% - bot 4%(25 portfolios), full sample

Monthly statistics for out-of-sample returns for a test portfolio of the top X $\%$ of stocks as sorted by expected return, vs. a portfolio of the top X $\%$ of stocks as sorted by expected return as estimated by the full model. To generate the expected returns for the full model a of the form $r_{i,t+1} = \alpha + \beta_1 ln(BE/ME)_{i,t} + \beta_2 ln(ME)_{i,t} + \beta_3 GPdat_{i,t} + \beta_4 R1to0_{i,t} + \beta_5 R12to2_{i,t} + \beta_6 INV_{i,t}$ was estimated and the estimates were averaged over an estimation window initialized with 60 months of data. Each subsequent time period the model was re-estimated by enlarging the sample and the average estimates were updated every June. These average estimates were used generate the expected returns. All averaged estimates at time period t were generated using estimates prior to time period t. For the test portfolios, e.g. M33Prof, the profitability variable was removed from the expected return regression and portfolios were reformed. Rows 1 - 6 represent the differences between the test model and the full, Rows 7 - 12 show the test model, and Rows 13 - 18 show the full model.

Stats	M101Prof	M102Inv	M103BEME	M104ME	M105Str	M106Mom
MeanDiff	-0.0787	-0.1847	-0.2857	0.0868	-1.4983	-0.0663
StdDiff	0.1222	0.4579	0.0620	-0.4382	0.1265	1.1478
SharpeRDiff	-0.0272	-0.0780	-0.0575	0.0701	-0.2765	-0.1165
CAPM AlphaDiff	-0.0769	-0.2290	-0.2908	0.0982	-1.3023	-0.0667
CAPM BetaDiff	-0.0039	0.0928	0.0108	-0.0240	-0.4107	0.0009
IRalphaDiff	-0.0271	-0.0801	-0.0575	0.0716	-0.2354	-0.1167
Mean	3.8299	3.7239	3.6230	3.9954	2.4104	3.8424
Std	5.7006	6.0362	5.6404	5.1402	5.7048	6.7262
SharpeR	0.5975	0.5467	0.5672	0.6948	0.3482	0.5082
CAPM Alpha	3.3480	3.1958	3.1340	3.5231	2.1226	3.3581
CAPM Beta	0.1215	0.2183	0.1363	0.1015	-0.2852	0.1264
IRalpha	0.5907	0.5377	0.5604	0.6894	0.3824	0.5011
Mean	3.9086	3.9086	3.9086	3.9086	3.9086	3.9086
Std	5.5784	5.5784	5.5784	5.5784	5.5784	5.5784
SharpeR	0.6247	0.6247	0.6247	0.6247	0.6247	0.6247
CAPM Alpha	3.4248	3.4248	3.4248	3.4248	3.4248	3.4248
CAPM Beta	0.1255	0.1255	0.1255	0.1255	0.1255	0.1255
IRalpha	0.6178	0.6178	0.6178	0.6178	0.6178	0.6178

All models are Equally Weighted

Table 9: Performance of hedged top 4% - bot 4%(25 portfolios), no micro sample

Monthly statistics for out-of-sample returns for a test portfolio of the top X $\%$ of stocks as sorted by expected return, vs. a portfolio of the top X $\%$ of stocks as sorted by expected return as estimated by the full model. To generate the expected returns for the full model a of the form $r_{i,t+1} = \alpha + \beta_1 ln(BE/ME)_{i,t} + \beta_2 ln(ME)_{i,t} + \beta_3 GPdat_{i,t} + \beta_4 R1to0_{i,t} + \beta_5 R12to2_{i,t} + \beta_6 INV_{i,t}$ was estimated and the estimates were averaged over an estimation window initialized with 60 months of data. Each subsequent time period the model was re-estimated by enlarging the sample and the average estimates were updated every June. These average estimates were used generate the expected returns. All averaged estimates at time period t were generated using estimates prior to time period t. For the test portfolios, e.g. M33Prof, the profitability variable was removed from the expected return regression and portfolios were reformed. Rows 1 - 6 represent the differences between the test model and the full, Rows 7 - 12 show the test model, and Rows 13 - 18 show the full model.

Stats	M107Prof	M108Inv	M109BEME	M110ME	M111Str	M112Mom
MeanDiff	-0.0724	-0.1442	-0.0776	0.1915	-0.3573	-0.5331
StdDiff	0.0514	0.1120	-0.0468	-0.0871	1.0671	0.6985
SharpeRDiff	-0.0168	-0.0337	-0.0118	0.0425	-0.1128	-0.1283
CAPM AlphaDiff	-0.0711	-0.2177	-0.0904	0.2291	-0.1541	-0.5338
CAPM BetaDiff	-0.0028	0.1540	0.0269	-0.0788	-0.4259	0.0014
IRalphaDiff	-0.0168	-0.0402	-0.0136	0.0481	-0.0757	-0.1286
Mean	2.1321	2.0603	2.1269	2.3961	1.8472	1.6714
Std	5.3392	5.3998	5.2411	5.2008	6.3549	5.9864
SharpeR	0.3199	0.3031	0.3249	0.3792	0.2240	0.2084
CAPM Alpha	1.6433	1.4967	1.6240	1.9435	1.5603	1.1806
CAPM Beta	0.1360	0.2928	0.1656	0.0600	-0.2871	0.1402
IRalpha	0.3103	0.2869	0.3135	0.3752	0.2514	0.1985
Mean	2.2046	2.2046	2.2046	2.2046	2.2046	2.2046
Std	5.2878	5.2878	5.2878	5.2878	5.2878	5.2878
SharpeR	0.3367	0.3367	0.3367	0.3367	0.3367	0.3367
CAPM Alpha	1.7144	1.7144	1.7144	1.7144	1.7144	1.7144
CAPM Beta	0.1388	0.1388	0.1388	0.1388	0.1388	0.1388
IRalpha	0.3271	0.3271	0.3271	0.3271	0.3271	0.3271

All models are Equally Weighted

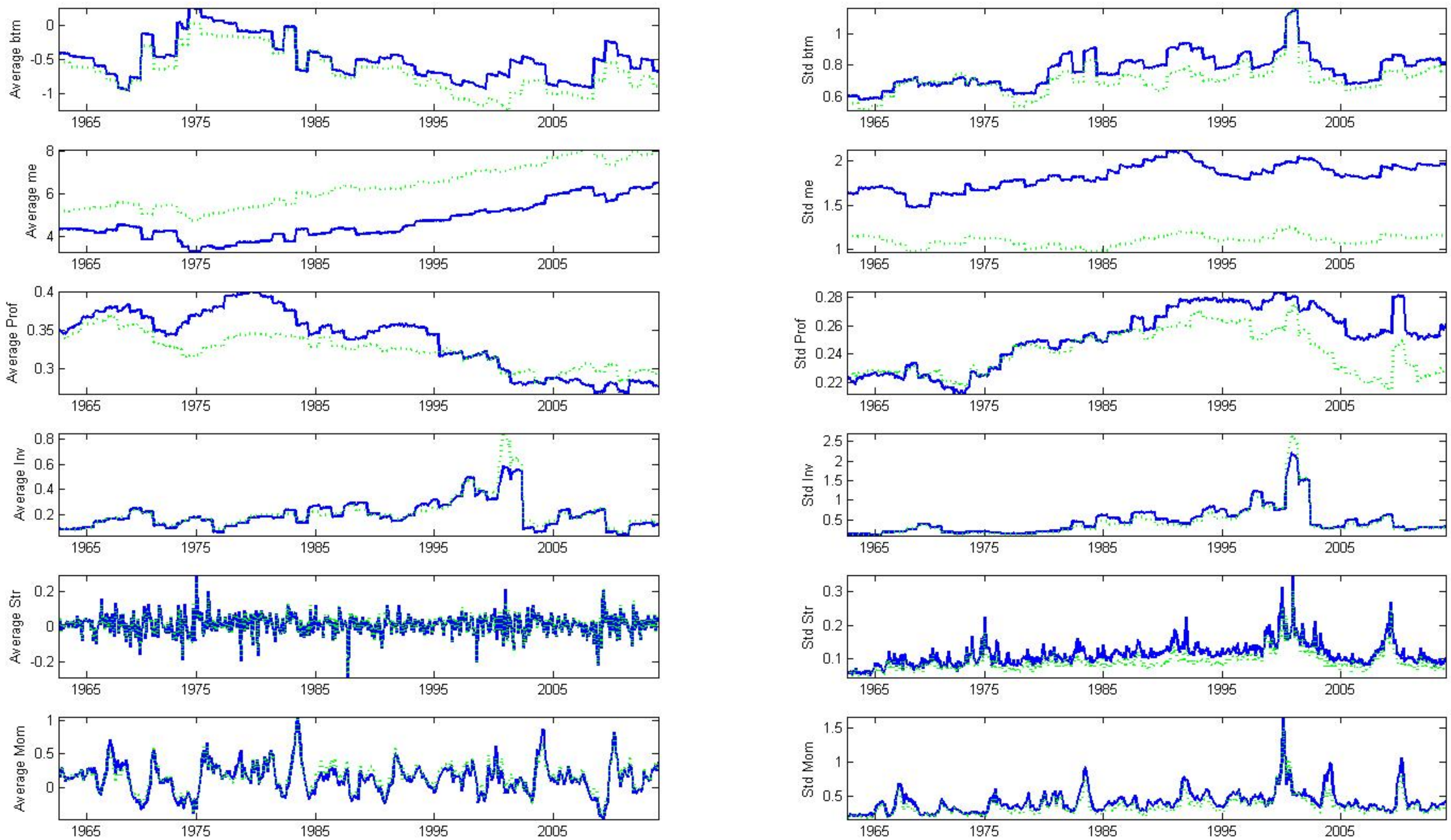


Figure 1: Firm-level summary statistics

This figure displays the cross-sectional means and standard deviations of the firm level characteristics. The left column shows average values and the right column shows standard deviations. The blue line is the full sample and green is the no MicroCap sample.

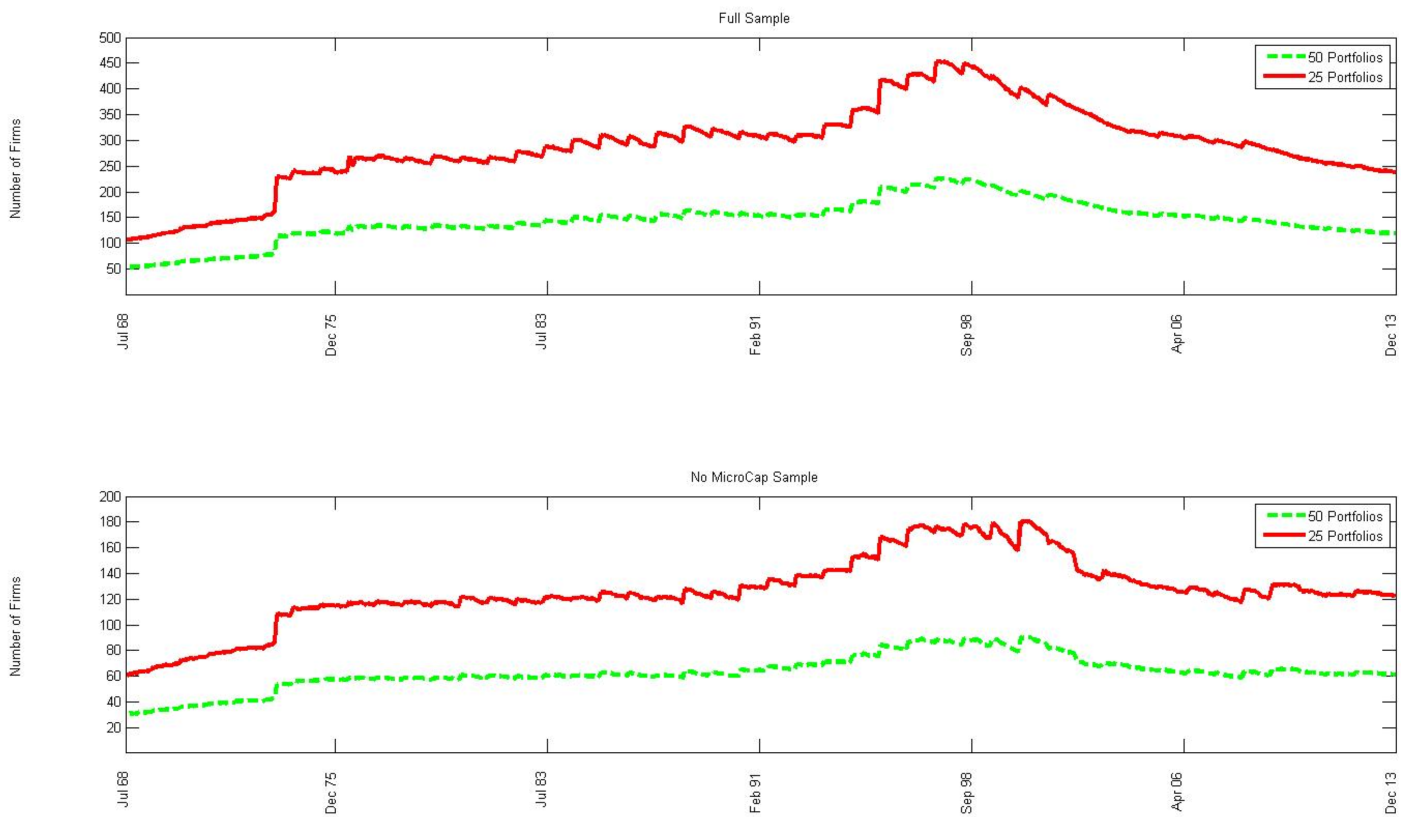


Figure 2: Number of firms over time

The number of firms per month over my out-of-sample testing period, Jul 1968 - Dec 2013. These plot show the number of firms for the zero-cost hedge portfolios. Of course for the short-sale constrained portfolio the number of firms just half as large. The top figure shows the full sample and the bottom figure shows the no MicroCap sample.

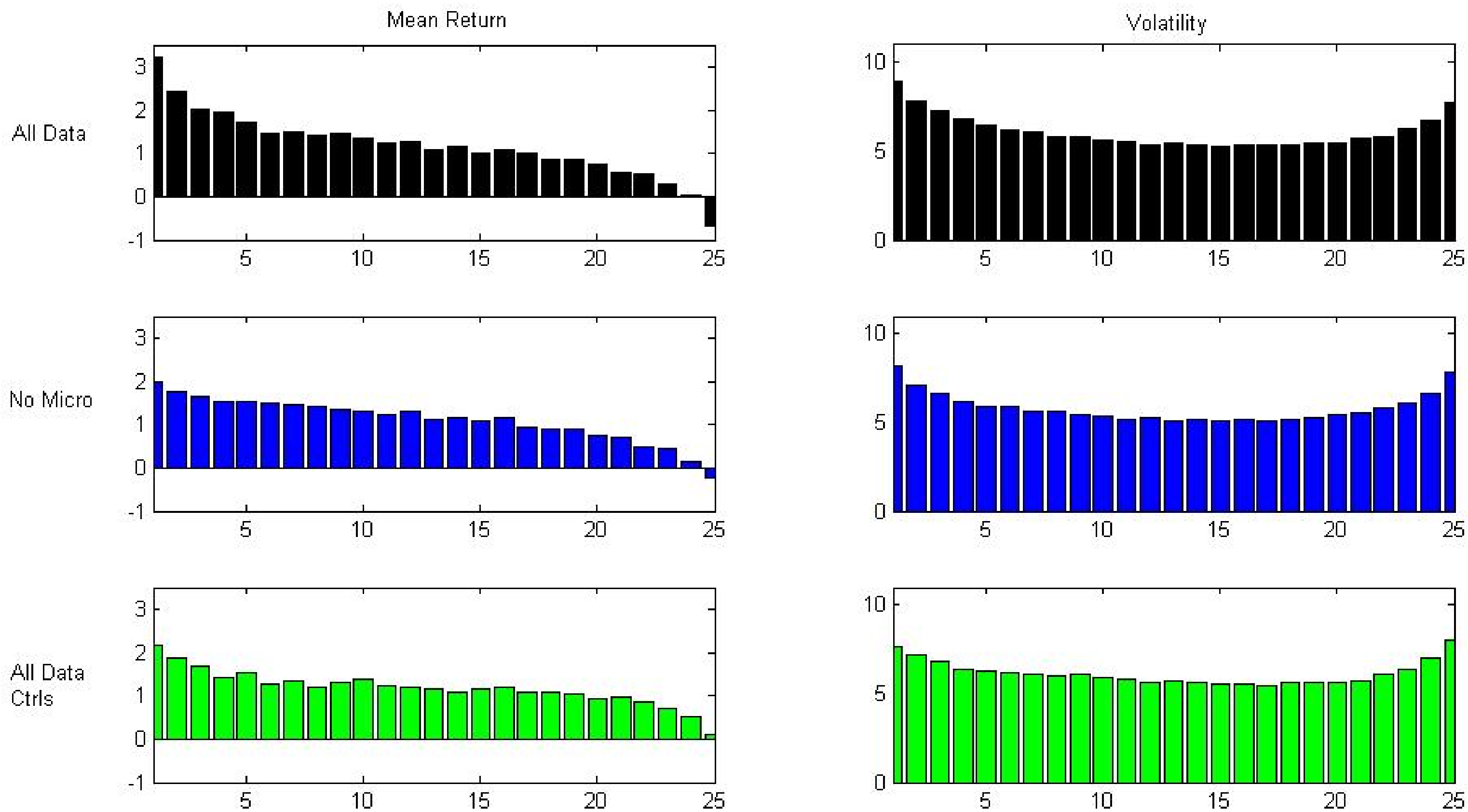


Figure 3: Portfolios sorted by expected return

25 Portfolios Sorted by Expected Returns from highest to lowest, 1968 - 2013. Row 1 consists of Model 1 with the Full Sample, Row2 consists of Model 1 with the No MicroCap Sample, and Row3 consists of Model 2 with the Full Sample.

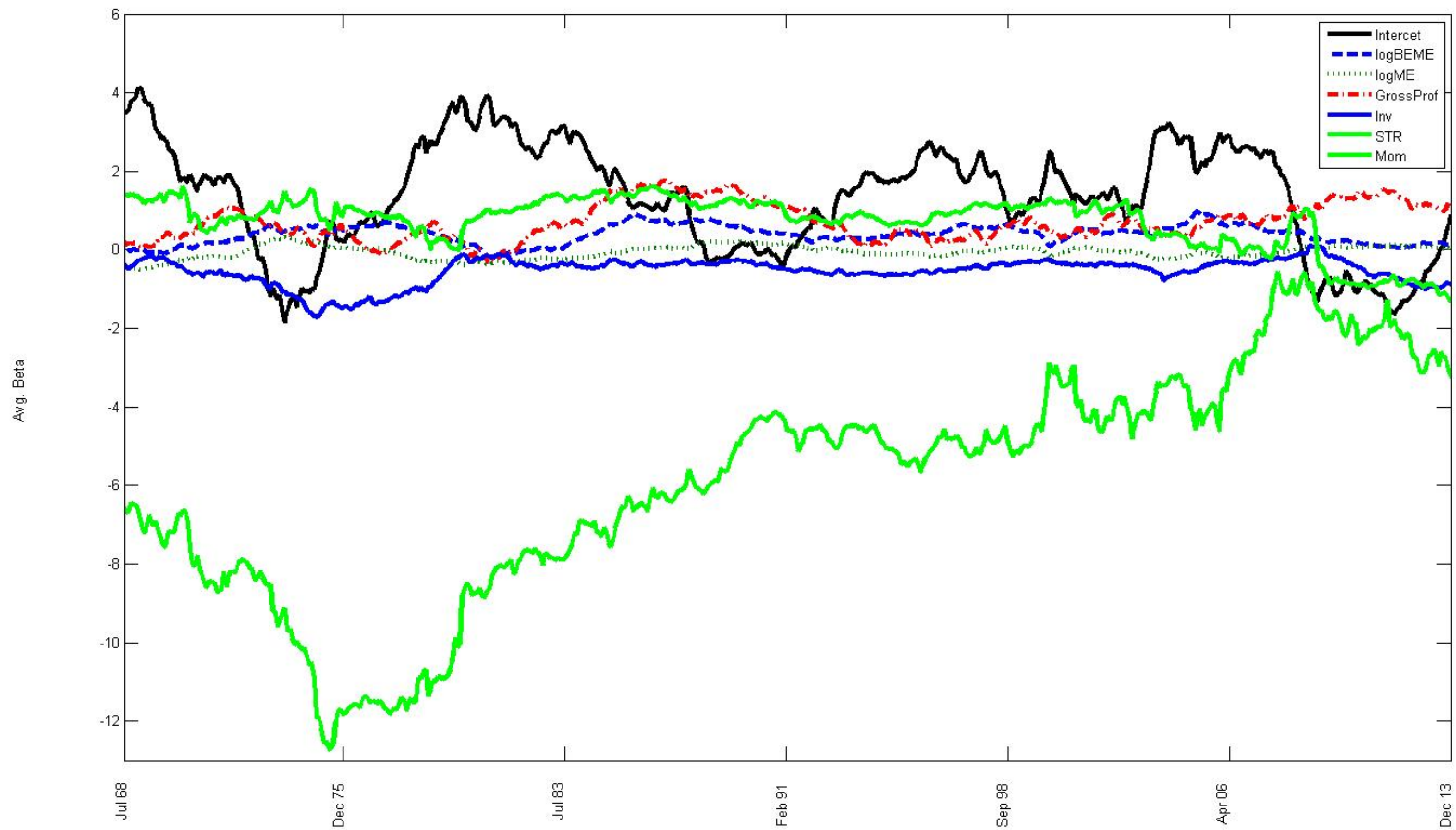


Figure 4: Rolling 60 month averages of Fama and Macbeth betas

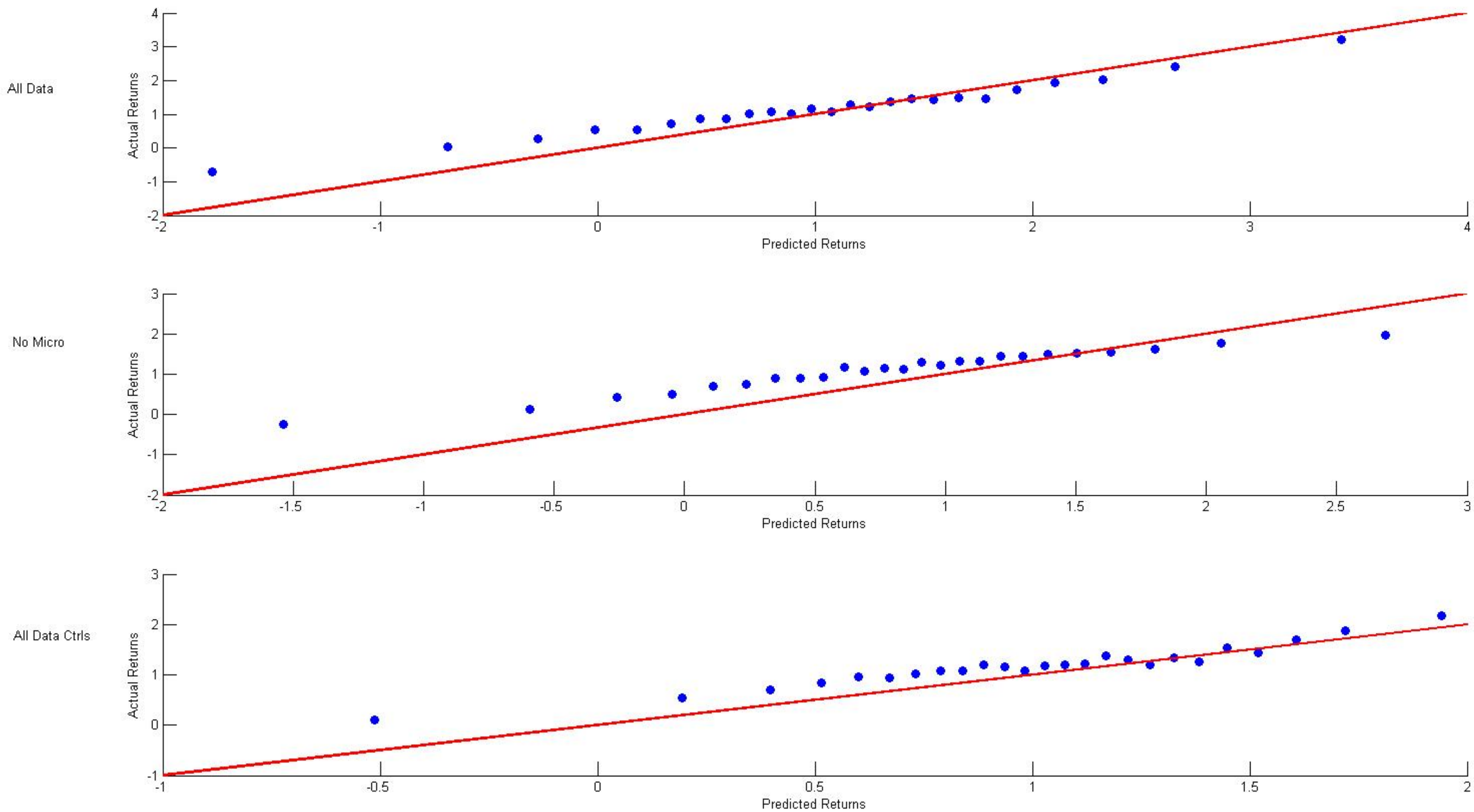


Figure 5: Predicted vs actual returns for 25 Portfolios

Row 1 consists of Model 1 with the Full Sample, Row2 consists of Model 1 with the No MicroCap Sample, and Row3 consists of Model 2 with the Full Sample.

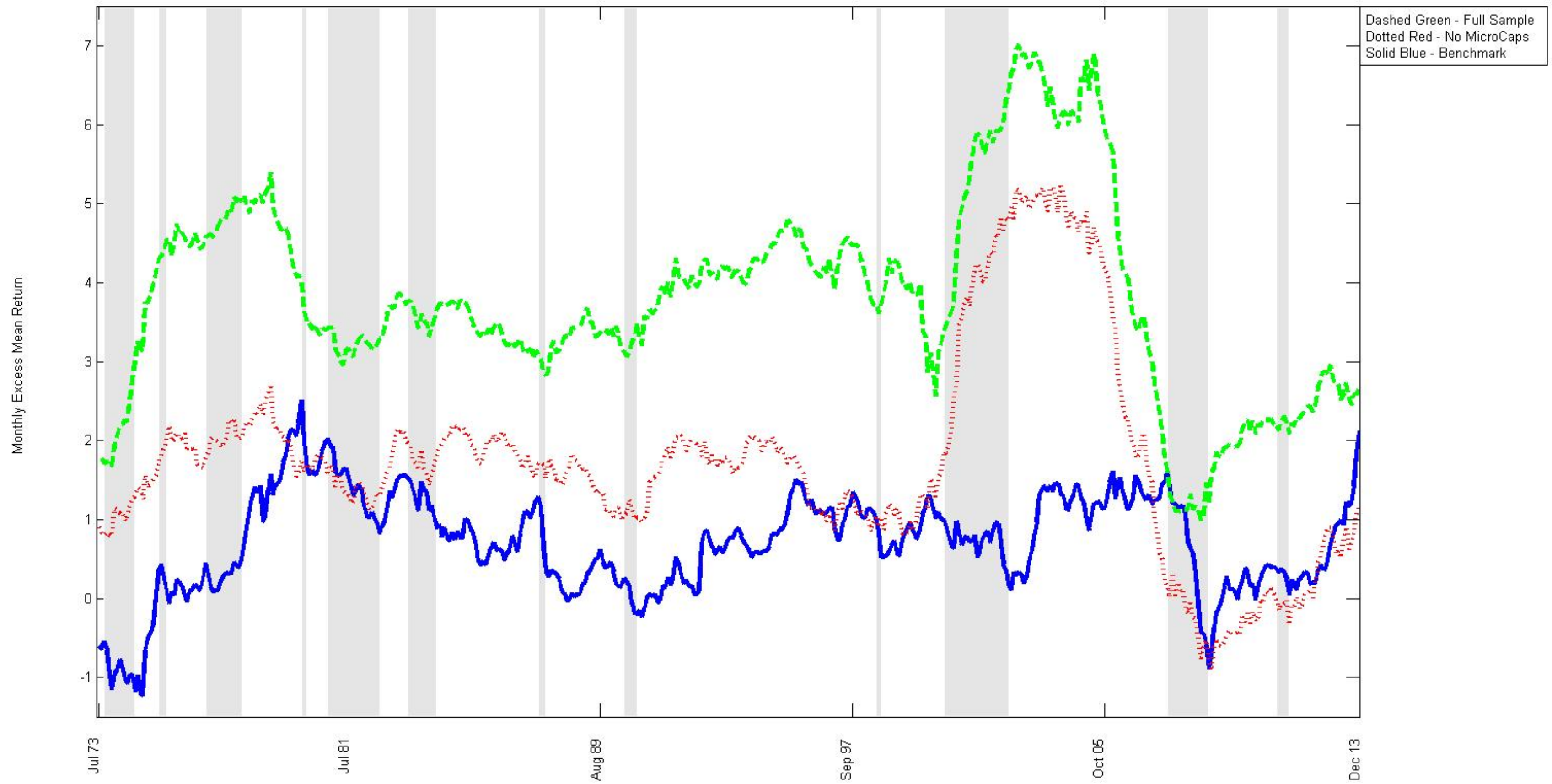


Figure 6: Rolling 60-month excess mean returns

The Best Performing Strategies(equally-weighted Hedge Portfolios, Top and Bottom 2% using Model 1) for both the full sample and the no microcap sample. The shaded rectangles indicate bear markets or time periods of greater than 10% decline in the S&P 500 index.

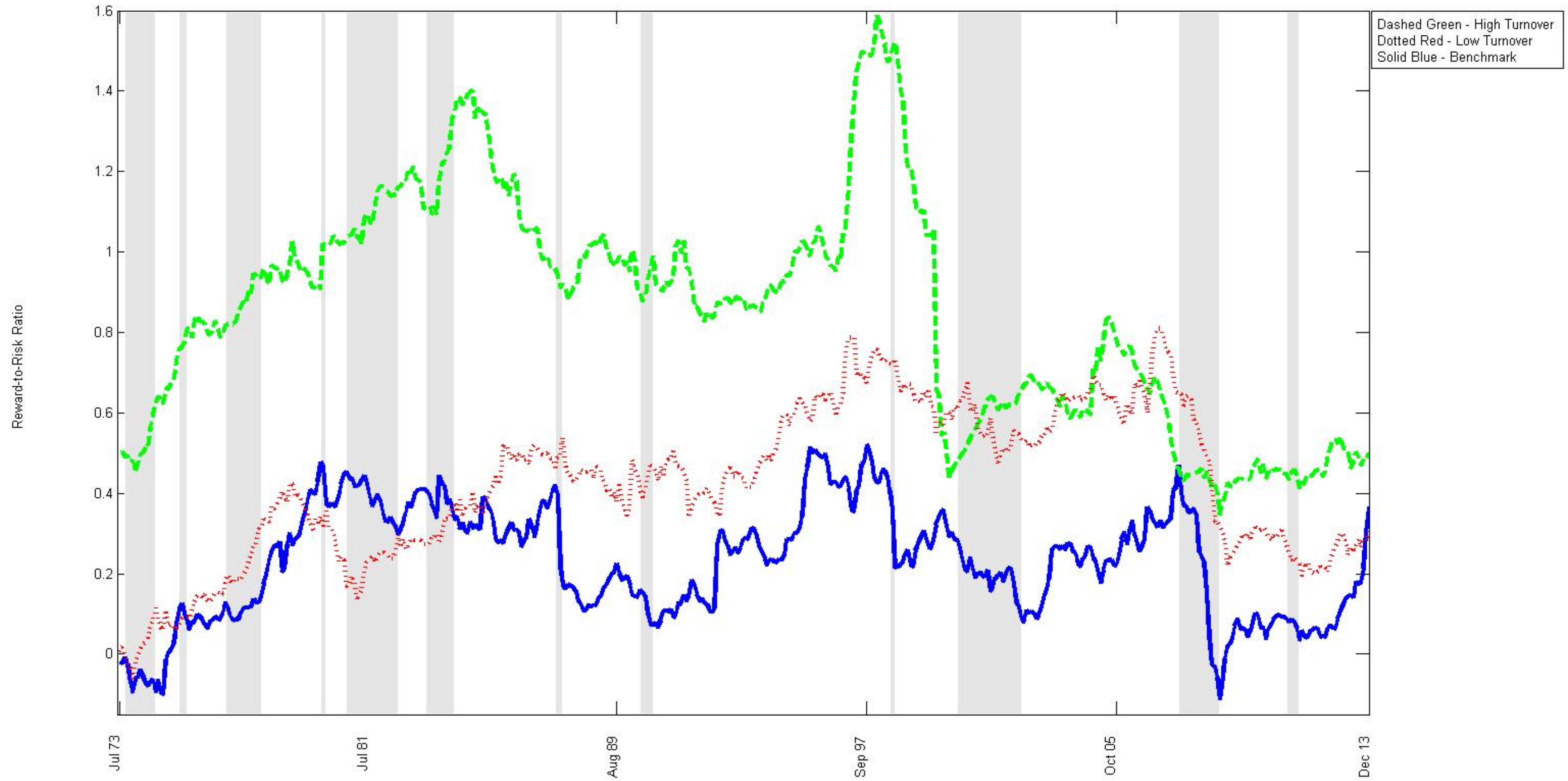


Figure 7: Rolling 60-month reward-to-risk ratios

The Best Performing(equally-weighted Hedge Portfolios, Top and Bottom 2%) High(Model 1) and Low(Model 2) Turnover Strategy for the full sample. The shaded rectangles indicate bear markets or time periods of greater than 10% decline in the S&P 500 index.

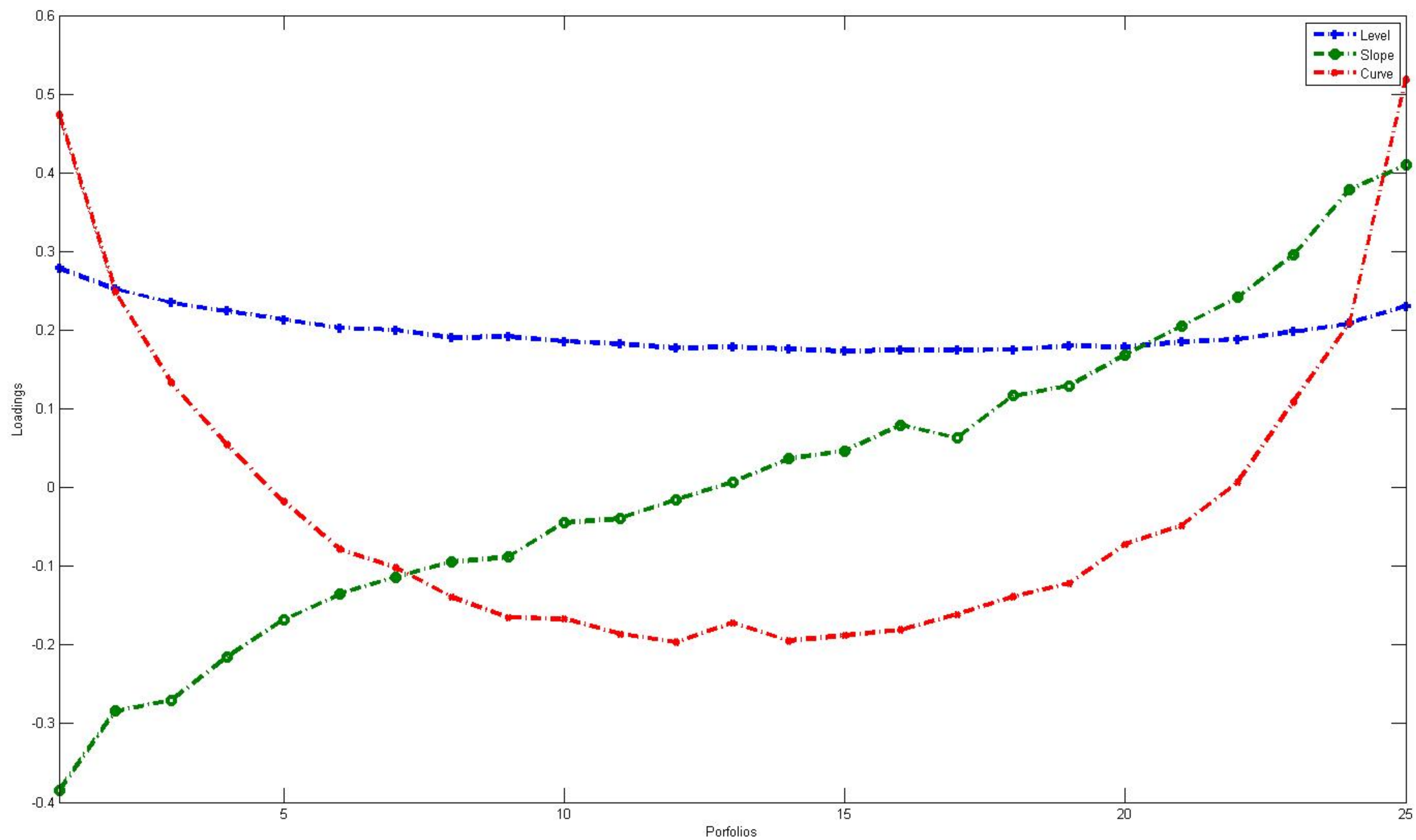


Figure 8: Factor loadings of the first three principal components for 25 equally-weighted portfolios

These portfolios were formed from the forecasted expected return using Model 1 in the full sample.

CHAPTER 3: NAIVE DIVERSIFICATION ISN'T SO NAIVE AFTER ALL

Introduction

Modern Portfolio Theory of course began with the seminal work of Markowitz (1952) who developed the workhorse theory of mean-variance efficiency. The two central conditions of Markowitz's fundamental model are that: (1) investors must desire to act according to the mean-variance efficient outcome and (2) investors must be able to arrive at a reasonable estimate for the mean return and covariance structure of asset choices. Estimation risk is a common term which emphasizes the failure of Markowitz's second condition. Estimation risk is the foundation of my work and the related literature. Samuelson (1967) notably added to Markowitz's work by proving that a $1/N$ (naive), equally weighted portfolio was the optimal strategy when return distributions are independently and identically distributed (*IID*). An implication of naive investing is an assumption that all moments of returns are equal; specifically the mean and variance for mean-variance investors . These assumptions, while likely not realistic, give rise to some appealing features regarding naive diversification. As discussed by Demiguel et al. (2009) and Kirby and Ostdiek (2012), some of these features include: no estimation error, no optimization, no matrix inversion, no shorts, extremely low turnover, and easy application to a large number of assets.

This paper directly compliments the work of Demiguel et al. (2009) and Kirby and Ostdiek (2012). These authors both present compelling arguments but offer very different conclusions. This study presents a middle ground between these authors' positions and provides evidence consistent with both. I find simply that naive diversification is hard to beat and perhaps naive investing isn't so naive after all. I conduct a horse-race of the most recent innovations in portfolio optimization techniques using actual stock data, most similar to the study presented by Demiguel et al. (2009). To create my empirical datasets I use the sequential cross-sectional regression methodology described in Dickson (2015) to predict top performing stocks. Following from comments by both Demiguel et al. (2009) and Kirby and Ostdiek (2012), I introduce a statistic to measure the cross-sectional dispersion of the conditional means and volatilities of my data. Consistent with Kirby and Ostdiek (2012), I conclude that the cross-sectional dispersion of the Sharpe ratios in my top performing stock portfolios are simply too small for mean-variance extensions to outperform naive diversification. These results add to the mounting evidence of the poor performance of portfolio optimization techniques. Using robust simulations, I do confirm that the extensions proposed by Kirby and Ostdiek (2012) are successful at improving the performance of mean-variance optimization. However, due to the very issues that these authors discuss, these techniques are incapable of outperforming naive diversification using actual stock data.

Demiguel et al. (2009) compared the out-of-sample performance of 14 competing portfolio strategies and ultimately determined that estimation risk eroded nearly all the gains from sophisticated optimization techniques. Simply put, no strategy consistently outperformed naive diversification. Kirby and Ostdiek (2012) presented evidence to refute this claim and argued that the research design of Demiguel et al. (2009) placed mean-variance optimization at a severe disadvantage in terms of estimation risk and turnover. They developed simple extensions of mean-variance optimization designed to reduce estimation risk and portfolio turnover, and showed that their extensions outperform naive diversification even in the presence of high trading costs. Furthermore, they present evidence that the performance of their extensions is driven by characteristics of the datasets tested. Specifically, datasets that do not have large cross-sectional dispersions in means and variances will likely not perform well using strategies designed to exploit this dispersion, i.e. mean-variance strategies.

I contribute to this line of work in several key ways. First I conduct a comprehensive portfolio analysis using individual stock data and not portfolios of stocks. While some authors such as Green and Hollifield (1992), Jagannathan and Ma (2003), and Brandt et al. (2009) used individual stock data in a portfolio analysis, my study is the first to present a horse-race of the most recent innovations in portfolio optimization using individual stocks. Both Demiguel et al. (2009) and Kirby and Ostdiek (2012), as well as many other authors such as Kan and Z. (2007), Garlappi, Uppal, and Wang (2007), Kirby and Ostdiek (2015), and DeMiguel, Martin-Utrera, and Nogales (2013) *et al.*, only consider portfolios of stocks when testing competing models. Demiguel et al. (2009) comment that this gives naive diversification an advantage because diversified portfolios have lower idiosyncratic volatility than stocks, so the loss from using naive diversification as opposed to optimal strategies is smaller (p.1920). Their comment hints at what is claimed by Kirby and Ostdiek (2012), i.e. that the cross-sectional dispersions in means and variances of the datasets drives performance. In addition to attenuating the apparent advantage towards naive diversification, using individual stock data has another practical benefit; that is, it more accurately replicates a fund manager's portfolio construction dilemma. Common datasets used in the aforementioned studies include the readily available characteristic based portfolios from Ken French's data library, industry portfolios, or country indices. Fund managers are generally more interested in investing in a particular group of stocks or sectors that are more likely to outperform as provided by their team of analysts. They are likely less interested in investing in portfolios of stocks that are not traded, like the data found in Ken French's data library, or simply all sectors and countries taken as a whole. The procedure described by Dickson (2015) aggregates forecasting signals from multiple sources known at time t, and predicts stock returns. These portfolios then serve as an excellent proxy for a group of actual stocks that are likely to outperform. The added advantage of an analysis with these data is that these results are tradeable and complete with aggressive adjustments for portfolio turnover.

Second, Demiguel et al. (2009) conclude that the approach proposed by Brandt et al. (2009) shows the most promise in for the performance of optimized portfolios. Therefore I include this model in my tests and use multiple sets of stock characteristics not examined by Brandt et al. (2009).

Third, I provide a robust Monte Carlo analysis of the models introduced by Kirby and Ostdiek (2012) to further test if these recent innovations are successful at reducing estimation risk and turnover. Additionally I present a fund of funds approach to further reduce estimation error and test several extensions of the reward-to-risk timing methodology proposed by Kirby and Ostdiek (2012).

In Section 2 I describe all of the portfolio strategies tested in the study. In Section 3 I describe the sources of data, the construction of the top performing stocks used in the analysis, the methodology for evaluating performance, and the details of the simulation experiment. Section 4 presents and discusses all results and Section 5 concludes.

Portfolio Strategies

All of the asset-allocation models considered are some variant of the solution to an expected utility maximization problem. Details for the much of the motivating theory for these models can be found in Brandt et al. (2009) and Kirby and Ostdiek (2012). The model of Brandt et al. (2009) is generalized for any utility function while the variants of mean-variance optimization are most commonly stated in terms of quadratic utility. However, mean-variance solutions can also be solved explicitly for CARA utility functions with normal *IID* returns. Mean-variance preferences can also be generalized for other utility functions with the assumption of normal *IID* returns, but the solutions may not be solved explicitly. If returns are normally distributed, then a 2nd-order Taylor expansion of any utility function will yield a problem in which only estimates for the first two moments of returns are required. Given this generality I will describe the tested strategies for two categories: mean-variance strategies and parametric portfolio choice. Table 11 summarizes the models presented below.

Mean-Variance Strategies

To introduce some notation $r_t = R_t - R_{f,t}\mathbf{1}$, $\mu_t = E_t[r_{t+1}]$, and $\Sigma_t = E[(r_{t+1} - \mu_{t+1})(r_{t+1} - \mu_{t+1})']$ $R_{f,t}$ equals the risk-free rate at time t and R_t equals the vector of gross risky excess returns at time t. I also denote w_t as the vector risky-asset weights and $w_{rf,t}$ as weight in the risk-free asset. With this notation the conditional portfolio mean vector is computed as $\mu_{p,t} = w_t'\mu_t$ and the conditional portfolio variance is computed as $\sigma_p^2 = w_t'\Sigma_t w_t$. The standard conditional mean-variance optimization problem solves the following optimization problem:

$$\min_{w_t} \ \frac{1}{2} w_t' \Sigma_t w_t \ , \ s.t. \ \ w_t' \mu = \mu_{p,t} \tag{20}$$

Equivalently one could maximize the portfolio expected return subject to a given level of portfolio risk. The well known solution is:

$$w_t = \frac{\mu_{p,t} \Sigma^{-1} \mu_t}{\mu_t' \Sigma_t^{-1} \mu_t} \tag{21}$$

For the tangency portfolio the risky asset weights sum to one so imposing this restriction and solving yields the following well known tangency portfolios weights:

$$w_{tp} = \frac{\Sigma^{-1} \mu_t}{\mathbf{1}' \Sigma_t^{-1} \mu_t} \quad \text{where } \mu_{p,t} = \frac{\mu_t' \Sigma_t^{-1} \mu_t}{\mathbf{1}' \Sigma_t^{-1} \mu_t} \tag{22}$$

In terms of utility functions, either quadratic utility or CARA utility with normal returns, an investor solves the following problem where g is equal to an investor's coefficient of relative risk aversion.

$$\max_{w_t} \ \mu_{p,t} - \frac{g}{2} \sigma_p^2 \ \text{ With a solution of: } w_t = \frac{\Sigma^{-1} \mu_t}{g} \tag{23}$$

Comparing equation 21 with equation 23 we note the inverse relationship between g and $\mu_{p,t}$. As an investor becomes more (less) risk averse they choose a lower (higher) target expected portfolio return. Furthermore, all investors hold weights in the same relative proportions given by the tangency portfolio weights w_{tp} and w_{rf}. Under the usual assumption that all investors agree on the distribution of returns, their portfolios differ only in the amount of wealth they allocate to the risky assets. Using the result from the mutual fund separation theorem we can decompose any efficient portfolio return in terms of weights invested in the tangency portfolio (x) and the risk-free asset ($1 - x$):

$$\mu_{p,t} = (1 - x)R_{f,t} + (x)(\frac{\Sigma^{-1}\mu_t}{\mathbf{1}'\Sigma_t^{-1}\mu_t})'R_t \tag{24}$$

Using the solution to equation 23 this implies that $x = \mathbf{1}'\Sigma_t^{-1}\mu_t/g$. Again we can see the inverse relationship between an investors' risk-aversion, g, and the target expected portfolio return. With these basic equations I will express each of the mean-variance strategies in terms of the weights given by the fundamental result in equation 21. For all of these results we must arrive at some "plug-in" estimate for the unobserved population values of μ_t and Σ_t. Additionally, all of the strategies discussed can be described as a shrinkage estimator of equation 21 and some prior, as well constraints on the weights.

As noted by Demiguel et al. (2009), a prominent role in the vast literature on estimation risk is played by the *Bayesian approach* and in particular, shrinkage estimators.[10] Put simply shrinkage involves a convex linear combination of two estimators, δF + (1 - δ)S, where δ is a number between 0 and 1. This technique is named shrinkage, since the estimator S is *shrunk* toward a more structured estimator, i.e. a prior, F.

Naive Diversification

There is certainly no shortage of proponents backing the use of an equally weighted portfolio. Some of the more recent an compelling evidence can be found in Demiguel et al. (2009), Pflug, Pichler, and Rendek (2012), and Murtazashvili and Vozlyublennaia (2013). Even authors such as Tu and Zhou (2011) and Kirby and Ostdiek (2012) who ultimately argue in favor of Markowitz theory note the strong empirical performance of

[10]For a relevant review of the development and application of shrinkage estimators to finance I refer the reader to Demiguel et al. (2009), Ledoit and Wolf (2003, 2003).

naive diversification. Computing these portfolio weights is simple, the naive strategy holds weights equal to $1/N$ for each of the N risky assets. While the implications of investing in a naive portfolio is an assumption that all moments of returns are equal, it can be expressed as a shrinkage estimator. To do this I consider the tangency portfolio weights given in equation 22 and a shrinkage constant of $\delta = 1$. The prior for the variance-covariance matrix of returns is an identity matrix and the prior for the vector of gross excess returns is a conformable vector of ones. In my tables of results this estimator is labeled as Naive and the weights can be computed as:

$$w_{naive} = \frac{I_N^{-1}\mathbf{1}}{\mathbf{1}' I_N^{-1}\mathbf{1}} \tag{25}$$

Minimum Volatility

The global minimum volatility portfolio plays a special role in portfolio theory as shown by Kirby and Ostdiek (2012). When optimizing over the risky assets only, they show that a two-fund separation theorem still applies where the return on the minimum-variance portfolio takes the place of the risk-free asset. Recent works such as Jagannathan and Ma (2003) and Frahm and Memmel (2010) present strong performance of the global minimum variance portfolio in out-of-sample tests by imposing short-sale constraints and using optimal shrinkage targets. The minimum-variance portfolio can further be seen in the context of the tangency portfolio weights with a shrinkage estimator applied to the vector of excess returns. Simply define a shrinkage constant of $\delta = 1$ and a prior given as a conformable vector of ones. Applying this to the tangency weights in equation 22, the minimum volatility weights, labeled as Minv, can be computed as:

$$w_{Minv} = \frac{\Sigma_t^{-1}\mathbf{1}}{\mathbf{1}'\Sigma_t^{-1}\mathbf{1}} \tag{26}$$

To further reduce estimation risk I also consider a minimum variance model incorporating an extreme version of shrinkage estimation, (Fleming, Kirby, & Ostdiek, 2001, 2003; Kirby & Ostdiek, 2012). These strategies are referred to as "volatility-timing" strategies and they significantly outperform mean-variance efficient portfolios in Kirby and Ostdiek (2012). Specifically Kirby and Ostdiek (2012) restrict the covariance matrix of excess returns to be diagonal thereby requiring no estimation of conditional covariances. In terms of a shrinkage estimator I again use $\delta = 1$ as the shrinkage constant and prior is simply a diagonal matrix with the conditional sample volatility estimates along the main diagonal. In my tables this strategy is labeled, Voltiming, and can be computed as:

$$\hat{w}_{i,t} = \frac{(1/\hat{\sigma}_{i,t}^2)^\eta}{\sum_{1=1}^{N}(1/\hat{\sigma}_{i,t}^2)^\eta} \tag{27}$$

This estimator can also be thought of as a special case of the unlevered risk-parity portfolios introduced by Asness, Frazzini, and Pederson (2012) where portfolio weights are defined as, $w_{i,t} = k_t \sigma_{i,t}^{-1}$. Since this portfolio is unlevered they define $k_t = 1/\sum_i \sigma_{i,t}^{-1}$ and these weights are then same as the Voltiming weights in equation 27. Kirby and Ostdiek (2012) define η as a "tuning" parameter to control timing aggressiveness. When $\eta = 1$ we get the risk-parity weights, as $\eta \to 0$ we get naive diversification, and as $\eta \to \infty$ the least volatile asset gets all the weight.[11]

[11] I use $\eta = 1$ in my analysis although I did experiment with multiple values but found nothing of consequence.

Tangency Portfolio and the Risky Assets Only

The solution for the tangency portfolio weights can be found in equation 22, and this strategy is abbreviated in my tables as TP. Kirby and Ostdiek (2012) show in detail that the study by Demiguel et al. (2009) placed the mean-variance model at a large disadvantage relative to the naive portfolio by focusing on the tangency portfolio. Referencing equation 22 we can see that the tangency portfolio weights imply an expected target portfolio return and Kirby and Ostdiek (2012) show that this target is quite aggressive. To address this issue Kirby and Ostdiek (2012) consider optimization over the risky assets only yielding the following weights:

$$\begin{gathered} w_{OC1N} = X_{TP,t}\frac{\Sigma^{-1}\mu_t}{\mathbf{1}'\Sigma_t^{-1}\mu_t} + (1 - X_{TP,t})\frac{\Sigma_t^{-1}\mathbf{1}}{\mathbf{1}'\Sigma_t^{-1}\mathbf{1}} \\ \text{Where} \quad X_{TP} = \frac{\mu_{p,t} - \mu_{MV,t}}{\mu_{TP,t} - \mu_{MV,t}} \end{gathered} \tag{28}$$

In equation 28 $\mu_{p,t}$ is the target expected portfolio return, $\mu_{MV,t}$ is the expected return on the minimum volatility portfolio, and $\mu_{TP,t}$ is the expected return on the tangency portfolio. To reduce the aggressiveness of the tangency portfolio strategy they set $\mu_{p,t} = \hat{\mu_t}'\mathbf{1}/N$, i.e. the expected return of naive diversification. I also consider a version of this estimator that restricts all weights to be positive while still summing to one. I label the former as OC1N and the latter as OC1Npos. OC1Npos can be computed simply as:

$$w_{OC1Npos,i,t} = \frac{max(w_{OC1N,i,t}, 0)}{\sum_{i=1}^{N} max(w_{OC1N,i,t}, 0)} \tag{29}$$

Reward-to-Risk Timing

Kirby and Ostdiek (2012) introduce another timing strategy that does not ignore estimates of the conditional expected returns called reward-to-risk timing strategies. I refer to this strategy as RRT and the weights can be computed as:

$$w_{\hat{RRT},i,t} = \frac{(\hat{\mu}_{i,t}/\hat{\sigma}^2_{i,t})^\eta}{\sum_{i=1}^N (\hat{\mu}_{i,t}/\hat{\sigma}^2_{i,t})^\eta} \tag{30}$$

I also consider a constrained version of this strategy where all weights are restricted to be positive, but still sum to one. I refer to this strategy as RRTpos and this restriction means that I set the weight to zero for any asset with a negative estimate for the conditional expected return:

$$w_{\hat{RRTpos},i,t} = \frac{max(w_{\hat{RRTpos},i,t})}{\sum_{i=1}^N w_{\hat{RRTpos},i,t}} \tag{31}$$

Kirby and Ostdiek (2012) also implement an alternative estimator of the conditional expected returns implied by the conditional CAPM. In the conditional CAPM the market risk premium is just a scaling factor so all of the cross-sectional variation in conditional expected excess returns come from variation in conditional beta. Therefore they compute weights for this alternative estimator as:

$$w_{\hat{\beta},i,t} = \frac{(\hat{\beta}^+_{i,t}/\hat{\sigma}^2_{i,t})^\eta}{\sum_{i=1}^N (\hat{\beta}^+_{i,t}/\hat{\sigma}^2_{i,t})^\eta} \quad \text{and also} \quad \hat{w}_{\bar{\beta},i,t} = \frac{(\bar{\beta}^+_{i,t}/\hat{\sigma}^2_{i,t})^\eta}{\sum_{i=1}^N (\bar{\beta}^+_{i,t}/\hat{\sigma}^2_{i,t})^\eta} \tag{32}$$

The asymptotic variance of this alternative estimator is lower than then conditional mean for all values of $\beta \neq 0$.[12] Kirby and Ostdiek (2012) extend this idea to multi-factor models and implement this estimator using the Carhart (1997) 4-factor model. In their analysis they

[12]Details can be found in Kirby and Ostdiek (2012).

assume all factors have identical risk premiums and thus their estimator for β in equation 32 becomes $\bar{\beta}^{+}_{i,t} = max(\beta_{i,t}, 0)$ where $\bar{\beta}_{i,t} = (1/K)\sum_{j=1}^{K}\beta_{i,j,t}$. Therefore the conditional expected return proxy for each asset in equation 32 is just the average conditional beta for each of the K factors. I test six different variations of this model. I label the first three FF3m, FF4m, and FF5m to denote conditional β estimates with respect to the Fama and French (1993) 3-factor model, the Carhart (1997) 4-factor model, and the Fama and French (2015) 5-factor model. I label the next three FF3mw, FF4mw, and FF5mw and use the same three factor models but estimate $\bar{\beta}_{i,t}$ differently. Instead of assuming identical risk premiums, I estimate $\bar{\beta}_{i,t}$ as a weighted average where the weights are computed from the absolute values of the estimated time series factor risk premium: $\bar{\beta}_{i,t} = \sum_{j=1}^{K}\beta_{i,j,t}|\bar{r}_k| / \left(\sum_{j=1}^{K}|\bar{r}_k|\right)$.

Parametric Portfolio Choice

In Demiguel et al. (2009) the authors conclude that using information about the cross-sectional characteristics of assets adds substantial value in their portfolio optimization experiment. They derive these results through the implementation of the Brandt et al. (2009) procedure to two of their test datasets. In the conclusion of their paper they recommend expanding on this procedure as a means to improve performance of optimized portfolios. The approach of Brandt et al. (2009) is similar to mean-variance strategies only in that an investor continues to optimize expected utility, but differs sharply in that the joint distribution of asset returns is not estimated directly. Instead, they parameterize the portfolio weights directly as a function of the asset characteristics and maximize expected utility over historical data relative to the parameters, not the weights. This procedure reduces to a relatively simple statistical estimation problem that is implemented using the generalized method of moments estimator (GMM) of Hansen (1982). I will explain the general idea briefly for my application but refer the reader to Brandt et al. (2009) for additional details and discussion. As with mean-variance strategies the investor maximizes expected utility:

$$\max_{(w_{i,t})_{i=1}^{N_t}} \quad E_t[u(r_{p,t+1})] = E_t\left[u\left(\sum_{i=1}^{N_t} w_{i,t} r_{i,t+1}\right)\right]$$
$$w_{i,t} = f(x_{i,t;\theta}) \tag{33}$$
$$f(x_{i,t;\theta}) = \bar{w}_{i,t} + \frac{1}{N_t}\theta' \hat{x}_{i,t}$$

The attractiveness of the procedure lies in the parsimonious parameterization of the $w_{i,t}$; it is particularly well-suited to solve an optimization problem with a large number of stocks, N_t. In the above specification, $\bar{w}_{i,t}$ is the weight of stock i at time t defined by the investor, θ is a vector of coefficients, and $\hat{x}_{i,t}$ are the observed characteristics for each stock standardized cross-sectionally to have zero mean and unit standard deviation at each time period t. Since I am particularly interested in how this technique performs relative to naive diversification, I set $\bar{w}_{i,t} = 1/N_t$. This means that my estimates of θ will determine the deviations from naive diversification and because $\hat{x}_{i,t}$ are standardized, the weights always sum to one. As in their original paper, I also use a power utility function with standard CRRA preferences over wealth of the following form:

$$u(r_{p,t+1}) = \frac{(1 + r_{p,t+1})^{1-\gamma}}{1-\gamma} \tag{34}$$

To put this optimization problem into a GMM framework I set the sample analog of the first-order conditions of the k θs as the moments to be set to zero:

$$\frac{1}{T}\sum_{t=0}^{T-1} g(r_{t+1}, \hat{x}_{i,t}; \theta) = \frac{1}{T}\sum_{t=0}^{T-1} \begin{pmatrix} (1+r_{p,t+1})^{-\gamma}\left(\sum_{i=1}^{N_t}(1/N_t)\hat{x}_{1,i,t}r_{i,t+1}\right) \\ (1+r_{p,t+1})^{-\gamma}\left(\sum_{i=1}^{N_t}(1/N_t)\hat{x}_{2,i,t}r_{i,t+1}\right) \\ \vdots \\ (1+r_{p,t+1})^{-\gamma}\left(\sum_{i=1}^{N_t}(1/N_t)\hat{x}_{k,i,t}r_{i,t+1}\right) \end{pmatrix} \tag{35}$$

I provide additional details on the GMM procedure including statistical inference in section 3. For the explanatory variables I use two different groups of observed characteristics. The first group consists of the level variables log market equity (me), log book-to-market equity (btm), gross profitability (Prof), and investment (Inv). The second group adds the flow variables short-term reversal (Str) and momentum (Mom). These variables are the same characteristics used to generate the stock portfolios most likely to outperform and are described in detail Section 3. I follow the estimation procedure described by Brandt et al. (2009) and initialize my out-of-sample returns with 120 months of historical data. I use the first 120 months of data to estimate the coefficient estimates, and use these estimates to form out-of-sample returns for the next year (12 months). Every subsequent year I reestimate the coefficients by enlarging the sample, and use the estimated coefficients to form the next 12 months of out-of-sample returns.

Empirical Application

Data

Recent studies commonly compare the performance of various portfolio strategies using portfolios of stocks rather than individual stocks. Some examples of this include Kan and

Z. (2007), Demiguel et al. (2009), Kirby and Ostdiek (2012), Kirby and Ostdiek (2015), and DeMiguel et al. (2013). However some other studies on mean-variance efficiency such as Green and Hollifield (1992), Jagannathan and Ma (2003), and Brandt et al. (2009) used individual stock data in their analysis. To more accurately replicate the problem faced by a fund manager I focus on individuals stocks. My sample consists of 8 datasets of individual stocks spanning July 1963 to Dec 2013 with monthly holding period returns obtained from the Center for Research in Security Prices (CRSP) and accounting data from Compustat. The sample includes common equity securities (share codes 10 and 11) for all firms traded on the NYSE, NASDAQ or AMEX (exchange codes 1, 2, and 3) who have a continuous time series of monthly returns for the previous 120 months at any time t.[13] To further enhance the consistency of this experiment with a fund manager's actual portfolio construction dilemma, I consider stocks which are deemed most likely to outperform. To define stocks that are more likely to outperform I use the approach discussed by Dickson (2015) that relies on sequential cross-sectional regressions of key stock characteristics to forecast next period's returns. I consider the following stock characteristics consistent with Dickson (2015): size (log(ME)), book-to-market (log(BE/ME))[14], profitability (ratio of gross profits to assets), past performance measured at horizons of one month ($r_{1,0}$) to capture short-term reversals, and 12 to two months ($r_{12,2}$), to capture momentum, and investment (growth of total assets from previous fiscal year). I provide exact definitions for these variables and their construction in the appendix. As in Novy-Marx (2013), to reduce the effect of outliers, I trim all independent variables at the 1% and 99% levels. I further break-up the

[13]This restriction facilitates the computation of the second moments of stock returns

[14]Taking logs makes the cross-sectional distribution of market equity and book-to-market more symmetric, reducing the impact of outliers

stocks into two groups denoted the "Full Sample" and the "No Micro Sample". Fama and French (2008) define microcaps as stocks with a market value of equity below the 20th percentile of the NYSE market capitalization distribution. Microcaps make up about one half of the stocks on NYSE, AMEX, and NASDAQ, but account for only about 3% of the total market cap. As they note, these small stocks may be less liquid than the representative sample and thus result in above average transaction fees. I estimate all models using the full sample without the microcap stocks to examine the extent to which the performance gains are driven by microcap stocks.

Generating the Expected Returns

To generate the expected returns I follow the approach used by Dickson (2015) and begin with the following specification for the month by month cross sectional regressions:

$$\begin{aligned} r_{i,t+1} = \alpha + \beta_1 ln(BE/ME)_{i,t} + \beta_2 ln(ME)_{i,t} + \beta_3 GPdat_{i,t} + \beta_4 R1to0_{i,t} \\ + \beta_5 R12to2_{i,t} + \beta_6 INV_{i,t} + \epsilon i, t \end{aligned} \tag{36}$$

I estimate this model for each of the 606 months from July 1963 through December 2013. Since the conditional coefficient estimates from month to month are quite noisy, I apply a simple rolling average to the coefficients to filter out the signal. This smoothed estimator takes the following form:

$$\hat{\beta}_t = \frac{1}{t-1} \sum_{i=1}^{t-1} \beta_i$$

Dickson (2015) shows that these regression coefficients are relatively stable over time and that this method significantly reduces estimation and sampling error. To curb excessive turnover and transaction costs, I update the smoothed coefficients annually, specifically every June. As discussed in Section 3 I use two different groups of observed characteristics. The first group consists of the level variables log market equity (me), log book-to-market equity (btm), gross profitability (Prof), and investment (Inv) denoted “No Momentum.” The second group adds the flow variables short-term reversal (Str) and momentum (Mom) denoted “All Variables.” The estimated expected returns from these two models are computed as follows:

$$\hat{r}_{i,t+1} = \hat{\alpha}_s + \hat{\beta}_{1,s} ln(BE/ME)_{i,t} + \hat{\beta}_{2,s} ln(ME)_{i,t} + \hat{\beta}_{3,s} GPdat_{i,t} + \hat{\beta}_{4,s} R1to0_{i,t} + \hat{\beta}_{5,s} R12to2_{i,t} + \hat{\beta}_{6,s} INV_{i,t} \quad \text{(All Variables)}$$

$$\hat{r}_{i,t+1} = \hat{\alpha}_s + \hat{\beta}_{1,s} ln(BE/ME)_{i,t} + \hat{\beta}_{2,s} ln(ME)_{i,t} + \hat{\beta}_{3,s} GPdat_{i,t} + \hat{\beta}_{6,s} INV_{i,t} \quad \text{(No Momentum)}$$

Note that the mean parameter estimates for month t are based on estimates before month t and the stock characteristics are known in month t. This ensures that I have an implementable trading strategy. This procedure should yield a reliable proxy for a universe of stocks that are likely to outperform. Hedge funds and fund managers are generally interested in optimizing over a subset of stocks that they think will do well, not just many different portfolios as used in the aforementioned studies. For my analysis I focus on the top 20% and top 10% of stocks from a sort of these expected returns. Table 10 lists the 8 datasets, the average number of stocks, and the average size of the stocks in each dataset over the time period of July 1973 through December 2013. In the full sample the number of stocks ranges from 88.77 to 177.06 and in the no micro sample the number of stocks ranges from 54.88 to 109.231. In both the full sample and the no micro sample the model excluding the momentum variables picks smaller stocks. For the full sample smallest average size is \$87.25 M for the top 10% of stocks excluding momentum variables and the largest average size is \$697.20 M for the top 20% using all variables. In the no micro sample the smallest average size \$741.24 M for the top 10% of stocks excluding momentum variables and the largest average size is 1904.21 M for the top 20% using all variables.

Methodology for Performance Comparison

To compare the performance of the competing strategies I employ the common rolling-sample approach of Demiguel et al. (2009) and Kirby and Ostdiek (2012). To implement this procedure a historical window of j months is used to estimate the conditional moments of returns and then these estimates are used to generate 1 out-of-sample return that is a function of the actual returns in month $j + 1$. The historical window of j months remains fixed as I iterate forward one month at a time through the end of my sample. Given my time series of 606 monthly returns this procedure yields 606 - j out-of-sample returns. I set $j = 120$ months consistent with both Demiguel et al. (2009) and Kirby and Ostdiek (2012). This choice generates 486 out-of-sample returns where my first out-of-sample return occurs in month 121, i.e. July 1973. Mathematically the rolling estimators take the following form:

$$\begin{aligned} \hat{\mu}_t &= \frac{1}{j}\sum_{i=0}^{j-1} r_{t-i} \\ \hat{\Sigma}_t &= \frac{1}{j}\sum_{i=0}^{j-1}(r_{t-i} - \hat{\mu}_t)(r_{t-i} - \hat{\mu}_t)' \end{aligned} \tag{37}$$

For the performance comparison the reported statistics are computed from the 486 out-of-sample returns. Adjustments for portfolio turnover and trading costs are explained in the next section.

Portfolio Turnover and Trading Costs

Portfolio turnover is an often overlooked but a very real cost to investors. Transactional brokerage fee costs are typically not included in the calculation of a fund's operating expense ratio and thus the true operating expense of high turnover funds can be significant. As long as transaction costs are greater than zero, anything that increases turnover directly reduces the true performance of a fund. The bid/ask spread represents perhaps the largest component of trading costs. In a practical application, the costs of the bid/ask spread would already be directly included in the return of a fund since assets are bought at the ask price and sold at the bid price. However, using CRSP data, the returns are computed from an average of the closing bid/ask spread, therefore not capturing the true costs of the bid/ask spread. To examine the amount of trading required to implement each strategy and approximate these real frictions, I follow Kirby and Ostdiek (2015). Turnover is simply the fraction of invested wealth traded each period needed to re-balance the portfolio to the desired weights. At any time t I calculate turnover as:

$$Turnover_t = \sum_{i=1}^{N} \frac{1}{2} |\hat{w}_{i,t+1} - \hat{w}_{i,t^+}| \tag{38}$$

This definition of turnover is consistent with what is used in the mutual fund industry, i.e. the lesser of the value of purchases or sales in the period divided by the net asset value (Kirby & Ostdiek, 2015). Since there are no fund inflows or outflows these must be equal. I define $\hat{w}_{i,t}$ as the portfolio weight in asset i at time t; $\hat{w}_{i,t^+}$ is the portfolio weight before re-balancing at time $t+1$; and $\hat{w}_{i,t+1}$ is the desired portfolio weight at time $t+1$, after re-balancing. To compute $\hat{w}_{i,t^+}$ I must consider the mechanical changes that occur within the portfolio. Assets that have done well over the time period will make up more than their starting share of weight at the end of the period, and assets that have done poorly will make up less than there starting share. I compute $\hat{w}_{i,t^+}$ as:

$$\hat{w}_{i,t^+} = \frac{\hat{w}_{i,t}(1 + r_{i,t})}{1 + \sum_{i=1}^{N} \hat{w}_{i,t} r_{i,t}} \tag{39}$$

Starting from the beginning of the sample, the first weights occur in month 121, therefore the first turnover calculation occurs in month 122. Studies such as Kirby and Ostdiek (2015, 2012) and Demiguel et al. (2009) do not ignore these mechanical weights while others such as Brandt et al. (2009) do ignore these mechanical changes. I have found that ignoring these mechanical changes is innocuous in this setting but do include them to capture the most conservative view of the trading costs.[15] Now the return of the portfolio net of the proportional transactions costs becomes:

$$r_{p,t+1} = \sum_{i=1}^{N} \hat{w}_{i,t} r_{i,t+1} - 2 \times c_{i,t} |\hat{w}_{i,t} - \hat{w}_{i,t-1^+}|, \tag{40}$$

where $c_{i,t}$ reflects the proportional transaction cost for stock i and time t. Since turnover is the value of assets both purchased and sold as a fraction of total wealth, and both purchases and sales incur transaction costs, I multiply the turnover in Eq. 40 by 2. Novy-Marx and Velikov (2014) find that a momentum based trading strategy had one of the largest time-series average costs of trading in his rigorous analysis of the trading costs of over twenty common anomalies. These costs were estimated at 48.39 basis points per month. While it has been noted by Domowitz et al. (2001) and Hasbrouck (2009) that the cost

[15]For the implementation of the Brandt et al. (2009) approach I do ignore the mechanical changes just as they did in their study.

of trading U.S. equities has declined over time, I want to be as conservative as possible in accounting for these limits to arbitrage. To do so I set $c = 50$ basis points consistent with the conservative measures used by Brandt et al. (2009), Demiguel et al. (2009), Kirby and Ostdiek (2012, 2015), and the even more recent estimates by Novy-Marx and Velikov (2014).

Finally, letting j reference the length of the rolling window (120 months), and T represent the total number of months in the study (606 months), numerically the average turnover I report in the tables is:

$$Turnover = \frac{1}{T-j-1} \sum_{t=j+1}^{T-1} \left(\frac{1}{2} \sum_{i=1}^{N} |\hat{w}_{i,t+1} - \hat{w}_{i,t^+}| \right) \tag{41}$$

Statistical Inference

To conduct statistical inferences about the relative performance of my various strategies using the Sharpe ratio, I follow Kirby and Ostdiek (2012) and use large sample t and $chi-squared$ statistics. I consistently compute these statistics using the generalized method of moments (GMM). For details of the proof of the general results see Hansen (1982). As Hansen (1982) shows, the Delta method, Slutky's theorem and LLN are all used to derive the asymptotic distribution of the GMM estimators. Recent asymptotic distribution derivations for Sharpe ratios are also provided by Opdyke (2007) and Bailey and de Prado (2011) who also use these theorems in their derivations. However I use GMM standard errors to appeal to these more recent derivations while still applicable in a more general context. As with any GMM analysis I begin with a set of moment conditions of the form $E(g(R_t, \theta)) = 0$, where $g(R_t, \theta)$ is a $J \times 1$ vector of moments, analogous to disturbances,

R_t is a vector or returns, and θ is $J \times 1$ vector of parameters. The fundamental result from Hansen (1982) shows that subject to general conditions, the limiting distribution of $\hat{\theta}$ is given by:

$$\sqrt{T}(\hat{\theta} - \theta) \xrightarrow{d} N(0, V) \tag{42}$$

With the following definitions:

$$V = D^{-1}SD^{-1'},\ D = E(\partial g(R_t, \theta)/\theta'),\ S = \sum_{-\infty}^{\infty} E(g(Y_t, \theta)g(Y_{t-j}, \theta)')$$

This definition of the limiting distribution for the parameter estimates also holds for the Brandt et al. (2009) estimation procedure described in Section 3. For this GMM application for Sharpe ratio comparisons the moment conditions are specified as follows:

$$g(R_t, \theta) = \begin{pmatrix} R_{bench,t} - \sigma_{bench} \times SR_{bench} \\ R_{test,t} - \sigma_{test} \times SR_{test} \\ (R_{bench,t} - \sigma_{bench} \times SR_{bench})^2 - (\sigma_{bench})^2 \\ (R_{test,t} - \sigma_{test} \times SR_{test})^2 - (\sigma_{test})^2 \end{pmatrix} \tag{43}$$

Using Eq. 42 I have the asymptotic standard errors of the Sharpe ratios and can now easily conduct a Wald test of linear restrictions to determine if the Sharpe ratios are statistically different. To do so I consider the following test statistic:

$$(\hat{SR}_{test} - \hat{SR}_{bench})(R_{SR}VR'_{SR})^{-1}(\hat{SR}_{test} - \hat{SR}_{bench}) \sim \mathcal{X}(1) \tag{44}$$

In Eq. 44 the discrepancy vector $R_{SR} = (-1, 1, 0, 0)$, V is the asymptotic covariance matrix described in Eq. 42, and the parameter vector $\theta = (SR_{bench}, SR_{test}, \sigma_{bench}, \sigma_{test})$.

Details for the Simulation Experiment

To provide further insight into the severity of estimation risk and turnover I conduct a Monte Carlo analysis using simulated returns. Simulating the returns provides a direct comparison of estimation risk as a function of the number of assets and the length of the estimation window. A similar experiment was conducted by Demiguel et al. (2009) to examine the severity of estimation risk of competing strategies. Demiguel et al. (2009) simulate returns using a single-factor model with a normally distributed risk-free rate. I do not follow this approach but instead simulate returns using a multivariate geometric Brownian motion model to more accurately proxy actual returns. To determine the expected return vector and the variance-covariance matrix I use 492 actual monthly stock returns from Jan 1973 to December 2013. I gathered this data from CRSP using all firms traded on the NYSE, AMEX, or Nasdaq and any firm without a complete time series of returns was removed from the sampling pool. To create excess returns I simulate the risk-free rate as a log-normal random variable where the moments matched the historic mean and standard deviation of the 1-month T-bill over the same time period.[16] I use a log-normal distribution to ensure that the risk-free rate is never negative since it never drops below zero over this 40 year period. For each iteration of the simulation a random time series of N stocks is chosen from the actual stock data file. The expected return vector and variance-covariance matrix of these returns are then computed and a time series of 240 returns are generated according to a multivariate geometric Brownian motion model. I fix the rolling window for moment estimation at 120 months, which also yields 120 months of out-of-sample returns. During

[16]This data was retrieved from the Ken French's data library.

each iteration I compute a Sharpe ratio for the out-of-sample returns as generated by 8 of the asset allocation models. I do not simulate the factor models and so I am restricted to the asset allocation models that only use the moments of the stock returns. For each Monte Carlo experiment I run 1000 iterations of the aforementioned procedure. What makes this experiment so effective is that each of the 1000 iterations is composed of a random group of N stocks estimated from actual stock returns. This means that each simulation is entirely independent and the expected return vector and variance-covariance matrix are allowed to vary from iteration to iteration. The Sharpe ratios for each strategy are then averaged over the 1000 iterations and these are the results presented in Figures 9, 11, and 10. Models were estimated using 5 to 200 assets in increments of 5 such that each strategy consists of 40 data points. By fixing the estimation window at 120 months and varying the number of assets, I can directly asses the relationship between the amount of data used to estimate the moments and the number of assets. I also measure the turnover of each strategy during each iteration and measure Sharpe ratios assuming a constant transaction cost of $c = 50$ basis points. Figures 9, 11, and 10 show the results from this experiment both including transaction costs and not including transaction costs. This distinction allows for the severity of estimation risk and turnover to be viewed separately over a set of independent simulations.

An Optimal Fund of Funds

As shown in Table 3, the average number of stocks in my dataset is much larger than the more common universe of 10 to 25 assets as used in Demiguel et al. (2009), Kirby and Ostdiek (2012), and Kirby and Ostdiek (2015). Without adjustments, traditional mean-variance optimization is not even possible when the number of stocks exceeds the historical estimation window of 120 months, since the variance-covariance matrix is not invertible. This was described in Ledoit and Wolf (2003) by writing the variance-covariance matrix in the following form:

$$S = \frac{1}{T}\mathbf{X}\left(\mathbf{I} - \frac{1}{T}\mathbf{1}\mathbf{1}'\right) \tag{45}$$

Where **X** is an NxT matrix of N net stock returns and **I** is a conformable identity matrix and **1** is a conformable vector of ones. The rank of this variance-covariance matrix is at most equal to the rank of $\mathbf{I} - \mathbf{1}\mathbf{1}'/T$ which is $T - 1$. This means that if the number of returns N is larger than $T - 1$, the variance-covariance matrix is rank deficient and thus not invertible. To address this issue I impose additional structure on the variance-covariance matrix to ensure that mean-variance estimation is possible. This additional structure is imposed by grouping a small number of stocks together to create a pseudo fund that is equal-weighted, and then optimizing over these pseudo funds. By grouping stocks together I am implicitly assuming that all moments of the stocks within each pseudo fund are equal. This fund of funds procedure serves two purposes: 1) as long as the number of funds is less than 120, it ensures that the variance-covariance matrix is invertible, and 2) serves as a shrinkage estimator by equating all moments of the stocks within each pseudo fund. I use the results of the simulation experiment to help guide me on the selection of the number of pseudo funds. The results suggest that 20 pseudo funds should be close to optimal for a rolling estimation window of 120 months.[17] To create the pseudo funds I estimate each stocks' historical sample standard deviation over the same rolling window of 120 months, and then sort these from high to low. I take the total number of stocks in each dataset at each time t, and divide this number by 20, to determine how stocks to hold in each pseudo

[17]More details on these results are available in the next section.

fund. Since this division is likely never exact I create the pseudo funds by making them as equally balanced as possible. Given the evidence of the low-volatility anomaly recently described in Baker, Bradley, and Wurgler (2011) and Andrea and Pederson (2014), I assign fewer stocks to the funds with the lowest volatility, effectively giving those stocks slightly larger weights in the pseudo funds.[18] I then optimize over these pseudo funds, avoiding the problem of a non-invertible variance-covariance matrix. Once I obtain the weights for the pseudo funds, I then apply these to the equally-weighted stocks within each pseudo fund to determine the final weights of each stock and generate the out-of-sample returns.

Results

Simulated Data Results

Figure 9 plots the Sharpe ratios of naive diversification, the actual tangency portfolio (TP) using the known mean vector and known variance-covariance matrix, the estimated TP, the estimated TP targeting naive diversification (TP1N), and the estimated TP targeting naive diversification with short-sale constraints (TP1Np). The top panel does not include trading cost adjustments and therefore provides a clean view of the effects of estimation error as the number of assets increases. As expected, the actual TP has the largest Sharpe ratio in every case since this involves no estimation error and is computed from the exact moments used to simulate the data. The worst performing strategy is the estimated TP which always has the lowest Sharpe ratio. This also means that even with only 5 assets, 120 months is not enough data for the estimated TP to outperform naive diversification. The performance of the two TP strategies targeting naive diversification show that this restric-

[18]For example, if the total number of stocks was 123, division by 20 would leave a remainder of 3. I assign 6 stocks (120/20) to pseudo funds 1 through 17 and 7 stocks to pseudo funds 18 - 20.

tion enhances the empirical performance of TP strategies. Furthermore, the TP strategy targeting naive diversification with short-sale constraints does just as good as the naive strategy, corroborating evidence from Jagannathan and Ma (2003), Demiguel et al. (2009) *et al.* that short-sale constraints improve empirical performance.

The plots in the bottom panel show the large reduction in performance due to trading costs. The Sharpe ratio of the estimated tangency portfolio drops below zero after only 35 stocks and along with the TP portfolio targeting naive diversification, it drops as low as -0.6 for 200 stocks. The turnover is not nearly as severe for the TP strategy targeting naive diversification with short-sale constraints suggesting that constraints not only reduce estimation error but also reduce turnover.

To determine the number of pseudo funds used in my optimal funds of funds analysis I examine this figure by visual inspection. The performance of the estimated tangency portfolio and its extensions begins to deteriorate after about 20 stocks. However, given these plots I tested a variety of pseudo funds and found the choice to be rather innocuous. Therefore I used 20 funds for the remainder of the analysis.

Figure 10 plots the Sharpe ratios of naive diversification, the actual TP, the estimated TP, RRT, and RRTpos. The results show that the restrictions imposed by Kirby and Ostdiek (2012) do achieve the desired results. Examining the plots both before and after trading costs reveals commensurate performance with naive diversification. While 120 months is still not long enough for these strategies to outperform naive diversification, at least it isn't so short that they vastly under perform.

Finally Figure 11 plots naive diversification, the actual global minimum variance portfolio using the known variance-covariance matrix, the estimated global minimum variance portfolio (MinV), and the volatility-timing portfolio (Volt). As expected the actual MinV portfolio has the largest Sharpe ratios. Further inspection of the figure shows that the VT portfolio is highly effective at curbing both estimation risk and turnover. The VT portfolio results are commensurate with the naive portfolio even after incorporating trading costs. The MinV Sharpe indicates a severe loss in performance due to both estimation risk and turnover. After incorporating trading costs the Sharpe ratio for the MinV portfolio turns negative after 90 stocks.

The general conclusions from this experiment show that the models recently proposed by Kirby and Ostdiek (2012) were successful in improving out-of-sample performance. However, the results also suggest that the performance of naive diversification is relatively strong and will be hard to beat with actual stock data.

Empirical Dataset Results

Full Sample Results

From Table 10 I find that both the number of stocks is larger and their average size smaller, for the full sample stocks compared to the no micro sample. I first examine the results for the full sample that were generated from my model excluding the momentum variables, Datasets 1 and 2. Judging by the small average firm size from Table 10, it is clear that the model used to create the datasets loads heavily on the small-cap anomaly. Referencing Table 12, no competing strategy has a statistically larger Sharpe ratio than naive diversification. The only statistically significant result is the under performance of the RRT strategy. The p-value of this difference is 0.0086 with an estimated Sharpe ratio after transaction costs of 0.2511. By comparison the Sharpe ratio of naive diversification is 0.7402. The largest Sharpe ratio after transaction costs is 0.7743 for Voltiming strategy but this is only significant at the 15% level. All of the factor model RRT strategies have marginally larger Sharpe ratios than naive diversification before transaction costs, but these gains disappear after incorporating trading costs. The turnover for these RRT strategies is on the order of 20% to 60% larger than the naive strategy, which is enough to reduce after transaction fee performance below that of the naive strategy. Consistent with prior literature, the TP portfolio performs poorly and has extreme weights and turnover. The expected return is 631%, the volatility is 3,454.6%, and the portfolio turnover is over 4600% per year. After transaction costs the Sharpe ratio is barely above zero with a value of 0.0232. The OC1N strategy does handedly beat the TP strategy, but it also has the second largest turnover, over 81% per year. The OC1Npos is much more successful at producing commensurate performance with naive diversification before transaction fees, but trading costs are much larger and erode all gains.

Referencing Table 13 the competing strategies do not fare as well as they did in the larger universe of stocks from Dataset 1. No Sharpe ratios are statistically different than the naive strategy and both the naive and Voltiming strategies have the largest raw Sharpe ratios after trading costs of 0.7313 and 0.7315 respectively. The Voltiming strategy has a Sharpe ratio before transaction fees of 0.8001 compared to 0.7844 for naive diversification. However, Voltiming also has a slightly higher turnover, 0.0772 compared 0.0647, such that after transaction fees the results are indistinguishable. In both Tables 12 and 13 the short-sale constraints greatly enhance performance through a reduction in estimation risk and turnover. As an example the turnover of the RRT strategy falls from 426% per year to only 26.75% per year after imposing short-sale constraints. The BSV strategy produces results in both of these samples similar to the RRT factor models, in neither case beating naive diversification.

I turn my attention now to the results in Tables 14 and 15 which are the datasets containing all variables. From Table 10 these models pick much larger stocks, although the average size is still relatively small, \$697.2 M in the largest case. Not only does this models pick larger stocks, it also contributes substantially to an increase in turnover. Even for naive diversification turnover increases from 5.83% to 34.86% per year. Referencing Table 14 no Sharpe ratios are statistically different than the naive strategy. Additionally, the naive strategy has the largest raw Sharpe ratio after trading costs of 0.7616. Including these momentum variables greatly increases the Sharpe ratio before transaction costs, 0.9924 compared to 0.7897 for naive diversification. However, the increase in turnover erodes nearly all gains such that the after transaction cost performance is about the same as the datasets excluding the momentum variables. With the exception of the TP strategy, which still produces extreme results, the naive strategy also produces the largest excess mean return after transaction fees of 14.49% per year. The BSV strategy exhibits much higher turnover in this sample as compared to Datasets 1 and 2, over 130% larger than the RRT factor models. While this model does yield the largest Sharpe ratio before transaction fees with a value of 1.079, after trading costs the Sharpe ratio falls to 0.610, well below the value of naive diversification.

Table 15 tells much the same story as Table 14. The turnover is slightly larger, for the naive strategy as it increased to 39.62% per year from 34.86% per year. However this is a function of the model used for stock selection and the fact that the this is a smaller sample of stocks, i.e. the top 10% compared to top 20%, and not a result of the strategy performance. Again the only statistically significant result is the under performance of RRT. The p-value of this difference is 0.0300 with an estimated Sharpe ratio after transaction costs of 0.1263 compared to the value of naive diversification of 0.8240. The Voltiming strategy has a marginally higher Sharpe ratio after transaction fees of 0.8334 but the p-value for this difference is 0.1572. The largest excess mean return is 26.15% per year for the BSV model but the turnover for this model is large, 88.21% year, such that after transaction fees this strong result disappears. As before the short-sale constraints improve performance by reducing estimation risk and reducing turnover.

Overall the results from these datasets show that naive diversification is hard to beat. Naive diversification routinely has one of the largest mean returns, largest Sharpe ratios, and smallest measures of turnover. By comparison the Voltiming strategy is the most competitive with the naive strategy routinely having one of the smallest volatility measures, largest Sharpe ratios, and smallest measures of turnover. The mean-variance extensions generally perform poorly with the exception of the short-sale constrained portfolio. In all cases the RRT with factor models outperform the RRT strategies both before and after transaction fees, and have lower turnovers in all cases. However, the choice of factor model and whether or not I assume equal or unequal risk premiums is innocuous. Finally, the BSV model performs unexpectedly quite poor considering the evidence in Demiguel et al. (2009) that suggest strategies of this type represent "a promising direction to pursue" (p.1923). I suspect the disappointing performance of this strategy is due to the characteristics of the datasets. Recall that these data were created by exploiting the cross-section of expected returns using the sequential cross-sectional regression methodology of Dickson (2015). Therefore, the same stock characteristics used in the parametric portfolio choice optimization algorithm were also used to pick these top performing stocks in the first place. So any gains from exploiting the cross-sectional stock characteristics were already realized when forming the datasets.

No Micro Sample Results

Table 10 shows that the average stock size is much larger for the no micro sample compared to the full sample. The largest average size for the no micro sample is $1904.18 M compared to $697.20 M for the full sample. The average number of firms however is smaller with the largest measured at 109.21 firms compared to 177.06 firms for the full sample. In Table 16 the largest Sharpe ratio after trading costs is 0.7019 for Voltiming compared to 0.6719 for the naive strategy, although the p-value for the difference between them is 0.2064. The largest excess mean return is 14.39% for the BSV model but the corresponding volatility is also the highest on the whole table with the exception of the TP strategy. The BSV strategy also results in the lowest turnover, only 5.18% per year, but the high volatility measure yields in an unimpressive Sharpe ratio of 0.6215. The lowest volatility measure is achieved by the MinV strategy but this strategy also has a low annual return of 8.55% and a large turnover of 78.52% per year. All mean-variance strategies and RRT strategies perform poorly. However, the results for the factor model RRT strategies are comparable to naive diversification.

Table 17 displays results that are much the same. Voltiming has a marginally larger Sharpe ratio, although statistically insignificant, with a value of 0.6373 compared to 0.6121 for the naive strategy. Again BSV has the largest mean but a relatively high volatility. All mean-variance strategies and RRT strategies perform poorly and again the results for the factor model RRT strategies are comparable to naive diversification.

I now turn my attention to the results in Tables 18 and 19 which are the datasets containing all variables. As in the full sample these models pick much larger stocks with a reasonable average size of $1904.12 M and $1736.57 M. Also as in the full sample, this model contributes substantially to an increase in turnover. Even for naive diversification turnover increases from 6.27% to 36.79% per year. Referencing Table 18 the largest Sharpe ratio after trading costs is 0.5880 for the FF5m strategy and this result is statistically different than naive diversification at the 1% significance level, i.e. a p-value of 0.0044. For naive diversification the Sharpe ratio is 0.5655 so this marginally larger measure for FF55m is not economically meaningful. Two other models, FF3m and FF5mw, also have statistically significant Sharpe ratios that are marginally larger than the naive strategy. While not significant, the naive strategy yields a larger Sharpe ratio than the Voltiming strategy whose value is 0.5570. The BSV strategy again has the largest mean return of 19.11%, with the exception of the extreme TP strategy with a value of 68.63%. Even though this table yields some significant results, the economic content is weak as the annualized Sharpe ratios are all within 2% of the value of naive diversification.

Table 19 reinforces the conclusions from Table 18. The FF5m strategy produces a statistically significant Sharpe ratio compared to naive diversification. The value is 0.6096 compared to 0.5820 with a p-value of 0.0399. The Sharpe ratio for Voltiming is smaller than naive diversification with a value of 0.5700. I also observe that again the naive strategy has the smallest turnover and the BSV strategy has the largest excess mean return, 21.95% per year. However, all results are economically weak as the best performing strategies have annualized Sharpe ratios all within 2% of the value of naive diversification.

The results from the no micro sample reinforce what was found in the full sample, i.e. naive diversification is hard to beat. Overall these results are economically weak as all of the top performing Sharpe ratios are close in value to naive diversification. The BSV model continues to show disappointing results despite its praise by Demiguel et al. (2009). In the next section I present some evidence for why these new models proposed by Kirby and Ostdiek (2012) have failed to outperform naive diversification even though they did so in the author's original analysis.

Cross-sectional Dispersion Statistics of the Datasets

In Kirby and Ostdiek (2012) the authors attest that the performance of volatility timing strategies is driven by the cross-sectional (CS) dispersion in conditional volatility and the performance in RRT strategies is related to the CS dispersion in conditional expected excess returns (p. 464). The authors simply report the range of volatilities and means to make this claim. I investigate this claim within my datasets as well but I measure the CS dispersion differently. Instead of the CS range, I present the CS mean absolute deviation. That is, at each t I measure the following mean absolute deviation of the sample statistic λ.

$$MAD_\lambda = E[|\lambda_j - \bar{\lambda}|] \tag{46}$$

I report the time series average of this statistic for my datasets and also for the three of the datasets used in Kirby and Ostdiek (2012) to facilitate comparisons with their study. These results are presented in Table 20. Kirby and Ostdiek (2012) find that the RRT strategies do poorly in the 10 Industry dataset but much better in the Momentum and Size/BTM portfolios. For the Voltiming strategies performance is weak in both the 10 Industry and

Size/BTM portfolios but stronger in the Momentum dataset. My measure of CS dispersion does in fact match the claims made in Kirby and Ostdiek (2012). Namely that the poor performance of the RRT strategies in the 10 Industry dataset is due to the small cross-sectional dispersion in the conditional expected excess return. My CS MAD does show that both the CS dispersion in conditional expected excess return and Sharpe ratios are about 20% lower than that of the Momentum and Size/BTM portfolios from Kirby and Ostdiek (2012). Using these CS MAD measures as a benchmark, I can compare the 8 additional datasets constructed in this paper to help explain the results. In all 8 of the datasets I see the same pattern. The CS MAD of the conditional expected excess return is smaller than both the Momentum and Size/BTM portfolios, the CS MAD of the conditional volatilities are as high as 50 % larger than both the Momentum and Size/BTM portfolios, and the CS MAD of the conditional Sharpe ratios are smaller than all three datasets from Kirby and Ostdiek (2012). So to answer the question, why do the VT strategies perform relatively well in these dataset, I point to the large CS MAD of the conditional volatilities of my 8 datasets. To answer the question, why did the RRT strategies fail to outperform naive diversification, I point to the small CS MAD of the conditional *Sharpe ratios*, not the conditional expected returns. My claim still supports the comments by Kirby and Ostdiek (2012), as they only stated that the performance was *related* to the conditional expected return, not that this was the only reason.

This discussion presents interesting conclusions for portfolio analysis using individual stock data. That is, optimal portfolios formed from individual stocks may fail to provide a large enough CS dispersion in the conditional *Sharpe ratios* for mean-variance extensions to outperform naive diversification. Of course this may not always be true but it does appear to be true here. This small dispersion may be due to the fact that portfolios of stocks are less diversified and thus have a larger dispersion in non-priced idiosyncratic risk, manifesting itself into smaller Sharpe ratios. Another explanation could be that the small dispersion in conditional Sharpe ratios is driven by the fact that the stocks in my datasets where all predicted to outperform, and therefore share similar conditional moment estimates. Regardless of the interpretation adopted by the reader, this analysis does provide a pattern between a measure, CS MAD, and portfolio strategy performance. Additionally, these results should serve as a warning to investors when attempting to optimize over a portfolio of stocks as compared to optimizing over a set of portfolios. Simply put, naive diversification is hard to beat.

Commentary on the BSV Model Loadings and Portfolio Weights

Table 21 presents statistics related to the loadings on the stock characteristics as well as the portfolio weights for the Brandt et al. (2009) model. In their original paper, these authors focused on the key stock characteristics of market equity, book-to-market equity, and momentum to maximize an investor's expected utility. By examining the parameter estimates of the stock characteristics the authors find that the investor overweights small firms, value firms, and past winners (p.3429). Overall, all of my results are consistent with this but I also use additional characteristics not included in their study. For the full sample, Datasets 1 and 2, my results indicate that investors overweight small firms, value firms, highly profitable firms, and firms with low investment. These results are consistent with the expected return relationships presented by Fama and French (2015) for their 5-factor asset pricing model. Analysis of the standard errors reveals that book-to-market is approximately 3 standard errors from zero, size is 2.5 standard errors from zero, but both gross-profitability and investment are less than 2 standard errors from zero. For the no micro sample, in Datasets 5 and 6, my results again indicate that investors overweight small firms, value firms, highly profitable firms, and firms with low investment. Compared to the full sample dataset, the loadings on size are much larger and more significant. For example, in Dataset 1 the loading on size is -2.68 with a standard error of 1.04 while in Dataset 5 the loading on size is -5.98 with a standard error of 1.37. This result indicates a larger sensitivity towards the small firm effect in the no micro sample. The average firm size in Dataset 5 is six times larger than that of Dataset 1, which is likely the cause of the larger loadings on the small firm effect.

Turning my attention to the datasets formed using all variables the interpretations are similar. For the full sample, Datasets 3 and 4, investors overweight small firms, value firms, highly profitable firms, firms with low investment, recent losers, and longer term winners. However, likely due to the inclusion of the flow variables R1to0 and R12to2, the standard errors for size, profitability, and investment are now much smaller and no longer significant at standard confidence levels. The magnitude and significance of book-to-market is still preserved and both R1to0 and R12to2 appear highly significant. For the no micro sample in Datasets 7 and 8, I see similar interpretations. As was the case without the flow variables, the no micro sample indicates a large sensitivity towards the small firm effect, the estimate is -3.73 with a standard error of 1.44 for Dataset 7 and -2.58 with a standard error of 1.04 in Dataset 8.

The largest maximum weight in any individual stock is 0.248 for Dataset 8 and the smallest minimum weight in any individual stock is -0.169 in Dataset 6. The average weight is always larger for the no micro samples for the top 10% of stocks sorted be expected return, simply because these datasets have the smallest number of stocks. Most importantly these weight statistics show that extreme positions, and thus extreme deviations from equal weights, is not an issue.

Conclusions

I conducted a horse-race of 15 different portfolio construction techniques using individual stock data. The data was constructed from an implementable trading strategy that yields stocks that are most likely to outperform. My results indicate that naive diversification consistently produces one of the largest out-of-sample mean returns, largest Sharpe ratios, and smallest turnover measures. Through a robust simulation experiment, I validate that the mean-variance extensions developed by Kirby and Ostdiek (2012) do indeed reduce estimation risk and turnover. Using these extensions with actual stock data however, these improvements in performance are not large enough to consistently top naive diversification using a traditional estimation window of 120 months. Confirming the claims made by Kirby and Ostdiek (2012), I conclude that this lack of performance is driven by the characteristics of the data. I introduce a statistic, the time series average of the cross-sectional mean absolute deviation of risk and return, to reinforce this claim. Specifically I find that my datasets have larger dispersions in cross-sectional volatility and smaller dispersions in cross-sectional Sharpe ratios than the data analyzed by Kirby and Ostdiek (2012). These facts explain why the volatility timing strategies fare well in my samples but the RRT strategies fail to perform as well as they did the data used by Kirby and Ostdiek (2012). Since Demiguel et al. (2009) conclude that the approach proposed by Brandt et al. (2009) shows the most promise, I also test various extensions of this strategy. However, after considering transaction costs, this strategy does not perform nearly as well as naive diversification. I attest that this finding is also a function of the data analyzed. My results add to the mounting evidence that naive diversification is hard to beat, particularly in a universe of stocks that are likely to outperform. These findings have important implications for fund managers and practitioners as these test datasets closely mimic a fund manager's portfolio construction dilemma. An obvious extension of this study would involve providing concrete recommendations for when, and when not, to apply optimization techniques. For example, even the seemingly successful RRT strategies fail to consistently beat naive diversification when the CS MAD of the data's Sharpe ratio falls below 0.04, using a 120 month estimation window for 20 funds. A complete analysis of many different combinations like this would be quite useful. Consistent with Demiguel et al. (2009), my evidence suggest that practical optimization techniques have a long way to go before they can be expected to do well in applications with limited data.

Appendix

Data Description

For all accounting variables we employ the standard fiscal year matching popularized by Fama and French (1992). The accounting variables for fiscal years that end in calendar year t are matched with stock returns for July of year $t + 1$ to June of year $t + 2$. So there is at least a six month lag for the accounting variables in each monthly cross-sectional regression.

1. log(ME): Market equity is defined as price per share times shares outstanding from

CRSP. To get ME for the firm, we aggregate values of all equity for a given permno and date. This aggregate value is assigned to the permno with the largest ME. Slightly deviating the Novy-Marx (2013) definition we update using the June market equity to compute this variable rather than the previous December value since his increases its explanatory power in the cross section.

2. log(BE/ME): Book-to-market is book equity scaled by market equity. Book equity is shareholder equity, plus deferred taxes, minus preferred stock, when available. The shareholder equity components follow the tiered definitions consistent with those used in Fama and French (1993) to construct the HML factor. Stockholder equity is defined in Compustat as (SEQ) if available, or else common equity plus the carrying value of preferred stock is available (CEQ + PSTX) if available, otherwise total assets minus total liabilities (AT - LT) is used. Deferred taxes is deferred taxes and investment tax credits (TXDITC) if available, or else deferred taxes and/or investment tax credit (TXDB and/or ITCB). Preferred stock is redemption value (PSTKR) if available, or else liquidating value (PSTKRL) if available, or else carrying value (PSTK).

3. GPdat: Gross profits and earnings before extraordinary items are Compustat data items GP and IB, respectively. For free cash flow we employ net income plus depreciation and amortization minus changes in working capital minus capital expenditures (NI + DP - WCAPCH - CAPX). Gross profits are also defined as total revenue (REVT) minus cost of goods sold (COGS).

4. Inv: Investment for firms in year t is the growth of total assets for the fiscal year

ending in year $t-1$ divided by total assets at the end of year $t-2$. This matches the definition used by (Fama & French, 2015). In their valuation equation, the investment variable is actually defined as the expected growth of book equity, not assets. However, as they state, sorts on asset growth result in larger spreads of average return and using growth in book equity produces similar results.

5. R1to0: The short-term reversal measure is simply the return at time t, lagged by one period.

6. R12to2: The momentum measure is the previous year's 11 month return, skipping the previous month to prevent capturing short-term reversal in the momentum measure.

Table 10: List of datasets

All data spans the same time period from July 1963 to Dec 2013, a total of 606 months. I also use the same rolling estimation window of 120 months so each dataset has 486 out-of-sample months. Each data source is denoted either Full Sample or No-Micro sample. For the Full Sample all stock data with complete information was included while the No-Micro sample excluded all stocks below the 20% Percentile of Market Equity for stocks on the NYSE. The All Variables data source designation means all variables were used in the model to pick the top performing stocks. The No Momentum data source designation means all variables, excluding R1to0 and R12to2, were used in the model to pick the top performing stocks.

Data Source	Top %	Avg. # Stocks	Avg. Size	Abbreviation
Panel A: Full Sample				
No Momentum	20%	177.06	155.45 M	Dataset 1
No Momentum	10%	88.77	87.25 M	Dataset 2
All Variables	20%	177.06	697.20 M	Dataset 3
All Variables	10%	88.77	541.54 M	Dataset 4
Panel B: No Micro Sample				
No Momentum	20%	109.21	921.18 M	Dataset 5
No Momentum	10%	54.88	741.24 M	Dataset 6
All Variables	20%	109.21	1904.12 M	Dataset 7
All Variables	10%	54.88	1736.57 M	Dataset 8

Table 11: List of all asset allocation models and a brief description

Model Description	Abbreviation
Panel A: Naive Diversification	
Equally weighted portfolio	Naive
Panel B: Minimum Volatility	
Volatility Timing	Voltiming
Traditional global minimum variance portfolio	MinV
Panel C: Mean Variance Extensions	
Tangency Portfolio, i.e. Maximum Sharpe Ratio	TP
Optimization over risky assets only targeting Naive diversification	OC1N
OC1N with no shorts allowed	OC1Npos
Panel D: Reward-to-Risk Timing	
Reward-to-Risk Timing with no constraints	RRT
Reward-to-Risk Timing with no shorts allowed	RRTpos
Panel E: Reward-to-Risk Timing with Factor Models	
RRT using Fama and French 3-factor model with equal risk-premiums	FF3m
RRT using Fama and French 4-factor model with equal risk-premiums	FF4m
RRT using Fama and French 5-factor model with equal risk-premiums	FF5m
RRT using Fama and French 3-factor model with un-equal risk-premiums	FF3mw
RRT using Fama and French 4-factor model with un-equal risk-premiums	FF4mw
RRT using Fama and French 5-factor model with un-equal risk-premiums	FF5mw
Panel F: Parametric Portfolio Choice	
The Brandt et al. (2009) model using various characteristics	BSV

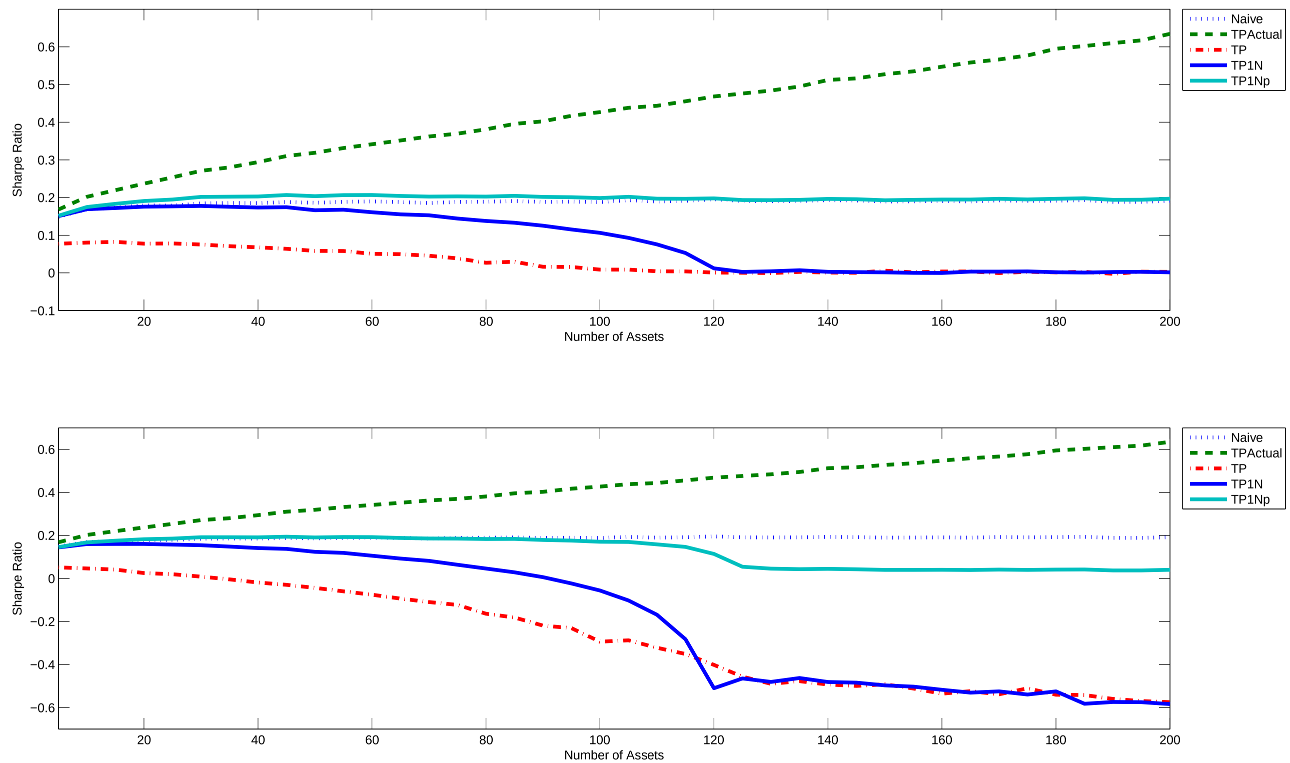
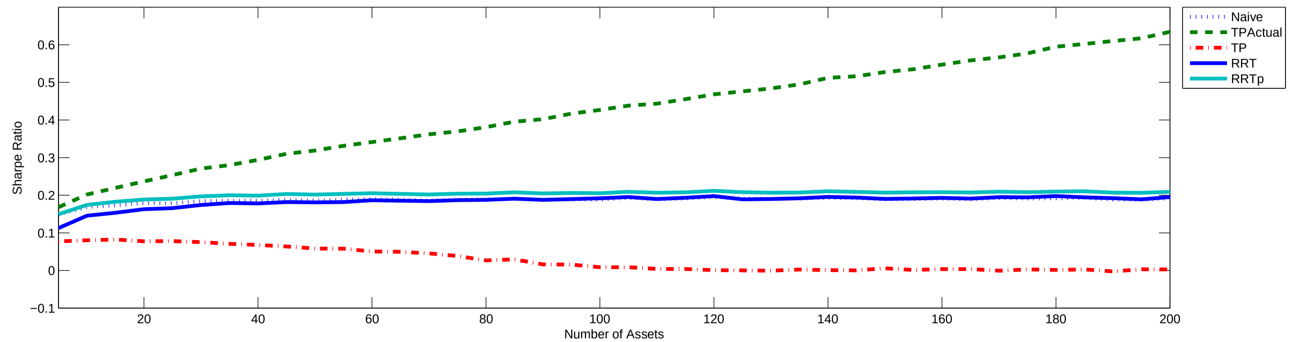


Figure 9: Monte Carlo simulations for tangency portfolio
The top figure plots Monte Carlo simulations for out-of-sample Sharpe ratios for the following portfolio strategies: Naive diversification, the true Tangency Portfolio (TPActual), the Tangency Portfolio (TP), TP targeting Naive diversification, and TP targeting Naive diversification with short-sale constraints. The rolling estimated window length was 120 months and the out-of-sample window was 240 months. Models were estimated using 5 to 200 assets in increments of 5 such that each strategy consists of 40 data points. The bottom figure plots adjustments for estimated transactions costs of c = 50 basis points. There were 1,000 simulations ran for each data point.



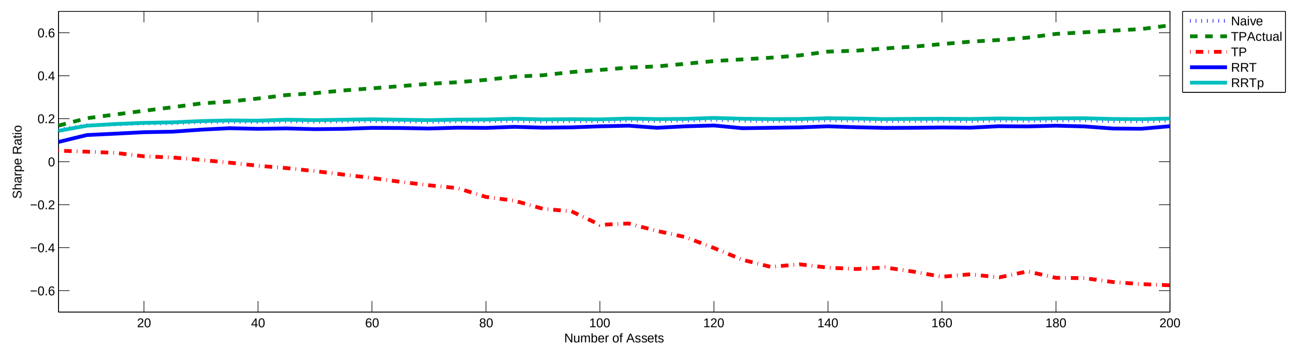
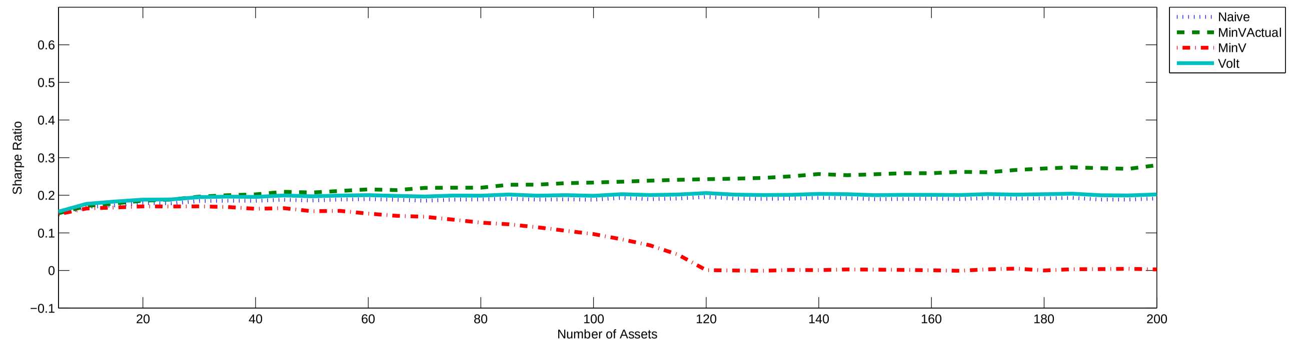


Figure 10: Monte Carlo simulations for tangency portfolio and rrt

The top figure plots Monte Carlo simulations for out-of-sample Sharpe ratios for the following portfolio strategies: Naive diversification, the true Tangency Portfolio (TPActual), the Tangency Portfolio (TP), Reward-to-Risk timing (RRT), and Reward-to-Risk timing (RRTp) with short-sale constraints. The rolling estimated window length was 120 months and the out-of-sample window was 240 months. Models were estimated using 5 to 200 assets in increments of 5 such that each strategy consists of 40 data points. The bottom figure plots adjustments for estimated transactions costs of c = 50 basis points. There were 1,000 simulations ran for each data point.



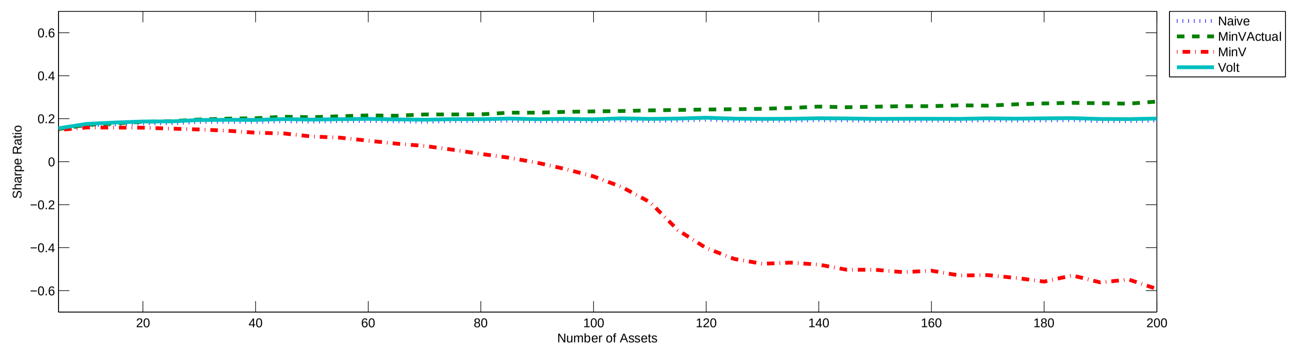


Figure 11: Monte Carlo simulations for minimum variance and volatlity timing
The top figure plots Monte Carlo simulations for out-of-sample Sharpe ratios for the following portfolio strategies: Naive diversification, the true Minimum Variance portfolio (MinVActual), the Minimum Variance portfolio (MinV), and the Volatility Timing portfolio (VT). The rolling estimated window length was 120 months and the out-of-sample window was 240 months. Models were estimated using 5 to 200 assets in increments of 5 such that each strategy consists of 40 data points. The bottom figure plots adjustments for estimated transactions costs of c = 50 basis points. There were 1,000 simulations ran for each data point.

Table 12: Annualized traditional performance statistics for dataset 1

Dataset 1: Full Sample, Top 20 %, 20 Funds, η = 1. Annualized traditional performance statistics for monthly out-of-sample returns for portfolios formed by stocks sorted by expected return. To generate the expected returns a model of the form $r_{i,t+1} = \alpha + \beta_1 ln(BE/ME)_{i,t} + \beta_2 ln(ME)_{i,t} + \beta_3 GPdat_{i,t} + \beta_4 R1to0_{i,t} + \beta_5 R12to2_{i,t} + \beta_6 INV_{i,t}$ was estimated and the estimates were averaged over an estimation window initialized with 60 months of data. Each subsequent time period the model was re-estimated by enlarging the sample and the average estimates were updated every June. These average estimates were used generate the expected returns. All averaged estimates at time period t were generated using estimates prior to time period t. $R1to0_{i,t}$ and $R12to2_{i,t}$ were used only as controls and thus β_4 and β_5 were not used to generate expected returns.

Model	$\hat{\mu}_{ex}$	$\hat{\sigma}$	SR	$\hat{\mu}_{TC}$	SR$_{TC}$	SR$_p$	Turnover
Panel A: Naive Diversification							
Naive	14.0242	17.7590	0.7897	13.1445	0.7402	1.0000	0.0583
Panel B: Minimum Volatility							
Voltiming	12.7162	15.1311	0.8404	11.7154	0.7743	0.1504	0.0695
Minv	7.7567	12.1334	0.6393	-0.0403	-0.0033	0.6965	0.6396
Panel C: Mean Variance Extensions							
TP	631.0075	3454.5941	0.1827	80.0568	0.0232	1.0000	46.4002
OC1N	7.5447	12.4627	0.6054	-2.3540	-0.1889	0.5472	0.8174
OC1Npos	11.0118	13.4950	0.8160	7.7856	0.5769	0.8830	0.2569
Panel D: Reward-to-Risk Timing							
RRT	17.9146	43.4881	0.4119	10.9183	0.2511	0.0086	0.5723
RRTpos	12.3493	15.8516	0.7791	9.8787	0.6232	0.9579	0.1932
Panel E: Reward-to-Risk Timing with Factor Models							
FF3m	13.3840	16.5048	0.8109	12.3956	0.7510	0.2467	0.0681
FF4m	13.3311	16.4736	0.8092	12.2807	0.7455	0.3739	0.0734
FF5m	13.1661	16.4124	0.8022	11.8778	0.7237	0.7251	0.0941
FF3mw	13.3253	16.3676	0.8141	12.3001	0.7515	0.2132	0.0712
FF4mw	13.2770	16.3275	0.8132	12.0687	0.7392	0.3718	0.0869
FF5mw	13.1467	16.2555	0.8088	11.9225	0.7334	0.4824	0.0882
Panel F: Parametric Portfolio Choice							
BSV Model	15.6206	20.0793	0.7779	14.5671	0.7255	0.7689	0.0620

Table 13: Annualized traditional performance statistics for dataset 2

Dataset 2: Full Sample, Top 10 %, 20 Funds, η = 1. Annualized traditional performance statistics for monthly out-of-sample returns for portfolios formed by stocks sorted by expected return. To generate the expected returns a model of the form $r_{i,t+1} = \alpha + \beta_1 ln(BE/ME)_{i,t} + \beta_2 ln(ME)_{i,t} + \beta_3 GPdat_{i,t} + \beta_4 R1to0_{i,t} + \beta_5 R12to2_{i,t} + \beta_6 INV_{i,t}$ was estimated and the estimates were averaged over an estimation window initialized with 60 months of data. Each subsequent time period the model was re-estimated by enlarging the sample and the average estimates were updated every June. These average estimates were used generate the expected returns. All averaged estimates at time period t were generated using estimates prior to time period t. $R1to0_{i,t}$ and $R12to2_{i,t}$ were used only as controls and thus β_4 and β_5 were not used to generate expected returns.

Model	$\hat{\mu}_{ex}$	$\hat{\sigma}$	SR	$\hat{\mu}_{TC}$	SR$_{TC}$	SR$_p$	Turnover
Panel A: Naive Diversification							
Naive	14.4194	18.3819	0.7844	13.4436	0.7313	1.0000	0.0647
Panel B: Minimum Volatility							
Voltiming	12.9125	16.1377	0.8001	11.8054	0.7315	0.8175	0.0772
Minv	7.3010	14.4615	0.5049	0.9886	0.0684	0.1924	0.5046
Panel C: Mean Variance Extensions							
TP	602.6906	4148.4942	0.1453	-92.5350	-0.0223	1.0000	57.8546
OC1N	6.6876	14.5919	0.4583	-0.6924	-0.0474	0.1050	0.6001
OC1Npos	10.5267	14.4282	0.7296	7.5648	0.5243	0.7047	0.2301
Panel D: Reward-to-Risk Timing							
RRT	43.5196	135.8104	0.3204	-7.5024	-0.0552	0.9951	4.2783
RRTpos	13.5133	17.3453	0.7791	10.2089	0.5886	0.9944	0.2675
Panel E: Reward-to-Risk Timing with Factor Models							
FF3m	13.9001	17.5821	0.7906	12.7252	0.7238	0.8867	0.0821
FF4m	13.7946	17.5143	0.7876	12.5458	0.7163	0.9759	0.0886
FF5m	13.6634	17.5601	0.7781	12.0323	0.6852	0.9537	0.1206
FF3mw	13.7851	17.4831	0.7885	12.5576	0.7183	0.9565	0.0864
FF4mw	13.5838	17.3326	0.7837	12.1466	0.7008	0.9992	0.1051
FF5mw	13.5879	17.4065	0.7806	12.0531	0.6925	0.9819	0.1126
Panel F: Parametric Portfolio Choice							
BSV Model	15.7707	20.9188	0.7539	14.6125	0.6985	0.5780	0.0711

Table 14: Annualized traditional performance statistics for dataset 3

Dataset 3: Full Sample, Top 20 %, 20 Funds, η = 1. Annualized traditional performance statistics for monthly out-of-sample returns for portfolios formed by stocks sorted by expected return. To generate the expected returns a model of the form $r_{i,t+1} = \alpha + \beta_1 ln(BE/ME)_{i,t} + \beta_2 ln(ME)_{i,t} + \beta_3 GPdat_{i,t} + \beta_4 R1to0_{i,t} + \beta_5 R12to2_{i,t} + \beta_6 INV_{i,t}$ was estimated and the estimates were averaged over an estimation window initialized with 60 months of data. Each subsequent time period the model was re-estimated by enlarging the sample and the average estimates were updated every June. These average estimates were used generate the expected returns. All averaged estimates at time period t were generated using estimates prior to time period t. All variables were used to generate expected returns.

Model	$\hat{\mu}_{ex}$	$\hat{\sigma}$	SR	$\hat{\mu}_{TC}$	SR$_{TC}$	SR$_p$	Turnover
Panel A: Naive Diversification							
Naive	18.8862	19.0302	0.9924	14.4925	0.7616	1.0000	0.3486
Panel B: Minimum Volatility							
Voltiming	17.5354	17.0894	1.0261	12.8091	0.7495	0.3265	0.3792
Minv	13.1349	14.4346	0.9100	-2.7854	-0.1930	0.8573	1.3227
Panel C: Mean Variance Extensions							
TP	286.6487	1064.6237	0.2692	-27.0767	-0.0254	0.9996	26.1063
OC1N	12.0002	14.7844	0.8117	-6.0655	-0.4103	0.4029	1.4998
OC1Npos	16.1923	15.7265	1.0296	10.1826	0.6475	0.7038	0.4886
Panel D: Reward-to-Risk Timing							
RRT	18.3833	19.5181	0.9419	11.9076	0.6101	0.6155	0.5174
RRTpos	16.9636	17.7450	0.9560	11.7641	0.6630	0.4389	0.4113
Panel E: Reward-to-Risk Timing with Factor Models							
FF3m	18.3002	18.0238	1.0153	13.7224	0.7613	0.1920	0.3651
FF4m	18.2292	17.9899	1.0133	13.6069	0.7564	0.3043	0.3681
FF5m	18.5185	18.0369	1.0267	13.7324	0.7613	0.1429	0.3811
FF3mw	18.0990	17.9219	1.0099	13.4947	0.7530	0.4712	0.3681
FF4mw	18.0782	17.9362	1.0079	13.3522	0.7444	0.6520	0.3764
FF5mw	18.2133	17.9028	1.0173	13.4699	0.7524	0.2946	0.3780
Panel F: Parametric Portfolio Choice							
BSV Model	24.5786	22.7854	1.0787	13.9037	0.6102	0.6053	0.8568

Table 15: Annualized traditional performance statistics for dataset 4

Dataset 4: Full Sample, Top 10 %, 20 Funds, η = 1. Annualized traditional performance statistics for monthly out-of-sample returns for portfolios formed by stocks sorted by expected return. To generate the expected returns a model of the form $r_{i,t+1} = \alpha + \beta_1 ln(BE/ME)_{i,t} + \beta_2 ln(ME)_{i,t} + \beta_3 GPdat_{i,t} + \beta_4 R1to0_{i,t} + \beta_5 R12to2_{i,t} + \beta_6 INV_{i,t}$ was estimated and the estimates were averaged over an estimation window initialized with 60 months of data. Each subsequent time period the model was re-estimated by enlarging the sample and the average estimates were updated every June. These average estimates were used generate the expected returns. All averaged estimates at time period t were generated using estimates prior to time period t. All variables were used to generate expected returns.

Model	$\hat{\mu}_{ex}$	$\hat{\sigma}$	SR	$\hat{\mu}_{TC}$	SR_{TC}	SR_p	Turnover
Panel A: Naive Diversification							
Naive	21.9969	20.6264	1.0664	16.9960	0.8240	1.0000	0.3962
Panel B: Minimum Volatility							
Voltiming	20.7859	18.6397	1.1151	15.5350	0.8334	0.1572	0.4175
Minv	17.3693	16.8396	1.0315	4.1278	0.2451	0.9771	1.0938
Panel C: Mean Variance Extensions							
TP	-159.9085	744.3043	-0.2148	-375.2759	-0.5042	1.0000	17.9497
OC1N	19.5019	17.5698	1.1100	5.3301	0.3034	0.9426	1.1725
OC1Npos	20.3220	17.3229	1.1731	14.2783	0.8242	0.2046	0.4868
Panel D: Reward-to-Risk Timing							
RRT	15.4710	54.2181	0.2853	6.8472	0.1263	0.0300	0.6974
RRTpos	21.0679	19.3471	1.0889	15.3923	0.7956	0.8368	0.4508
Panel E: Reward-to-Risk Timing with Factor Models							
FF3m	21.5068	19.7428	1.0893	16.3384	0.8276	0.3189	0.4093
FF4m	21.4853	19.7020	1.0905	16.2765	0.8261	0.3838	0.4122
FF5m	21.6498	19.6861	1.0998	16.2708	0.8265	0.3539	0.4254
FF3mw	21.2588	19.6860	1.0799	16.0788	0.8168	0.7240	0.4115
FF4mw	21.2123	19.6464	1.0797	15.9386	0.8113	0.8408	0.4189
FF5mw	21.3044	19.6159	1.0861	15.9850	0.8149	0.6556	0.4221
Panel F: Parametric Portfolio Choice							
BSV Model	26.1527	23.7954	1.0991	15.1250	0.6356	0.9788	0.8821

Table 16: Annualized traditional performance statistics for dataset 5

Dataset 5 No Micro Sample, Top 20 %, 20 Funds, η = 1. Annualized traditional performance statistics for monthly out-of-sample returns for portfolios formed by stocks sorted by expected return. To generate the expected returns a model of the form $r_{i,t+1} = \alpha + \beta_1 ln(BE/ME)_{i,t} + \beta_2 ln(ME)_{i,t} + \beta_3 GPdat_{i,t} + \beta_4 R1to0_{i,t} + \beta_5 R12to2_{i,t} + \beta_6 INV_{i,t}$ was estimated and the estimates were averaged over an estimation window initialized with 60 months of data. Each subsequent time period the model was re-estimated by enlarging the sample and the average estimates were updated every June. These average estimates were used generate the expected returns. All averaged estimates at time period t were generated using estimates prior to time period t. $R1to0_{i,t}$ and $R12to2_{i,t}$ were used only as controls and thus β_4 and β_5 were not used to generate expected returns.

Model	$\hat{\mu}_{ex}$	$\hat{\sigma}$	SR	$\hat{\mu}_{TC}$	SR$_{TC}$	SR$_p$	Turnover
Panel A: Naive Diversification							
Naive	13.4765	18.7203	0.7199	12.5779	0.6719	1.0000	0.0549
Panel B: Minimum Volatility							
Voltiming	12.5950	16.5117	0.7628	11.5888	0.7019	0.2064	0.0659
Minv	8.5502	14.4617	0.5912	-0.9352	-0.0647	0.6241	0.7852
Panel C: Mean Variance Extensions							
TP	-328.0044	1460.7560	-0.2245	-471.2107	-0.3226	1.0000	30.3710
OC1N	8.4968	14.6460	0.5801	-2.0281	-0.1385	0.5765	0.8734
OC1Npos	11.4487	15.4110	0.7429	7.8612	0.5101	0.8871	0.2857
Panel D: Reward-to-Risk Timing							
RRT	19.6106	34.3230	0.5714	12.8037	0.3730	0.5171	0.5626
RRTpos	12.1773	16.8102	0.7244	9.9244	0.5904	0.9927	0.1708
Panel E: Reward-to-Risk Timing with Factor Models							
FF3m	13.0687	17.6047	0.7423	11.9771	0.6803	0.1458	0.0715
FF4m	12.9304	17.4152	0.7425	11.7597	0.6753	0.3273	0.0778
FF5m	12.9751	18.0251	0.7198	11.4692	0.6363	1.0000	0.1066
FF3mw	12.9604	17.4463	0.7429	11.8496	0.6792	0.2837	0.0732
FF4mw	12.7085	17.2062	0.7386	11.3157	0.6577	0.6565	0.0952
FF5mw	12.9080	17.6860	0.7298	11.5599	0.6536	0.7722	0.0929
Panel F: Parametric Portfolio Choice							
BSV Model	14.3909	21.6964	0.6633	13.4845	0.6215	0.0832	0.0518

Table 17: Annualized traditional performance statistics for dataset 6

Dataset 6 No Micro Sample, Top 10 %, 20 Funds, η = 1. Annualized traditional performance statistics for monthly out-of-sample returns for portfolios formed by stocks sorted by expected return. To generate the expected returns a model of the form $r_{i,t+1} = \alpha + \beta_1 ln(BE/ME)_{i,t} + \beta_2 ln(ME)_{i,t} + \beta_3 GPdat_{i,t} + \beta_4 R1to0_{i,t} + \beta_5 R12to2_{i,t} + \beta_6 INV_{i,t}$ was estimated and the estimates were averaged over an estimation window initialized with 60 months of data. Each subsequent time period the model was re-estimated by enlarging the sample and the average estimates were updated every June. These average estimates were used generate the expected returns. All averaged estimates at time period t were generated using estimates prior to time period t. $R1to0_{i,t}$ and $R12to2_{i,t}$ were used only as controls and thus β_4 and β_5 were not used to generate expected returns.

Model	$\hat{\mu}_{ex}$	$\hat{\sigma}$	SR	$\hat{\mu}_{TC}$	SR$_{TC}$	SR$_p$	Turnover
Panel A: Naive Diversification							
Naive	13.5641	20.5566	0.6598	12.5822	0.6121	1.0000	0.0627
Panel B: Minimum Volatility							
Voltiming	12.9849	18.6540	0.6961	11.8888	0.6373	0.2872	0.0730
Minv	10.8687	17.3561	0.6262	4.2204	0.2432	0.9711	0.5437
Panel C: Mean Variance Extensions							
TP	-113.8624	679.7078	-0.1675	-348.0275	-0.5120	1.0000	19.9724
OC1N	10.8660	17.4142	0.6240	3.5294	0.2027	0.9694	0.6011
OC1Npos	12.9111	17.4018	0.7419	9.8233	0.5645	0.3521	0.2414
Panel D: Reward-to-Risk Timing							
RRT	-8.8590	146.5847	-0.0604	-13.4206	-0.0916	0.9503	2.0932
RRTpos	13.3075	18.3821	0.7239	10.7695	0.5859	0.4489	0.1969
Panel E: Reward-to-Risk Timing with Factor Models							
FF3m	13.3226	19.7941	0.6731	12.0476	0.6086	0.6035	0.0868
FF4m	13.4418	19.5993	0.6858	12.0681	0.6157	0.2798	0.0945
FF5m	13.1686	19.9292	0.6608	11.4508	0.5746	0.9985	0.1249
FF3mw	13.4312	19.7320	0.6807	12.1414	0.6153	0.3524	0.0880
FF4mw	13.5868	19.3913	0.7007	11.9799	0.6178	0.1585	0.1126
FF5mw	13.2438	19.7494	0.6706	11.6675	0.5908	0.8351	0.1123
Panel F: Parametric Portfolio Choice							
BSV Model	14.1503	23.2425	0.6088	12.9944	0.5591	0.2697	0.0637

Table 18: Annualized traditional performance statistics for dataset 7

Dataset 7 No Micro Sample, Top 20 %, 20 Funds, η = 1 Annualized traditional performance statistics for monthly out-of-sample returns for portfolios formed by stocks sorted by expected return. To generate the expected returns a model of the form $r_{i,t+1} = \alpha + \beta_1 ln(BE/ME)_{i,t} + \beta_2 ln(ME)_{i,t} + \beta_3 GPdat_{i,t} + \beta_4 R1to0_{i,t} + \beta_5 R12to2_{i,t} + \beta_6 INV_{i,t}$ was estimated and the estimates were averaged over an estimation window initialized with 60 months of data. Each subsequent time period the model was re-estimated by enlarging the sample and the average estimates were updated every June. These average estimates were used generate the expected returns. All averaged estimates at time period t were generated using estimates prior to time period t. All variables were used to generate expected returns.

Model	$\hat{\mu}_{ex}$	$\hat{\sigma}$	SR	$\hat{\mu}_{TC}$	SR_{TC}	SR_p	Turnover
Panel A: Naive Diversification							
Naive	15.7034	19.5278	0.8042	11.0425	0.5655	1.0000	0.3679
Panel B: Minimum Volatility							
Voltiming	14.9151	17.9232	0.8322	9.9835	0.5570	0.3824	0.3935
Minv	12.2548	16.0982	0.7613	-4.7126	-0.2927	0.9546	1.4112
Panel C: Mean Variance Extensions							
TP	68.6387	626.1987	0.1096	-88.2678	-0.1410	0.9981	13.0091
OC1N	11.6455	16.0521	0.7255	-6.3579	-0.3961	0.8349	1.4968
OC1Npos	13.8660	16.7771	0.8265	7.7294	0.4607	0.8796	0.4975
Panel D: Reward-to-Risk Timing							
RRT	5.6832	67.3230	0.0844	-4.6000	-0.0683	0.0051	0.8332
RRTpos	14.1615	18.3372	0.7723	8.8188	0.4809	0.5984	0.4205
Panel E: Reward-to-Risk Timing with Factor Models							
FF3m	15.6092	18.7333	0.8332	10.7364	0.5731	0.0402	0.3857
FF4m	15.3581	18.6022	0.8256	10.4314	0.5608	0.2860	0.3894
FF5m	16.2285	18.8495	0.8610	11.0829	0.5880	0.0044	0.4054
FF3mw	15.4022	18.6426	0.8262	10.5065	0.5636	0.2401	0.3885
FF4mw	15.0339	18.5845	0.8089	9.9870	0.5374	0.9625	0.3989
FF5mw	15.7247	18.6953	0.8411	10.6612	0.5703	0.0370	0.3992
Panel F: Parametric Portfolio Choice							
BSV Model	19.1192	23.4713	0.8146	9.0496	0.3856	0.8212	0.8030

Table 19: Annualized traditional performance statistics for dataset 8

Dataset 8 No Micro Sample, Top 10 %, 20 Funds, η = 1. Annualized traditional performance statistics for monthly out-of-sample returns for portfolios formed by stocks sorted by expected return. To generate the expected returns a model of the form $r_{i,t+1} = \alpha + \beta_1 ln(BE/ME)_{i,t} + \beta_2 ln(ME)_{i,t} + \beta_3 GPdat_{i,t} + \beta_4 R1to0_{i,t} + \beta_5 R12to2_{i,t} + \beta_6 INV_{i,t}$ was estimated and the estimates were averaged over an estimation window initialized with 60 months of data. Each subsequent time period the model was re-estimated by enlarging the sample and the average estimates were updated every June. These average estimates were used generate the expected returns. All averaged estimates at time period t were generated using estimates prior to time period t. All variables were used to generate expected returns.

Model	$\hat{\mu}_{ex}$	$\hat{\sigma}$	SR	$\hat{\mu}_{TC}$	SR_{TC}	SR_p	Turnover
Panel A: Naive Diversification							
Naive	17.5648	21.2427	0.8269	12.3638	0.5820	1.0000	0.4141
Panel B: Minimum Volatility							
Voltiming	16.5526	19.5311	0.8475	11.1330	0.5700	0.6154	0.4327
Minv	13.5650	18.3560	0.7390	0.0281	0.0015	0.8013	1.1209
Panel C: Mean Variance Extensions							
TP	-112.6459	501.8284	-0.2245	-267.6213	-0.5333	0.9941	12.8302
OC1N	13.3198	17.6518	0.7546	-0.6327	-0.0358	0.8577	1.1546
OC1Npos	15.3939	18.3192	0.8403	9.2890	0.5071	0.9535	0.4931
Panel D: Reward-to-Risk Timing							
RRT	18.9744	26.9599	0.7038	10.2776	0.3812	0.6157	0.7007
RRTpos	16.0934	19.8762	0.8097	10.3342	0.5199	0.9305	0.4558
Panel E: Reward-to-Risk Timing with Factor Models							
FF3m	17.4095	20.4954	0.8494	12.0190	0.5864	0.2925	0.4276
FF4m	17.3909	20.3703	0.8537	11.9508	0.5867	0.3069	0.4311
FF5m	18.0749	20.4502	0.8839	12.4670	0.6096	0.0399	0.4412
FF3mw	17.3531	20.4426	0.8489	11.9658	0.5853	0.4108	0.4292
FF4mw	17.2370	20.3618	0.8465	11.7558	0.5773	0.6908	0.4367
FF5mw	17.6701	20.3390	0.8688	12.1342	0.5966	0.1050	0.4371
Panel F: Parametric Portfolio Choice							
BSV Model	21.9537	25.7327	0.8531	11.4811	0.4462	0.9994	0.8401

Table 20: Cross-sectional dispersion statistics for all models

Dataset	Top X %	Mom Vars Included?	Mu	Std	SR
Panel A: Fama and French Datasets from Kirby (2012)					
Momentum	–	–	0.2672	0.9166	0.0541
Size/BTM	–	–	0.2348	0.7889	0.0506
Industry	–	–	0.1803	0.7458	0.0454
Panel B: Full Sample Datasets					
Dataset 1	20%	No	0.1829	1.3512	0.0311
Dataset 2	10%	No	0.2312	1.4238	0.0346
Dataset 3	20%	Yes	0.2146	1.2830	0.0346
Dataset 4	10%	Yes	0.2788	1.3587	0.0387
Panel C: No-Micro Sample Datasets					
Dataset 5	20%	No	0.1810	1.2651	0.0305
Dataset 6	10%	No	0.2463	1.3152	0.0343
Dataset 7	20%	Yes	0.2135	1.1810	0.0344
Dataset 8	20%	Yes	0.2862	1.2730	0.0390

Table 21: Performance statistics for BSV approach

BSV (2009) approach statistics for stock characteristic estimates and portfolio weights. This application uses a power utility function with relative risk aversion coefficient, $\gamma = 5$. The parameter estimates and their standard errors are the time series averages of these statistics.

Stat	Dataset 1	Dataset 2	Dataset 3	Dataset 4	Dataset 5	Dataset 6	Dataset 7	Dataset 8
Panel A: Stock Characteristic Estimates								
θ_{btm}	3.8114	2.4151	5.1988	2.9576	2.6459	1.6391	3.2057	1.6797
Std. err.	1.1796	0.8237	1.5485	1.0935	1.2760	0.9683	1.4950	1.1017
θ_{me}	-2.6841	-2.0246	-0.2562	-0.4378	-5.9759	-3.7308	-3.6322	-2.5844
Std. err.	1.0366	0.8092	1.2058	0.9160	1.3717	0.9835	1.4427	1.0381
θ_{GP}	2.3704	1.6962	0.8032	1.0337	1.6095	1.2072	0.4744	0.3393
Std. err.	1.3368	0.9285	1.5512	1.0768	1.3015	0.9924	1.4447	1.0766
θ_{Inv}	-2.0302	-1.1070	-1.4853	-1.1473	-0.9718	-0.5599	-0.9578	-0.9393
Std. err.	1.1728	0.8158	1.3848	0.9684	1.2647	0.9267	1.3773	0.9790
θ_{R1to0}	–	–	-8.1571	-4.7122	–	–	-5.8722	-3.3108
Std. err.	–	–	1.2854	0.8917	–	–	1.1971	0.8771
θ_{R12to2}	–	–	3.8648	2.1285	–	–	3.2138	1.9581
Std. err.	–	–	1.0669	0.7994	–	–	1.0148	0.7963
Panel B: Portfolio Weight Statistics								
max w_i	0.1259	0.1381	0.2134	0.2164	0.1711	0.1869	0.2426	0.2482
min w_i	-0.0678	-0.0762	-0.1270	-0.1350	-0.1679	-0.1690	-0.1453	-0.1583
Avg. w_i	0.0294	0.0360	0.0453	0.0532	0.0535	0.0671	0.0547	0.0673

CHAPTER 4: FACTOR LOADINGS AS PROXIES FOR EQUITY CHARACTERISTICS, ROUND THREE

Introduction

I provide an updated analysis on the factors vs. characteristics argument originally discussed by Daniel and Titman (1997) and Davis, Fama, and French (2000). Daniel and Titman (1997) concluded that characteristics, not factors, better explain the cross-section of expected returns due to behavioral biases. Davis et al. (2000) provide compelling evidence that the short sample analyzed by Daniel and Titman (1997) was insufficient to settle the debate, and after analyzing a larger sample period, concluded that their proposed factor structure explained the cross-section of expected returns at least as good as the characteristic-based model of Daniel and Titman (1997). I provide an alternative and more direct approach to compare factors and characteristics which focuses on exactly what we care about the most: portfolio performance. For my null hypothesis, I assume that portfolios formed from individual equities based on characteristics, and portfolios formed from individual equities based on factor loadings, are equivalent. My analysis shows that some characteristics are not well approximated by factor loadings for individual equities but that this disagreement is mitigated for portfolios of equities. Specifically, I find that portfolios formed by pre-formation book-to-market factor loadings closely matches the empirical return distribution of portfolios formed by the book-to-market characteristic. I also find that the top performing portfolios formed by the characteristics of market equity, gross-profitability, and investment, are well approximated by their factor loadings. Furthermore, I use the cross-sectional regression methodology described in Dickson (2015), to combine the signals from multiple sources, and predict expected returns. These results show that significant excess returns can be earned when using factor loadings as proxies for equity characteristics. The implications of this work are directly applicable to the selection of mispriced constituents of mutual funds or ETFs where data on the characteristics of individual holdings may be difficult to obtain. While I study only the most relevant stock characteristics, extensions of this study open up the whole universe of asset pricing anomalies, nearly 314 according to Harvey et al. (2013).

Daniel and Titman (1997) focus on the fundamental question of whether the return patterns of characteristic-sorted portfolios are consistent with an underlying factor structure. Their tests rely on finding firms with characteristics that do not match their risk loadings. For example, a firm that is actually a growth firm according to its book-to-market-equity but is currently "acting" like a value firm. This would be the case for a growth firm which had a conditional factor loading on a distress factor more similar to value firms than growth firms. Daniel and Titman (1997) propose three competing models: 1) A model consistent with Fama and French (1993, 1996), 2) A model with a stable factor structure and time-varying risk return premia, and 3) A model where firm characteristics not factor loadings determine expected returns. Their results support a characteristic based model while Davis et al. (2000) argue that these findings are sample specific. Expanding on this work I study the characteristic vs. factors argument using all of the stock characteristics recently included in the Fama and French (2015) five-factor model. I argue that the awkward sorting procedures used by both Daniel and Titman (1997) and Davis et al. (2000) result in the authors searching for "outliers." This statement is difficult to refute as the portfolios in these aforementioned studies are highly unbalanced, resulting in some portfolios containing only one stock. This is a lesser-known but real concern when forming intersection portfolios across multiple dimensions. I instead study portfolios with an equal number of securities and aggregate forecasting signals across multiple dimensions in a straightforward manner.

Decomposing Expected Returns

Cross-sectional regressions

My analysis focuses on the model implied by the sequential cross-sectional regressions pioneered Fama and Macbeth (1973) (*FM hereafter*). These tests are simple, effective, and have likely been used more than any other asset pricing test to date. Their original paper tested the equilibrium relationships implied by the CAPM, but more generally, FM regressions can be used to test if any factor or characteristic is "priced." Priced means simply that the candidate variable exhibits a statistically significant risk-premium using FM standard errors. The procedure first fits cross-sectional regressions using OLS on the following functional form:

$$r_{i,t} = \alpha + x_{i,t}^{'}\beta + \epsilon_{i,t}, \quad i = 1, 2, ..., N, \quad \forall t \tag{47}$$

Following these t cross-sectional regressions there exists a sequence of t estimated intercepts, $\hat{\alpha}$, and t estimated slope coefficients, $\hat{\beta}$. The second step of the procedure involves taking the time-series average of these estimates, using the usual standard errors for the average; these standard errors are known as FM standard errors. If any element is statistically significant, then that candidate explanatory is said to be "priced," which means that the there is evidence that the explanatory variable is related to expected stock returns. While the implications of this relationship may not be immediately clear, Kirby and Cordis (2015) describe in detail that the elements of $\hat{\beta}_t$ are linear combinations of the cross-section of returns at time t. This is a result of the usual formulas for OLS estimators given by $\hat{\beta}_t = \hat{\Sigma}^{-1}_{xx,t}\hat{\Sigma}^{-1}_{xr,t}$. This result is well-known and expresses the OLS $\hat{\beta}$ vector as simply a weighted average. If we have J explanatory variables then $\hat{\Sigma}^{-1}_{xx,t}$ is the JxJ sample covariance matrix of $x_{i,t}$ and $\hat{\Sigma}^{-1}_{xr,t}$ is the $Jx1$ vector of sample covariances of $x_{i,t}$ and ri, t. By factoring out each $r_{i,t}$ Kirby and Cordis (2015) express the $\hat{\beta}$ vector in terms of $z_{i,t}$:

$$\begin{aligned} \hat{\beta}_t &= \frac{1}{N}\sum_{i=1}^{N} z_{i,t} r_{i,t} \\ z_{i,t} &= \hat{\Sigma}^{-1}_{xx,t}\left(x_{i,t} - \frac{1}{N}\sum_{i=1}^{N} x_{i,t}\right) \end{aligned} \tag{48}$$

Taking expectations of the term in parentheses results in an average of zero, meaning that the weights of each element of the $\hat{\beta}_t$ vector sum to zero. This means that each element is a weighted sum of returns where some returns are associated with a firm that has an explanatory variable higher than the cross-sectional average of that candidate variable, and some returns are associated with a firm that has an explanatory variable lower than the cross-sectional average of that candidate variable.[19] More importantly the $\hat{\beta}_t$ vector can be thought of as a vector of returns on hedge portfolios, or more commonly known as self-financing portfolios. Proceeds from shorting firms where $x_{i,t} < \bar{x}_t$ are used to finance investments in firms where $x_{i,t} > \bar{x}_t$. This results in a zero net investment, hence the phrase self-financing. Most importantly, this discussion illustrates that the $\hat{\beta}_t$ coefficients from the common FM regressions are just hedge portfolios, much like the factors found in the asset

[19]See Kirby and Cordis (2015) for more details.

pricing models of (Fama & French, 1993, 2015), Carhart (1997), and Hou, Xue, and Zhang (2014).

Time-series regressions

Time-series regression tests of asset pricing models have been around since the earliest tests of the CAPM, perhaps most notably Black, Jensen, and Scholes (1972). More recently, Fama and French (1993) used the time series tests of Black et al. (1972) to evaluate the effectiveness of their two additional factors, market equity and book-to-market equity, to the market portfolio return of the CAPM. The aforementioned models of Carhart (1997), Hou et al. (2014), and (Fama & French, 2015) all offer extensions of this basic concept pioneered by Fama and French (1993). These models attempt to describe the cross-section of expected returns and are tested with time-series regressions. In these models, the factors, in addition to the market return, can be thought of as priced factors that are approximately orthogonal to the overall market return, and therefore would be consistent with the multi-factor models of R. C. Merton (1973) and Ross (1976). All of these models imply the following equilibrium relationship:

$$E[R_t] = R_f + \sum_{k=1}^{N} \beta_k E[R_k] \tag{49}$$

This equilibrium relationship is formulated in terms of an expectation and therefore the variables are expressed in terms of future, ex-ante values. In that sense it also true that the β_k that we care about is also the future the β_k. However these models are always tested with ex-post, i.e. observed historical data. Validating these tests is discussed by Elton, Gruber, Brown, and Goetzmann (2010). Using the basic CAPM we can begin with the market model:

$$\tilde{R}_{i,t} = \alpha_i + \beta_i \tilde{R}_{M,t} \tilde{\epsilon}_{i,t} \tag{50}$$

Taking the expected value of this equation yields the following:

$$E[R_i] = \alpha_i + \beta_i E[R_M]$$

Therefore the following condition holds (51)

$$E[R_i] - \alpha_i + \beta_i E[R_M] = 0$$

Now adding this value of zero to Equation 50 and rearranging yields:

$$\tilde{R_{i,t}} = E[R_i] + \beta_i \left[\tilde{R}_{M,t} - E[R_m]\right] + \tilde{\epsilon}_{i,t} \tag{52}$$

Finally, the equilibrium relationship implied by the CAPM is given by:

$$E[R_i] = R_f + \beta_i \left[E[R]_{M,t} - R_f\right] + \tilde{\epsilon}_{i,t} \tag{53}$$

Plugging this into Equation 52 and we arrive at the testable form of the CAPM:

$$\tilde{R_{i,t}} = R_f + \beta_i \left[\tilde{R}_{M,t} - R_f\right] + \tilde{\epsilon}_{i,t} \tag{54}$$

Which can be extended to the larger class of multi-factor models:

$$\tilde{R_{i,t}} = R_f + \sum_{j=1}^{K} \beta_{i,j} \left[\tilde{R}_{j,t} - R_f\right] + \tilde{\epsilon}_{i,t} \tag{55}$$

There are several implicit assumptions in testing multi-factor equilibrium models with time-series tests: 1) the multi-factor versions of the market model must hold every period, 2) the equilibrium relationship must hold every period, and 3) the $\hat{\beta}$ is stable over time. Therefore, using a time-series regression as an asset pricing test, or using a time-series regression to estimate factor loadings, implies these assumptions.

Relating cross-sectional and time-series regressions

Relating these two approaches is key to understanding why forming portfolios on factor loadings would be expected to produce similar performance as portfolios formed from characteristics. Consider taking the time series average of the t estimates of Equation 47:

$$E[r_i] = E[\alpha] + \sum_{j=1}^{K} E[\beta_j]E[x_j] \tag{56}$$

Now comparing this to the equilibrium relationship in Equation 49 and the similarities become obvious. Recall that one of the time-series assumptions is that the $\hat{\beta}$ estimates are stable over time, much like the expectations in Equation 56. Following the discussion from Section 2.1, recall the $\hat{\beta}$ estimates in Equation 56 are zero-cost hedge portfolios and the x_j variables are observed stock characteristics. Now referencing Equation 55, the aforementioned popular asset pricing models all define $\tilde{R}_{j,t}$ as a zero cost hedge portfolio. In equilibrium, this means that the $\hat{\beta}_j$ in time-series regression are analogous to the observed stock characteristics in cross-sectional regressions, and the $\hat{\beta}_j$ in cross-sectional regressions are analogous to the $\tilde{R}_{j,t}$ in time-series regressions. Therefore, forming portfolios on the basis of observed stock characteristics implies the same equilibrium relationship as forming portfolio portfolios on the basis of factor loadings, when the factors are constructed as zero-cost hedge portfolios of the same characteristics. Cochrane (2011) also notes that portfolios formed by sorts are the same thing as nonparametric cross-sectional regressions. Additionally univariate sorts and univariate cross-sectional regression will yield the same portfolios.

A final point to make regarding cross-sectional and time-series regressions are the time periods used to construct each of the inputs. For the cross-section regressions I match up the data consistent with the conventional Fama and French (1992) timing convention, which means at any point in time my returns are matched with accounting data that is a minimum of 6 months old and a maximum of 18 months old. Of course this ensures I make investment decisions only using information known at time t. Using a time-series regression to estimate a factor loading however computes each $\hat{\beta}_j$ as a weighted average of all the time periods in the sample. Daniel and Titman (1997) and Davis et al. (2000) use rolling estimates to compute pre-formation loadings as do I. For example, if I estimate the loadings using five years of historical data, then the factor loading is a weighted average over the last five years, but the characteristic is one observation at a specific point in time. The aforementioned studies did not consider this but to ensure I am fair with what information is used to construct my portfolio signals, I also compute rolling averages of the characteristics before sorting stocks in portfolios. That way both the factor loadings and the characteristics are represented as weighted averages over the same time periods.

Empirical Application

Data

My sample spans July 1963 to Dec 2013 with monthly holding period returns obtained from the Center for Research in Security Prices (CRSP) and accounting data from Compustat. The sample includes common equity securities (share codes 10 and 11) for all firms traded on the NYSE, NASDAQ or AMEX (exchange codes 1, 2, and 3). I consider the following stock characteristics consistent with Novy-Marx (2013) and Fama and French

(2015): size (log(ME)), book-to-market (log(BE/ME))[20], profitability (ratio of gross profits to assets), and investment (growth of total assets from previous fiscal year). These four explanatory variables have all been shown to have substantial risk premiums in the aforementioned literature. As in Novy-Marx (2013), to reduce the effect of outliers, I trim all independent variables at the 1% and 99% levels. Consistent with the prior literature, I use the standard Fama and French (1992) timing convention where I match monthly stock returns for July of year t to June or year $t+1$ with the Compustat variables for fiscal years ending in calendar $t-1$. This standard timing convention ensures that all accounting variables are known at the time of the investment decision. Additionally, these characteristics are updated annually and thus remain static from July of a given to June of the following year.

The log of market equity (ME) is defined as price per share times shares outstanding from CRSP. To get ME for the firm, I aggregate values of all equity for a given permno and date. This aggregate value is assigned to the permno with the largest ME. Slightly deviating the Novy-Marx (2013) definition I update using the June market equity to compute this variable rather than the previous December value, since this increases its explanatory power in the cross section. The log of book-to-market (BEME) is book equity scaled by market equity. Book equity is shareholder equity, plus deferred taxes, minus preferred stock, when available. The shareholder equity components follow the tiered definitions consistent with those used in Fama and French (1993) to construct the HML factor. Stockholder equity is defined in Compustat as (SEQ) if available, or else common equity plus the carrying value

[20]Taking logs makes the cross-sectional distribution of market equity and book-to-market more symmetric, reducing the impact of outliers

of preferred stock is available (CEQ + PSTX) if available, otherwise total assets minus total liabilities (AT - LT) is used. Deferred taxes is deferred taxes and investment tax credits (TXDITC) if available, or else deferred taxes and/or investment tax credit (TXDB and/or ITCB). Preferred stock is redemption value (PSTKR) if available, or else liquidating value (PSTKRL) if available, or else carrying value (PSTK). Profitability (PROF) is defined as gross profits and earnings before extraordinary items are Compustat data items GP and IB, respectively, scaled by total assets. Gross profits are also defined as total revenue (REVT) minus cost of goods sold (COGS). Investment (INV) for firms in year t is the growth of total assets for the fiscal year ending in year $t-1$ divided by total assets at the end of year $t-2$. This matches the definition used by (Fama & French, 2015). In their valuation equation, the investment variable is actually defined as the expected growth of book equity, not assets. However, as they state, sorts on asset growth result in larger spreads of average return and using growth in book equity produces similar results.

Generating the Factors and Factor Loadings

Generating the Factors

As my analysis relies on the performance of portfolios constructed from signals based on equity characteristics versus equity factor loadings, I must estimate pre-portfolio formation factor loadings. To construct the factors I follow the procedure described in Kirby and Cordis (2015). This procedure is an alternative to the sorting procedure popularized by Fama and French (1993), and fully accounts for the estimated cross-sectional correlations between the candidate characteristics. This ensures that the portfolios formed to capture patterns of a specific equity characteristic truly measures just the marginal return of that characteristic. Therefore sorts on decile portfolios show dispersion in the characteristic of interest, while all other characteristics remain approximately constant across the deciles. Another advantage to this procedure is that it is simple to implement with a large number of candidate explanatory variables. The Fama and French (1993) procedure of finding intersections of stocks, quickly becomes awkward as the number of explanatory variables increases and the portfolios are highly unbalanced. Daniel and Titman (1997) report that some of their test portfolios contain only one stock, producing results that are far too exposed to idiosyncratic shocks.

The Kirby and Cordis (2015) regression-based approach relies on sorts of stocks based on the residuals from cross-sectional regressions for each of the equity characteristic. Given k equity characteristics, the cross-sectional regression for equity characteristic x_j is as follows:

$$x_{i,j,t} = \alpha + \sum_{i=1, i \neq j}^{k} \beta_i x_{i,j,t} + e_{i,j,t} \tag{57}$$

This regression is repeated $\forall t$ and the regression residuals for each of the $j \in (1, ..., k)$ equity characteristics, $\hat{e}_{i,j,t}$ are saved for each firm. I then form decile portfolios of stocks based on sorts of the residuals. The factors used to estimate the factor loadings are then constructed as self-financing hedge portfolios of the top and bottom deciles. Therefore the factors themselves represent the difference in returns due to the marginal impact of each of the k candidate stock characteristics. These results of this procedure are found in Table 22 and discussed in more detail in Section 5.1.

Generating the Factor Loadings

To generate the pre-formation factor loadings, $b_{k,i}$,[21] I use rolling time-series regressions for each stock, $r_{i,t}$, using the factors,$f_{k,t}$, as constructed in the previous section. I also include an equally-weighted market factor in the regression to control for variation in overall systemic risk. The time-series regression takes the following form:

$$r_{i,t} = \alpha_i + b_{Mkt,i} f_{Mkt,t} + b_{ME,i,t} f_{ME,t} + b_{BEME,i,t} f_{BEME,t} + \\ b_{PROF,i,t} f_{PROF,t} + b_{INV,i,t} f_{INV,t} + \epsilon_{i,t} \tag{58}$$

To estimate these loadings my rolling window consists of the previous 60 months of returns. Daniel and Titman (1997) use only 36 months of data and Davis et al. (2000) require at least 36 months of the last 60 months. I use at least 60 months to reduce estimation error in the factor loadings. While not reported, I also experimented with 120 months and 36 months and achieved similar results.

Portfolio Construction

To conduct the empirical analysis, I focus on the performance statistics of returns to portfolios formed using signals derived from both the individual equity characteristics and the individual equity factor loadings. I use two main approaches that have similarities to those used by Daniel and Titman (1997) and Davis et al. (2000), but are also quite different. Daniel and Titman (1997) and Davis et al. (2000) both focus on finding risk loadings that do not match up with their associated equity characteristics. It seems very plausible that there are some cases where this is true, *e.g.* when a value stock acts more like a growth stock or a stock with high profitability acts more like a stock with low profitability. After all, the

[21] I use b here so as not to confuse the time-series factor loading from the cross-sectional estimates, β

factor loadings are conditional and just measured as a weighted average of "noisy" equity returns, where the weights are determined by the variance of "noisy" hedge portfolios. At the individual equity level, a firm may have attractive accounting measures but be part of a distressed industry which may cause patterns in returns that are not representative of other firms with similar characteristics. In that sense, the approaches used by Daniel and Titman (1997) and Davis et al. (2000) really are just finding firms in the tails of the distribution of returns sorted by a candidate equity characteristic, *i.e.* the outliers. This is also exactly why their triple-sorting procedure that results in 45 portfolios is very unbalanced, and in some cases producing portfolios containing only one stock – they are searching for outliers. My procedure is more robust to outliers as all of my portfolios analyzed contain the same number of stocks. The triple-sorting procedure of Daniel and Titman (1997) and Davis et al. (2000) is also completely intractable using four equity characteristics as I am analyzing. In fact, their triple-sorting procedure would result in 50,625 portfolios analyzing four equity characteristics, approximately 10 times more than the maximum number of stocks in any given year.

One-dimensional Sorts for Portfolio Construction

First I simply form portfolios based on observed stock characteristics and estimated pre-formation factor loadings. If time-series factor loadings really do proxy for equity characteristics, then portfolios formed by these estimated factor loadings should, on average, be the same as portfolios formed by the equity characteristics. I sort stocks into decile portfolios using both equity characteristics and pre-formation factor loadings to discern patterns across the empirical distribution of returns.

Multivariable Sorts for Portfolio Construction

My second approach relies on the sequential multivariable cross-sectional regressions described in Dickson (2015). This procedure aggregates forecasting signals from multiple sources and is far more tractable than the awkward univariate sorting procedure introduced by Fama and French (1993) for multiple characteristics. Cochrane (2011) stated in his AFA presidential address that following this work of Fama and French, the anomaly literature was once again "descending into chaos." Recently Harvey et al. (2013) documented over 314 different anomalous predictive variables. When discussing asset pricing as a function of characteristics, Cochrane (2011) also stated that "we will all end up running multivariate regressions" as we cannot "chop portfolios 27 ways (p. 1061)." This procedure was also used by Haugen and Nardin (1996), Hanna and Ready (2005), Fama and French (2006, 2008), and Lewellen (2014) to forecast expected stock returns and combine the signals from multiple stock characteristics. The approach not only exploits the univariate impacts of the candidate explanatory variables on the cross-section of expected returns, but also the cross-sectional covariance matrix of the explanatory variables. Just as as with my first approach, I sort stocks into decile portfolios using signals based on both equity characteristics and pre-formation factor loadings. In my first approach, the signals I used were the characteristics and loadings themselves, while in my second approach the signals I use are the expected returns from my sequential cross-sectional regressions. To generate the expected returns I begin with the following specification for the month by month cross-sectional regressions:

$$r_{i,t+1} = \alpha + \beta_{k1} x_{k1,i,t} + \beta_{k2} x_{k2,i,t} + \beta_{k3} x_{k3,i,t} + \beta_{k4} x_{k4,i,t} + \epsilon_{i,t} \quad (59)$$

For the portfolios formed by the equity characteristics, each $x_{k,i,t}$ is equal to each stock's observed characteristic. For the portfolios formed by the factor loadings, each $x_{k,i,t}$ is equal to a pre-formation estimated factor loading. Since my data spans from July 1963 through December 2013, and I smooth each stock characteristic over the previous 60 months to ensure the loadings and characteristics are estimated from the same time periods, my first estimated factor loading occurs in July 1973. Therefore I estimate this model for 486 months from July 1973 through December 2013. Since the conditional coefficient estimates from month to month are quite noisy, I apply a simple rolling average to the coefficients to filter out the signal. The smoothed coefficient estimates are generated as:

$$\hat{\beta}_t = \frac{1}{t-1} \sum_{i=1}^{t-1} \beta_i$$

Note that the mean parameter estimates for month t are based on estimates before month t, which ensures that I have an implementable trading strategy and incorporate no look-ahead bias. To curb excessive turnover and transaction costs, I update the smoothed coefficients annually, specifically ever June, and also begin with a sixty month burn-in period. For both the characteristic and factor based models I use this sixty month burn-in period, this brings my total out-of-sample count of month returns to 426 months. Denoting these smoothed estimates as $\hat{\beta}_{k1}$, the estimated expected returns are computed as:

$$\hat{r}_{i,t+1} = \hat{\alpha} + \hat{\beta}_{k1} x_{k1,i,t} + \hat{\beta}_{k2} x_{k2,i,t} + \hat{\beta}_{k3} x_{k3,i,t} + \hat{\beta}_{k4} x_{k4,i,t} \qquad (60)$$

Controlling for Estimation Error in the Loadings

To examine the extent to which estimation error may play a role in the estimation of the factor loadings, I also form equally weighted test portfolios containing on average, approximately thirty equities. I then compare the same performance statistics using characteristics and factor loadings. For each of these portfolios, I compute an equally weighted average of the characteristics of the equities including in the portfolio, and use these average characteristics as the observed signal when forming portfolios. Basically these average characteristics serve as the observed portfolio characteristic. The portfolios of equities have a substantially smaller amount idiosyncratic volatility due to diversification, which results in more accurate estimates for the $\hat{\beta}$ vector. This can easily be seen by examining the analytic formula for the variance of the OLS $\hat{\beta}$ vector: $\sigma_{\hat{\beta}_j} = \sigma_j^2 \Sigma_{x,x}^{-1}$. The sample covariance matrix is the same for both equities and portfolios of equities but the idiosyncratic volatility, σ_j^2, is much smaller for the portfolios than for the individual equities. This will result in more accurate estimates of the factor loadings. Another advantage of this approach is that it provides a good proxy for how well factor loadings can be used to form portfolios of ETFs, where aggregating individual equity data may be difficult or not even possible. These test portfolios serve as proxies for ETFs. According to my main hypothesis, as estimation error is reduced, the factor loadings should provide better proxies for the characteristics. Therefore I should observe portfolio returns that match up quite closely using either factor loadings or characteristics. I report results for these tests alongside the individual equities in Tables 27 through 35.

To form these portfolios I group stocks into 100 portfolios based on one-dimensional sorts of the following observable equity characteristics: size (log(ME)), book-to-market (log(BE/ME)), profitability (ratio of gross profits to assets), investment (growth of total assets from previous fiscal year), past performance measured at horizons of one month ($r_{1,0}$) to capture short-term reversals, and 12 to two months ($r_{12,2}$), to capture momentum.

This process results in 600 test portfolios, containing on average 30 equities. As previously mentioned, I also computed the average of the equity characteristics contained in each of these 600 portfolios to facilitate the same comparison as I did for individual equities.

Portfolio Turnover and Trading Costs

Portfolio turnover is an often overlooked but very real cost to investors. Transactional brokerage fee costs are typically not included in the calculation of a fund's operating expense ratio and thus the true operating expense of high turnover funds can be significant. As long as transaction costs are greater than zero, anything that increases turnover directly reduces the true performance of a fund. To examine the amount of trading required to implement each strategy I follow Kirby and Ostdiek (2015). Turnover is simply the fraction of invested wealth traded each period needed to re-balance the portfolio to the desired weights. At any time t I calculate turnover as:

$$Turnover_t = \sum_{i=1}^{N} \frac{1}{2} |\hat{w}_{i,t+1} - \hat{w}_{i,t^+}| \tag{61}$$

This definition of turnover is consistent with what is used in the mutual fund industry, i.e. the lesser of the value of purchases or sales in the period divided by the net asset value. (Kirby & Ostdiek, 2015) Since there are no fund inflows or outflows these must be equal. I define $\hat{w}_{i,t}$ as the portfolio weight in asset i at time t; $\hat{w}_{i,t^+}$ is the portfolio weight before re-balancing at time $t+1$; and $\hat{w}_{i,t+1}$ is the desired portfolio weight at time $t+1$, after re-balancing. To compute $\hat{w}_{i,t^+}$ I must consider the mechanical changes that occur within the portfolio. Assets that have done well over the time period will make up more than their starting share of weight at the end of the period, and assets that have done poorly will make up less than there starting share. I compute $\hat{w}_{i,t^+}$ as:

$$\hat{w}_{i,t^+} = \frac{\hat{w}_{i,t}(1 + r_{i,t})}{1 + \sum_{i=1}^{N} \hat{w}_{i,t} r_{i,t}} \tag{62}$$

Starting from the beginning of my sample, the first weights occur in month 61, therefore the first turnover calculation occurs in month 62. Studies such as Kirby and Ostdiek (2015, 2012) and Demiguel et al. (2009) do not ignore these mechanical weights while others such as Brandt et al. (2009) do ignore these mechanical changes. I have found that ignoring these mechanical changes is innocuous in this setting but do include them to capture the most conservative view of the trading costs. Now the return of the portfolio net of the proportional transactions costs becomes:

$$r_{p,t+1} = \sum_{i=1}^{N} \hat{w}_{i,t} r_{i,t+1} - 2 \times c_{i,t} |\hat{w}_{i,t} - \hat{w}_{i,t-1^+}|, \tag{63}$$

where $c_{i,t}$ reflects the proportional transaction cost for stock i and time t. Since turnover is the value of assets both purchased and sold as a fraction of total wealth, and both purchases and sales incur transaction costs, I multiply the turnover in Eq. 63 by 2. To be as conservative as possible I set I set $c = 50$ basis points consistent with the measures used by Brandt et al. (2009), Demiguel et al. (2009), Kirby and Ostdiek (2012, 2015), and the even more recent estimates by Novy-Marx and Velikov (2014). Finally, letting L reference the burn-in-period of 60 months, and T represent the total number of months in my study, numerically the average turnover I report in the tables is:

$$Turnover = \frac{1}{T - L - 1} \sum_{t=L+1}^{T-1} \left(\frac{1}{2} \sum_{i=1}^{N} |\hat{w}_{i,t+1} - \hat{w}_{i,t^+}| \right) \tag{64}$$

Statistical Inference

To conduct statistical inferences about the relative performance of the various strategies using the Sharpe ratio, I follow Kirby and Ostdiek (2012) and use large sample t and $chi-squared$ statistics. I consistently compute these statistics using the generalized method of moments (GMM). For details of the proof of the general results see Hansen (1982). As Hansen (1982) shows, the Delta method, Slutky's theorem and LLN are all used to derive the asymptotic distribution of the GMM estimators. Recent asymptotic distribution derivations for Sharpe ratios are also provided by Opdyke (2007) and Bailey and de Prado (2011) who also use these theorems in their derivations. However I use GMM standard errors to appeal to these more recent derivations while still applicable in a more general context. In the analysis I begin with a set of moment conditions of the form $E(g(R_t, \theta)) = 0$, where $g(R_t, \theta)$ is a $J \times 1$ vector of moments, analogous to disturbances, R_t is a vector or returns, and θ is $J \times 1$ vector of parameters. I use the fundamental result from Hansen (1982) that, subject to general conditions, the limiting distribution of $\hat{\theta}$ is given by:

$$\sqrt{T}(\hat{\theta} - \theta) \xrightarrow{d} N(0, V) \tag{65}$$

I have the following:

$$V = D^{-1}SD^{-1'},\ D = E(\partial g(R_t, \theta)/\theta'),\ S = \sum_{-\infty}^{\infty} E(g(Y_t, \theta)g(Y_{t-j}, \theta)')$$

The moment conditions are specified as follows:

$$g(R_t, \theta) = \begin{pmatrix} R_{bench,t} - \sigma_{bench} \times SR_{bench} \\ R_{test,t} - \sigma_{test} \times SR_{test} \\ (R_{bench,t} - \sigma_{bench} \times SR_{bench})^2 - (\sigma_{bench})^2 \\ (R_{test,t} - \sigma_{test} \times SR_{test})^2 - (\sigma_{test})^2 \end{pmatrix} \tag{66}$$

Using Eq. 65 I have the asymptotic standard errors of the Sharpe ratios and can also easily conduct a Wald test of linear restrictions to determine if the differences between the Sharpe ratios of the benchmark and test portfolios are statistically different from zero. To do so I consider the following test statistic:

$$(\hat{SR}_{test} - \hat{SR}_{bench})(R_{SR} V R'_{SR})^{-1}(\hat{SR}_{test} - \hat{SR}_{bench}) \sim \mathcal{X}(1) \tag{67}$$

In Eq. 67 the discrepancy vector $R_{SR} = (-1, 1, 0, 0)$ and V is the asymptotic covariance matrix described in Eq. 65. It can be shown that the square root of this statistic is equivalent to the following limiting distribution:

$$\sqrt{T}((\hat{SR}_{test} - \hat{SR}_{bench}) - (SR_{test} - SR_{bench}) \xrightarrow{d} N(0, R_{SR} V R'_{SR}) \tag{68}$$

So in the case where the population Sharpe ratios are equal in the benchmark and test portfolios, I have the following large-sample test statistic:

$$\sqrt{T}\left(\frac{\hat{SR}_{test} - \hat{SR}_{bench}}{(R_{SR}\hat{V}R'_{SR})^{1/2}}\right) \overset{\text{asy}}{\sim} N(0,1) \tag{69}$$

I also use this asymptotic covariance matrix to conduct a simple $t-test$ on the individual significance of the Sharpe ratio. Of course I no longer need a discrepancy vector and can simply take the square root of the appropriate diagonal element of the matrix V to compute the test statistic.

Results

Overall, the results show that not all factors are well approximated by their factor loadings, but in some cases the results are commensurate, specifically for BEME. Additionally, estimated factor loadings for portfolios of equities appear to provide better proxies than for the individual equities. Specifically, for portfolios of equities, an investor would earn nearly 400 basis points a year over a benchmark portfolio using estimated factor loadings as proxies for characteristics, when sorting by expected return. For individual equities, an investor would only earn around 100 basis points a year over a benchmark portfolio, and this includes nearly a 65% increase in volatility. After considering transaction costs, this higher average return completely disappears.

Regression-Based Hedge Portfolios

Table 22 reports summary statistics of the four different sets of characteristic-based decile portfolios. Columns one through four report the first four centralized moments of returns for each decile, and columns five through nine report the averages of each of the characteristics. Panel A shows results for portfolio sorted by the BEME residual. The pattern in mean returns is monotonically decreasing in nearly every decile. For decile 1 the mean return is 1.3377% per month and for decile 10 the mean return is 1.0718% per month. This is consistent with value firms outperforming growth firms, i.e. that distressed firms earn higher expected returns. The top decile also displays the highest volatility, smallest negative skewness, and largest kurtosis. Most importantly, columns five through nine show that the procedure of Kirby and Cordis (2015) achieved the desired result. The column labeled BEME decreases from its largest value in decile 1 of 0.2081, to its smallest value in decile 10 of -1.3249, while all other characteristics remains approximately constant. This procedure created a spread in returns due to the variation of the characteristic of interest, BEME in this case, while holding all other characteristics approximately constant.

Panel B shows results for portfolios sorted by the ME residual. Mean returns range from 1.1186% per month to 1.3154% per month, consistent with small-cap stocks outperforming large-cap stocks over the sample. The volatility and kurtosis are also the smallest for decile 1 and display a monotonically decreasing pattern. As in Panel A, the average ME for each decile shows a strict monotonically decreasing pattern, while all other characteristics are approximately constant. The BEME values range from -0.3121 to -0.5194, but the other characteristics show a much tighter pattern.

Panel C shows results for portfolios grouped by the PROF residual. The mean returns range from 1.4688% in decile 1 to 1.0432% in decile 10. These patterns show that firms with large measures of gross-profitability outperform those with smaller measures. The other centralized moments follow similar patterns as in the previous panels. These portfolios formed by the PROF residual display the largest spread in mean returns of all four of the characteristic based hedge portfolios. The PROF values range from 0.7615 for decile 1 to 0.0935 for decile 10. BEME again varies by a small amount, ranging from -0.3164 to -0.6652, while all other characteristics are relatively constant.

Panel D shows results for portfolios grouped by the INV residual. The mean returns range from 1.1440% per month to 1.2796% per month and display the smallest spread in mean returns of the four residual sorted portfolios. These patterns are also consistent with the literature that investment is inversely related to expected stock returns (see Fama and French (2015)). The other three centralized moments follow similar patterns as before. Examining patterns in columns five through nine and INV varies from 0.3978 for decile 1 to 0.0126 for decile. As before, all other characteristics remain approximately constant, with the exception of BEME which again varies slightly, -0.2914 to -0.6991.

Table 24 shows the correlation matrix of the factors. Both BEME and ME have low correlations with the Mkt factor, 0.0606 and -0.0354 respectively; while both PROF and INV have high correlations with the Mkt factor, 0.3298 and 0.5076 respectively. For the cross-correlations of the factors, the largest correlation occurs between BEME and PROF, with an estimate of 0.2354, while the smallest correlation occurs between ME and PROF, an estimate of 0.0101. Overall the correlations are all quite small, further corroborating the effectiveness of the procedure described by Kirby and Cordis (2015).

Cross-Sectional Regression Estimates

Tables 25 and 26 show the results from sequential cross-sectional regressions with Fama and Macbeth (1973) standard errors. I also report average values of the regression R-squared. Several interpretations of cross-sectional regressions are worth noting. As described in Section 2, cross-sectional slopes can be interpreted as returns on characteristic-based portfolios. Second, Fama (1976) describes how the cross-sectional R^2 reflects how much ex post volatility is explained, and is not indicator of the predictive ability of the characteristics. Finally, if I can interpret the slopes as returns on characteristic-based portfolios, then I can directly interpret the size of the t-stat as the size of the Sharpe ratios of characteristic-based portfolios. Consider the definition of the t-statistic and the Sharpe ratio:

$$t = \frac{E(R)}{\sigma(R)/\sqrt{T}} \text{ and } SR = \frac{E(R)}{\sigma(R)}$$
$$SR = \frac{t}{\sqrt{T}} \tag{70}$$

As long as T is the same for all regressions, then a higher t-stat indicates a higher Sharpe ratio for the return on a characteristic-based portfolio. First I discuss the results for individual equities in table 25. While not wholly convincing, I do find support that characteristics explain the cross-section of expected returns. Panel A shows results using the average characteristics as explanatory variables and Panel B shows results using the pre-formation factor loadings. In Panel A, the risk-premiums are quite similar to those obtained by Kirby and Cordis (2015). In column 1 I show the results using all characteristics. The largest risk-premium is 0.7908 for PROF while the smallest risk-premium is -0.0431 for ME. Only BEME and PROF show significant t-statistics with values of 4.1036 and 4.0356 respectively. This means that portfolios formed by BEME and PROF have approximately the same Sharpe ratio, which confirms the results documented by Novy-Marx (2013), that the cross-sectional explanatory power of gross-profitability rivals that of BEME. In column 2 I drop BEME from the regression, which causes the risk premiums of both PROF and INV to fall, while the risk-premium for ME remains relatively unchanged. Additionally, both ME and INV now appear significant. In column 3 I drop ME and observe little change in the risk premiums of the other characteristics. In column 4 I drop PROF and observe the risk premium for BEME to fall from 0.2833 to 0.0944, while also becoming insignificant. Finally in column 4 I drop INV, and again observe litle change in the risk premiums of the other characteristics.

In Panel B, the factor loadings yield smaller risk premiums in all cases. In column 1 I show the results using all of the pre-formation factors, and only the BEME factor yields significant risk premium, with an estimate of 0.1763. Examining columns 2, 3, and 4, there is nothing significant to note. Dropping each of the explanatory variables has little impact on the risk premium estimates of the other factors. Comparing Panel B and Panel A, the R-squared estimates are larger in all cases for Panel B. This suggests that more ex post volatility is explained using the factor loadings than using the characteristics themselves.

Examining the results for portfolios of equities in Table 26, shows even more interesting results. Here I find much stronger evidence for the hypothesis that characteristics explain the cross-section of expected returns. Perhaps most obvious are the much higher R-squared values. Panel A of Table 26 shows an R-squared of 0.1429 compared to 0.0329 in Table 25. All other columns also yield higher R-squared values. This is likely due to the smaller amount of idiosyncratic volatility present in the portfolios or equities, compare to the individual equities themselves. In regards to the risk-premiums, all estimates are similar except for those of INV in Panel A. The risk premium estimate for INV in Panel A is -0.7441 in Table 26, compared to -0.5501 in Table 25. Additionally this estimate is statistically significant with a t-statistic of -6.7299. Column 2 follows a similar pattern as in Table 25, that is, removing BEME causes large changes in the estimates of the other characteristics. In column 3, removing ME causes more significant changes to BEME and PROF than it did in Table 25. Similar conclusions hold in columns 4 and 5 when dropping PROF and INV.

In Panel B, the factor loading display much stronger risk premium estimates than in Panel B of Table 25, supporting the result factor loadings are reasonable proxies for latent characteristics. In column 1, BEME, PROF, and INV all display statistically significant risk premiums. Perhaps most importantly, the sign of the effect matches that of prior literature.

BEME and PROF are both positive showing that value firms (distressed) and firms with high measures of profitability, both have positive risk-premiums. ME and INV are negative showing that large firms and firms with high levels of investment have negative expected returns. In columns 2 through 5, dropping each of the explanatory variables seems to have a smaller impact on the other estimates than was the case using individual equities of Table 25. In conclusion, the factor loadings from portfolio of equities seem to provide promise for achieving a reduction in estimation error. Since using factor loadings show patterns in risk premiums for portfolios of equities, I suspect to find similar patterns in both one-dimensional and multivariable sorts.

Average Characteristic vs. Factor-Based Portfolios

Tables 27 through 33 show results for decile portfolios formed by sorts of the average characteristics over the previous sixty months, and sorts of the pre-formation time-series factor loadings estimated over the previous sixty months. Panels A and B show characteristic sorted and factor sorted portfolios respectively, for individual equities. Panels C and D show characteristic sorted and factor sorted portfolios respectively, for portfolios of equities. The results in panels C and D should be less sensitive to estimation error, and thus my null hypothesis of equal predictability should be more strongly supported by these results. To measure performance I report the first four centralized moments of monthly returns, the average firm size, the information ratio relative to the S&P 500, and the portfolio turnover.

Individual Equities

Table 27 displays performance statistics for BEME. In Panel A the mean return ranges from 1.5633% in decile 1 to 1.0631% in decile 10. In Panel B the mean returns are very similar ranging from 1.5415% for decile 1 to 1.0334% for decile 10. This pattern in mean returns strongly supports the hypothesis that factor loadings are good proxies for equity characteristics. In Panel A the standard deviation follows an unusual pattern ranging from 4.7442 in decile 1 to 5.0218 in decile 10. This results in decile 1 achieve the largest Sharpe ratio of 0.2404 per months and decile 10 achieve the smallest Sharpe ratio of 0.1276. In Panel B the pattern is not quite as strong but the Sharpe ratio ratios range from 0.1987 in decile 1 to 0.1029 in decile 10. A similar pattern emerges for the information ratio, which shows declining relative performance relative to the S&P 500. The largest standard deviation in Panel B occurs in decile 10, which also has the smallest return. In both Panels A and B decile 10 exhibits a sharp decline in performance relative to decile 9, nearly 20 basis points per month lower in each case with over a 25% reduction in Sharpe ratio. Only decile 1 of Panel B exhibits a slightly positive skewness while all other estimates are similar and slightly negative. The kurtosis in both panels exhibit slight excess kurtosis with an unusually large value of 12.3386 in decile 1 of Panel B, nearly 50% larger than any other estimate. In both panels, the smallest stocks occur in decile 1 and the average size increases as the deciles increase. However, Panel A shows a larger spread in average size ranging from 711.8802 to 3268.8946 while panel B only ranges from 916.4951 to 1637.2201. The turnover is quite small in Panel A, with the largest values occurring in the middle deciles and a maximum value of 0.0634. Panel B exhibits much larger turnover with a maximum value of 0.3733. Panel B also has the largest turnover measures in the middle deciles. In conclusion, investing in the top deciles by either characteristics or factors, would result in very similar portfolios. In the other deciles there are only slight differences in average size and turnover and in general these results strongly support my main hypothesis for BEME.

Table 29 displays performance statistics for ME. In Panel A the mean return ranges from 1.0953% for the largest stocks to 1.4431% for the smallest stocks. This pattern is not as strong in Panel B but is still present, ranging from 1.2078% for large stocks to 1.4071% for small stocks. Examining the average size of firms in these portfolios shows that sorts on the factor loadings do not result in the same spread in ME. In Panel A the average size ranges from 7639.4687 to 26.5111, while in panel B the average size ranges from 2722.9248 to 326.5065. Nonetheless, the performance in mean return and Sharpe ratios are similar for the top performing deciles. The turnover in Panel A is again much smaller than for Panel B in the middle deciles, but the top performing deciles are similar, 0.0271 for Panel A and 0.0917 for Panel B.

At first glance, Tables 31 and 33 would appear to display the weakest evidence in individual equities that factors proxy well for characteristics, due to the smaller return spreads in Panel B. However, the returns spreads are also smaller in Panel A for these characteristics as well. In Panel A of Table 31 the mean return ranges from 1.4651% in decile 1 to 1.2216% in decile 10. In Panel B this range is smaller, 1.2458% to 1.4067%. However the Sharpe ratios in all deciles are actually quite similar. Furthermore, the average firm sizes match up quite closely and are relatively constant. This pattern also emerges in Table 33 with Panel A exhibiting a mean return spread of 1.1968 to 1.5478 and 1.2417 to 1.4089 in Panel B. Both tables show patterns in Sharpe ratios that are commensurate with Tables 27 and 29. Turnover is larger for both characteristic and factor sorted portfolios, but otherwise follow similar patters as the previous tables, and the information ratios follow similar performance patterns as the Sharpe ratios. Even though the return spreads are smaller in Panel A and Panel B, the mean returns are still larger than an equally-weighted (EW) benchmark portfolio. For both PROF and INV and investor would sill earn approximately 180 basis points in excess of an EW benchmark using characteristic sorts, and 120 points in excess of EW benchmark using factor sorts by investing in the top deciles.

Referencing panel A of Table 35, the null hypothesis is that the Sharpe ratios for hedge portfolios of the top and bottom deciles formed by sorts of characteristics and factor loadings, are the same. I use GMM and Newey-West standard errors with five lags for the residual covariance matrix to facilitate the statistical tests. For BEME and PROF the null hypothesis cannot be rejected at any standard significance levels. However for ME, the null can be rejected at the 5% significance level and for INV, the null can be rejected at the 10% significance level.

Portfolios of Equities

The general conclusion examining the portfolios of equities in panels C and D of Tables 27 through 33, supports my main hypothesis, that factors proxy well for characteristics. The reduction in estimation error yields more accurate factor estimates and creates a more consistent pattern of performance, particularly for PROF and INV, which yielded weak support of my hypothesis using individual equities. In Table 27 the patterns are just as robust as they were for individual equities. Return spreads are slightly larger ranging from 1.6667% to 1.0793% in panel C and 1.5596% to 1.0962% Panel D. The Sharpe ratios match closely as does the information ratio relative to the S&P 500. Even the average firm size matches quite well ranging from 75.2817 to 782.0039 in panel C to 87.7710 to 440.8969 in panel D.

In Table 29 the return spreads range from 1.2007% to 1.6193% in panel C to 1.2627%

to 1.5893% in panel D. These mean return patterns are much closer than for the individual equities. Recall that for individual equities, sorts for ME resulted in portfolios of much different average firm size. This is not the case for the portfolios of equities. Average firm size ranges from 1381.2942 to 31.8611 in Panel C and 1282.4224 to 51.3290 in Panel D. These results in panels C and D provide much stronger support for my hypothesis for ME than did individual equities.

Tables 31 and 33 show that panels C and D much more closely than they did for panels A and B. In Table 31 this is particularly true for the top decile. The top decile in panel C yields a mean return of 1.3907%, a Sharpe ratio of 0.1522 and an information ratio of 0.1805. In panel C the mean return is 1.4076%, the Sharpe ratio is 0.1437, and the information ratio is 0.1599. The average firms sizes are also comparable, 189.3396 and 166.9053 respectively and the turnover of the factor sorted portfolio is quite small, 0.0910. In Table 33 the top decile also matches quite closely. In panel C the top decile yields a mean return of 1.6847%, a Sharpe ratio of 0.2199, and an information ratio of 0.2608. In panel D the top decile yields a mean return of 1.4780%, a Sharpe ratio of 0.1868, and an information ratio 0.2259.

Referencing panel B of Table 35, the null hypothesis is that the Sharpe ratios for hedge portfolios of the top and bottom deciles formed by sorts of characteristics and factor loadings, are the same. I use GMM and Newey-West standard errors with five lags for the residual covariance matrix to facilitate the statistical tests. For BEME, ME, and PROF the null hypothesis cannot be rejected at any standard significance levels. However for INV, the null can be rejected at the 1% significance level.

Aggregate Signal-Based Portfolios

Aggregating the signals from the candidate explanatory variables produces unambiguous improved performance over an equal-weight benchmark of all stocks. For the individual equities, the results suggest that factor loadings are unable to serve as proxies for the characteristics. However, when controlling for estimation error, my hypothesis is more strongly supported. That is, the results for the portfolios of equities support that signals derived from the factor loadings generate returns nearly 400 basis points per year in excess of the benchmark.

Individuals Equities

Panel A of Table 36 shows the results for the top portfolios formed from individual equity characteristics as sorted by expected return. I report performance statistics of the top 10%, 4%, 2% and 1% of stocks. The mean return ranges from 1.5844% for the top 10% to 2.0898% for the top 1%. The mean return is monotonically increasing as the number of socks gets smaller, suggesting that the model is produces consistent ordinal ranks of expected returns. The equally-weighted benchmark yields a mean return of 1.3114%, over 550 basis points smaller than the top 1% of stocks. As expected the standard deviation also increases with the mean return, ranging from 4.8435 for the top 10% to 5.9659 for the top 1%. This yields Sharpe ratios that are about the same for each of the four portfolios with a maximum of 0.2642 for top 4% of stocks. The equally-weighted benchmark has slightly smaller volatility and produces a Sharpe-ratio of 0.2009. The turnover is the smallest for the top 10% of stocks, a value of 0.0279, and largest for the top 1% of stocks, a value of 0.0456. This is an intuitive result since the portfolios are smaller, a larger percentage of the portfolio would turnover as stocks move in and out. All Sharpe ratios are statistically different that zero at standard levels, with P-values of 0.0000. The maximum Sharpe ratio of the test portfolio, the top 4% of stocks with a value of 0.2642, is statistically different the benchmark at the 10% significance level with a P-value of 0.0893. All others are not statistically significantly different at standard levels. All portfolios yield stocks that are relatively small, the largest group of stocks, the top 10%, have an average size of 152.2470 M.

Panel B of Table 36 show much less impressive results, suggesting that factor loadings act as poor proxies for the equity characteristics. If the factor loadings were good proxies for the equity characteristics, we would observe similar performance statistics in both panels. The largest mean return occurs for the top 4% of stocks, with a value of 1.3884%. This is marginally higher than the benchmark with a value of 1.3114. The standard deviation for these portfolios formed from factor loadings are around 20% higher than in panel A, with the largest value occurring in the top 1% of stocks, 7.4257. This is likely due to the larger turnover estimates, which range from 0.0992 for top 10% to 0.1542 for the top 1%. All Sharpe ratios are statistically significantly different than zero at standard significance levels, but are also smaller than the benchmark in every case. The largest estimate is 0.1746 for the top 10% compared to 0.2009 for the benchmark. All portfolios yield much larger stocks than in panel A. The average size ranges from 203.7652 M to 520.3163 M.

Portfolios of Equities

Panel A of Table 37 shows the results for the portfolios of stocks formed from the equity characteristics, while panel B shows the results for portfolios of stocks formed from the factor loadings. Comparing the performance statistics from panels A and B provides evidence to how well factor loadings proxy for the characteristics. In contrast to Table 36, these results strongly support my hypothesis that factor loadings proxy well for equity characteristics. This is likely due to the decreased estimation risk present in these portfolios of equities. The mean return in panel A ranges from 1.5844% for the top 10% to 2.0898% for the top 1%. In panel B the mean return ranges from 1.4676% for the top 10% to 1.6147% for the top 1%. While the mean return for the top portfolio in panel A is 550 basis points higher per year than in panel B, the top portfolio from panel B still beats the benchmark portfolio by 387 basis points per year. While the factor loadings do not yield as large of a mean return, this shows that the factor loadings still provide substantial predictive power. Looking at the risk of the portfolios, the top portfolio in panel B is less risky, an estimate of 6.3994 compared to 7.1381 for panel A. The Sharpe ratios are therefore quite similar, 0.2352 for the top portfolio in panel A compared to 0.1882 for the top portfolio in panel B. For panel A, all Sharpe ratios are statistically different than the benchmark value of 0.1588 at the 10% significance level. In fact the top 10%, 2%, and 1% are significant at better than the 5% level. Both models select relatively smaller stocks, but just as in Table 36 the factor loading select larger stocks on average.

While the results in these tables provide evidence that characteristics yield stronger patterns in expected returns than factor loadings, the factor loadings are still able to provide reasonable proxies that yield significant performance improvements over a standard benchmark. Furthermore, this analysis documents the danger of estimation error for individual equities. Since portfolios of equities contain much less idiosyncratic volatility, factor estimates in these cases retain strong patterns in expected returns that allows for the generation of excess return without actually observing the stock characteristics.

Conclusions

Expanding on the earlier studies by Daniel and Titman (1997) and Davis et al. (2000), I provide a comprehensive comparison of factors vs. characteristics using the four equity characteristics found in the new Fama and French (2015) five-factor asset pricing model. I provide an alternative methodology that does not rely on the awkward double and triple sorts employed in these after mentioned studies, and extend this debate to include a larger set of characteristics and factors. My main hypothesis is that factors provide reasonable proxies for equity characteristics. I use the recent procedures described by Kirby and Cordis (2015) to create empirical asset pricing factors, which fully marginalizes the impact of the other stock characteristics in the cross-section. For individual equities, my hypothesis is most strongly supported for BEME. The portfolios formed by BEME factor loadings matches the empirical distribution of portfolios formed by the BEME characteristic very closely. However, for ME, PROF, and INV, only the top portfolios are well approximated. In all cases, the spread in average firm size was much smaller for the factor sorted portfolios. To control for estimation risk, I also compare these estimates for portfolios of equities, which contain much smaller amounts of idiosyncratic volatility. In general, portfolios formed from factor loadings of these portfolios matched the characteristic based portfolios much more closely. Finally, I used the sequential cross-sectional regression methodology described in Dickson (2015) to aggregate the signals from multiple explanatory variables and forecast expected returns. This analysis revealed that factor loadings are poor proxies for characteristics with individual equities, but strong proxies for characteristics with portfolios of equities. Using just pre-formation factor loadings yielded a portfolio that beats an EW benchmark by nearly 400 basis points per year. Additionally, this portfolio yielded Sharpe ratios over 20% larger than the benchmark. The application of these results is quite useful as mutual funds and ETFs often contain large numbers of stocks making the aggregation of fundamental equity data cumbersome, unwieldy, and often times impossible. Therefore factor loadings should allow for a reasonable proxy for equity characteristics when constructing portfolios of mutual funds and ETFs.

Table 22: Summary statistics

(a) The following table shows summary statistics for the portfolios formed by the following firm characteristics: logarithm of the ratio of book to market equity (BEME), logarithm of market equity, ratio of gross profits to assets (PROF), and growth in total assets (INV). To form the portfolios at each month t, I conduct a cross-sectional regression of each characteristic on all others and a constant. The residuals are saved and the portfolios are formed on sorts of those residuals.

Panel A: Firms grouped by BEME residual

	Sample Moments of Returns				Sample Means of Characteristics			
Decile	Mean	Std	Skewness	Kurtosis	BEME	ME	PROF	INV
1	1.3377	5.2648	-0.2417	7.1553	0.2081	5.4716	0.3693	0.1521
2	1.2893	5.0827	-0.3076	7.1452	0.0125	5.4397	0.3255	0.1251
3	1.3448	4.9641	-0.3909	6.8768	-0.1156	5.4437	0.3164	0.1197
4	1.3350	4.8412	-0.3693	6.7458	-0.2209	5.4196	0.3137	0.1192
5	1.2914	4.8346	-0.3472	6.0046	-0.3147	5.3876	0.3108	0.1208
6	1.2475	4.8495	-0.5468	6.6435	-0.4105	5.3687	0.3053	0.1237
7	1.2546	4.8005	-0.3959	6.8752	-0.5172	5.3271	0.3114	0.1275
8	1.1589	4.8668	-0.4246	6.5925	-0.6680	5.3820	0.3256	0.1307
9	1.1229	4.8926	-0.4865	6.6299	-0.8757	5.4691	0.3451	0.1373
10	1.0718	5.1686	-0.3911	6.1215	-1.3249	5.5393	0.3686	0.1482

Panel B: Firms grouped by ME residual

	Sample Moments of Returns				Sample Means of Characteristics			
Decile	Mean	Std	Skewness	Kurtosis	BEME	ME	PROF	INV
1	1.1186	4.6663	-0.3519	5.4561	-0.3238	7.9370	0.3072	0.1136
2	1.1327	5.0275	-0.3111	6.0806	-0.4544	7.1564	0.3198	0.1266
3	1.1730	4.9489	-0.4568	5.9908	-0.5194	6.6135	0.3410	0.1329
4	1.2380	5.1680	-0.3492	5.9045	-0.4996	6.0964	0.3467	0.1386
5	1.2502	5.2369	-0.3383	6.3649	-0.4618	5.6404	0.3412	0.1404
6	1.3850	5.2414	-0.3059	6.6909	-0.4262	5.1967	0.3378	0.1372
7	1.3146	5.2581	-0.4514	6.3191	-0.4030	4.7563	0.3384	0.1381
8	1.2616	5.1773	-0.0997	8.4714	-0.3789	4.2636	0.3314	0.1364
9	1.2655	5.0115	-0.1729	6.9711	-0.3121	3.6636	0.3128	0.1266
10	1.3154	5.0734	-0.1010	7.4270	-0.4432	2.9085	0.3154	0.1136

Panel C: Firms grouped by PROF residual

	Sample Moments of Returns				Sample Means of Characteristics			
Decile	Mean	Std	Skewness	Kurtosis	BEME	ME	PROF	INV
1	1.4688	5.2568	-0.3041	7.2647	-0.5378	5.2385	0.7615	0.1269
2	1.3794	5.3741	-0.4175	6.2090	-0.4685	5.3082	0.5362	0.1360

3	1.3563	5.2248	-0.4364	6.3851	-0.4535	5.4828	0.4436	0.1327
4	1.3190	5.4095	-0.3586	6.5454	-0.3861	5.5216	0.3721	0.1350
5	1.2415	5.3112	-0.3365	6.8405	-0.3652	5.6398	0.3178	0.1303
6	1.2040	5.2510	-0.2529	6.1855	-0.3425	5.5967	0.2717	0.1306
7	1.1395	4.8021	-0.2311	6.4007	-0.3306	5.5671	0.2210	0.1285
8	1.1122	4.5812	-0.3536	6.8281	-0.3164	5.4581	0.1603	0.1272
9	1.1882	4.5092	-0.5358	6.6165	-0.3563	5.3488	0.1116	0.1263
10	1.0432	4.4973	-0.5385	6.9451	-0.6652	5.0864	0.0935	0.1307

Panel D: Firms grouped by INV residual

	Sample Moments of Returns				Sample Means of Characteristics			
Decile	Mean	Std	Skewness	Kurtosis	BEME	ME	PROF	INV
1	1.1440	6.0454	-0.2644	5.8190	-0.5952	5.3237	0.3326	0.3978
2	1.1921	5.4491	-0.3192	6.6530	-0.4037	5.3782	0.3412	0.2116
3	1.2668	5.0487	-0.3254	6.3720	-0.3275	5.3888	0.3411	0.1580
4	1.2692	4.8176	-0.4550	6.7707	-0.3082	5.4658	0.3302	0.1286
5	1.2625	4.6659	-0.4694	6.5900	-0.2914	5.5083	0.3211	0.1072
6	1.2518	4.5673	-0.4376	6.3077	-0.3137	5.5175	0.3111	0.0923
7	1.2617	4.5058	-0.4401	6.4096	-0.3601	5.4990	0.3137	0.0790
8	1.2661	4.5765	-0.4286	6.4746	-0.4244	5.4784	0.3216	0.0664
9	1.2619	4.7082	-0.4732	7.2960	-0.4989	5.4421	0.3348	0.0492
10	1.2796	5.0336	-0.4638	7.3292	-0.6991	5.2464	0.3443	0.0126

Table 24: Correlation matrix of factors

	EW Market	BEME	ME	PROF	INV
EW Market	1.0000	–	–	–	– –
BEME	0.0606	1.0000	–	–	– –
ME	-0.0354	0.1830	1.0000	–	– –
PROF	0.3298	0.2354	0.0101	1.0000	– –
INV	0.5076	0.1481	0.0597	0.2255	1.0000 –

Table 25: Fama-Macbeth regressions for individual equities

Time-series averages of cross-sectional regression estimates with Fama-Macbeth standard errors for Individual Equities.

	(1)	(2)	(3)	(4)	(5)
Panel A: Characteristic Sorted - Individual Equities					
Constant	1.3921	1.6906	1.1235	1.7706	1.3732
	5.6230	6.5528	6.0716	6.3354	5.2974
BEME	0.2833	–	0.3540	0.0944	0.3095
–	4.1036	–	4.4094	1.2635	4.4976
ME	-0.0431	-0.0760	–	-0.0696	-0.0430
–	-1.3287	-2.2997	–	-2.0269	-1.3203
PROF	0.7908	0.4988	0.8915	–	0.8229
–	4.0356	2.5711	4.1366	–	4.2200
INV	-0.1787	-0.9125	-0.1947	-0.4695	–
–	-0.5501	-2.8059	-0.5925	-1.3974	–
R2	0.0329	0.0287	0.0227	0.0252	0.0295
Panel B: Factor Sorted - Individual Equities					
	(1)	(2)	(3)	(4)	(5)
Constant	1.3323	1.3367	1.3144	1.3311	1.3305
	6.2388	6.2713	6.2212	6.2446	6.2325
BEME	0.1763	–	0.1923	0.1665	0.1611
–	2.8073	–	3.2324	2.5999	2.5569
ME	-0.1090	-0.1579	–	-0.0981	-0.1152
–	-0.9977	-1.5325	–	-0.8956	-1.0642
PROF	0.0806	0.0471	0.0740	–	0.0729
–	1.0586	0.6191	0.9662	–	0.9632
INV	-0.0268	-0.0622	-0.0244	-0.0236	–
–	-0.5815	-1.3183	-0.5265	-0.5150	–
R2	0.0401	0.0324	0.0287	0.0281	0.0338

Table 26: Fama-Macbeth regressions for portfolios of equities

Time-series averages of cross-sectional regression estimates with Fama-Macbeth standard errors for Portfolios of Equities.

	(1)	(2)	(3)	(4)	(5)
Panel A: Characteristic Sorted - Portfolios of Equities					
Constant	1.5165	1.6462	1.1727	1.7419	1.2980
	3.6971	4.2608	5.4016	4.2799	3.0333
BEME	0.1925	–	0.3451	0.1039	0.3748
–	2.0757	–	3.7122	1.1630	3.7220
ME	-0.0663	-0.0914	–	-0.0797	-0.0445
–	-1.2039	-1.8140	–	-1.4515	-0.7836
PROF	0.5913	0.3540	0.8248	–	0.8033
–	4.2983	2.2881	4.6750	–	6.1548
INV	-0.7441	-0.9129	-0.6480	-0.8167	–
–	-6.7299	-6.0905	-4.6217	-7.5581	–
R2	0.1429	0.1161	0.0812	0.1320	0.1306
Panel B: Factor Sorted - Portfolios of Equities					
	(1)	(2)	(3)	(4)	(5)
Constant	1.3121	1.2770	1.3822	1.3206	1.3192
	5.7050	5.4769	5.6333	5.7113	5.7421
BEME	0.2986	–	0.4205	0.2529	0.3746
–	3.1353	–	3.6808	2.6261	3.9884
ME	-0.2846	-0.3111	–	-0.3312	-0.2280
–	-1.7145	-1.8932	–	-1.9191	-1.3649
PROF	0.2403	0.1150	0.3776	–	0.1981
–	2.6122	1.1173	3.3210	–	1.8243
INV	-0.2485	-0.2684	-0.1882	-0.2155	–
–	-3.2845	-3.5092	-2.1370	-2.5986	–
R2	0.1342	0.1102	0.0823	0.1236	0.1237

Table 27: Performance statistics for BEME portfolios

(a) Decile portfolios sorted in descending order by the logarithm of the ratio of book equity to market equity (BEME) and the factor loading from the following multivariable regression model. A rolling window of 120 months was used to estimate each factor loading in a multivariable regression model of the form: $r_{i,t} = \alpha + f_t\beta + \epsilon_{i,t}$. The data sample includes all stocks in CRSP spanning from July 1963 - December 2013, consisting of 606 months. To be included each month, a stock must have at least 60 months of historical data. The first out of sample return occurred in July 1973 yielding a total of 486 monthly out-of-sample returns.

Decile	Mean	Std	SR	Skewness	Kurtosis	AvgSize	IRSP	Turnover
Panel A: Characteristic Sorted - Individual Equities								
1	1.5633	4.7472	0.2404	-0.3702	8.3196	711.8802	0.2734	0.0262
2	1.4882	4.4408	0.2400	-0.0707	8.4135	993.6654	0.2685	0.0478
3	1.3113	4.4024	0.2020	-0.4813	6.7614	1114.9336	0.2229	0.0572
4	1.2569	4.6338	0.1801	-0.7028	6.8077	1190.2610	0.2129	0.0628
5	1.2998	4.7031	0.1866	-0.4827	6.8493	1212.0947	0.2267	0.0634
6	1.3329	4.9907	0.1825	-0.4983	6.9116	1342.5765	0.2477	0.0615
7	1.2946	5.0120	0.1741	-0.4039	6.6249	1534.2828	0.2293	0.0567
8	1.2529	5.1694	0.1607	-0.5155	6.2270	1866.7373	0.2295	0.0494
9	1.2488	5.1214	0.1614	-0.3933	6.5013	2258.3903	0.2457	0.0379
10	1.0631	5.0218	0.1276	-0.4025	6.0153	3268.8946	0.1845	0.0198
Panel B: Factor Sorted - Individual Equities								
1	1.5415	5.6331	0.1987	0.3298	12.3386	916.4951	0.2361	0.1094
2	1.4395	4.6572	0.2184	-0.2915	7.4309	1349.5449	0.2601	0.2486
3	1.3741	4.4029	0.2162	-0.2308	7.1360	1469.1259	0.2584	0.3196
4	1.2633	4.2551	0.1977	-0.3624	6.6393	1568.7085	0.2298	0.3579
5	1.3285	4.4080	0.2056	-0.4852	7.1784	1604.2256	0.2502	0.3733
6	1.3408	4.6243	0.1986	-0.4725	7.1588	1704.1149	0.2679	0.3666
7	1.2432	4.5770	0.1794	-0.5966	6.3549	1730.6181	0.2275	0.3440
8	1.3092	4.7476	0.1868	-0.7361	6.6918	1705.7212	0.2615	0.3009
9	1.2386	5.0221	0.1625	-0.6132	6.3666	1799.5398	0.2350	0.2253
10	1.0334	5.9384	0.1029	-0.4426	5.3175	1637.2201	0.1175	0.0956
Panel C: Characteristic Sorted - Portfolios of Equities								
1	1.6667	5.6975	0.2184	0.1357	8.0025	75.2817	0.2428	0.0086
2	1.4497	5.4311	0.1892	-0.2079	7.2325	131.0011	0.2226	0.0223
3	1.3980	5.7908	0.1685	-0.0998	6.6490	149.9504	0.1976	0.0440
4	1.3001	5.8045	0.1512	0.0630	7.8197	156.9308	0.1703	0.0671
5	1.3068	5.5542	0.1593	-0.4015	6.8430	178.2079	0.1887	0.0776
6	1.3328	5.6590	0.1609	-0.4246	6.7800	186.8867	0.1917	0.0724
7	1.3507	5.9537	0.1560	-0.3651	6.5094	194.8831	0.1864	0.0572
8	1.4441	6.1121	0.1672	-0.3369	6.3211	281.4516	0.2064	0.0395
9	1.3534	5.9252	0.1572	-0.5171	6.0269	537.7850	0.2037	0.0216
10	1.0793	6.6085	0.0994	-0.3789	5.3013	782.0039	0.0975	0.0077
Panel D: Factor Sorted - Portfolios of Equities								
1	1.5596	5.7602	0.1975	0.0900	8.4152	87.7710	0.2215	0.0823
2	1.4458	5.6765	0.1803	-0.1266	7.5150	131.2032	0.2117	0.2120
3	1.4278	5.5947	0.1797	-0.2609	7.2887	165.8821	0.2158	0.2854
4	1.3905	5.5935	0.1731	-0.3240	7.2421	230.3414	0.2106	0.3202
5	1.4061	5.6730	0.1734	-0.2858	6.9491	342.3541	0.2131	0.3329
6	1.3943	5.7564	0.1689	-0.3028	6.7924	325.8211	0.2100	0.3322
7	1.3390	5.8628	0.1564	-0.2441	6.5603	373.5633	0.1868	0.3102
8	1.3264	5.8307	0.1551	-0.4593	6.2075	337.5814	0.1899	0.2687

9	1.2959	6.0164	0.1452	-0.3698	6.0454	440.8969	0.1721	0.1988
10	1.0962	6.6549	0.1013	-0.2891	5.1517	238.9677	0.0978	0.0765

Table 29: Performance statistics for ME portfolios

(a) Decile portfolios sorted in descending order by the logarithm of the market equity (ME) and the factor loading from the following multivariable regression model. A rolling window of 120 months was used to estimate each factor loading in a multivariable regression model of the form: $r_{i,t} = \alpha + f_t\beta + \epsilon_{i,t}$. The data sample includes all stocks in CRSP spanning from July 1963 - December 2013, consisting of 606 months. To be included each month, a stock must have at least 60 months of historical data. The first out of sample return occurred in July 1973 yielding a total of 486 monthly out-of-sample returns.

Decile	Mean	Std	SR	Skewness	Kurtosis	AvgSize	IRSP	Turnover
				Panel A: Characteristic Sorted - Individual Equities				
1	1.0953	4.6006	0.1463	-0.3177	5.0418	7639.4687	0.2234	0.0122
2	1.1482	4.8452	0.1498	-0.3452	6.1001	3231.8557	0.2248	0.0251
3	1.2290	4.8382	0.1668	-0.4488	6.5520	1762.0267	0.2541	0.0315
4	1.2143	4.9445	0.1602	-0.3245	6.7867	1112.0688	0.2189	0.0358
5	1.2689	5.0572	0.1674	-0.2790	6.1162	713.1073	0.2218	0.0396
6	1.3631	5.0555	0.1861	-0.5351	6.5586	449.4277	0.2323	0.0406
7	1.3919	5.0829	0.1908	-0.5281	6.7764	278.8805	0.2237	0.0410
8	1.5086	5.2477	0.2070	-0.3960	6.9515	158.8705	0.2454	0.0409
9	1.4576	4.9885	0.2075	-0.2959	6.9817	74.2806	0.2206	0.0380
10	1.4431	4.5621	0.2238	-0.0701	9.8236	26.5111	0.2040	0.0271
				Panel B: Factor Sorted - Individual Equities				
1	1.2880	5.0091	0.1728	-0.2599	5.6613	2722.9248	0.2488	0.1010
2	1.2176	4.7011	0.1692	-0.3581	5.6906	2537.4707	0.2363	0.2301
3	1.2078	4.4786	0.1754	-0.3839	5.6778	2360.6091	0.2215	0.2834
4	1.2955	4.6196	0.1890	-0.3923	6.0440	1983.8092	0.2586	0.3097
5	1.3470	4.6025	0.2009	-0.4250	6.9697	1670.2631	0.2741	0.3218
6	1.3304	4.7100	0.1928	-0.5698	7.2221	1403.8431	0.2544	0.3184
7	1.2940	4.8232	0.1807	-0.4968	7.2647	1068.3442	0.2253	0.2992
8	1.3356	4.8547	0.1881	-0.4767	7.8082	793.8067	0.2291	0.2638
9	1.4074	5.0233	0.1961	-0.4164	7.5966	603.3075	0.2269	0.2008
10	1.3962	5.6613	0.1720	-0.2078	7.8654	326.5065	0.1894	0.0917
				Panel C: Characteristic Sorted - Portfolios of Equities				
1	1.2007	5.5268	0.1409	-0.5180	5.8295	1381.2942	0.2003	0.0120
2	1.3509	5.3259	0.1744	-0.6047	6.4306	238.8767	0.2312	0.0430
3	1.3177	5.2371	0.1710	-0.5278	6.8905	205.5706	0.2155	0.0748
4	1.3039	5.3467	0.1649	-0.5042	6.7058	187.2592	0.2040	0.0843

5	1.3318	5.5134	0.1650	-0.4552	6.6749	170.3511	0.2028	0.0726
6	1.3203	5.7772	0.1555	-0.3694	6.7542	150.9825	0.1850	0.0586
7	1.4062	6.0811	0.1618	-0.2781	6.3726	130.7147	0.1926	0.0481
8	1.3944	6.2437	0.1557	-0.1274	6.7551	105.0560	0.1798	0.0330
9	1.4363	6.7959	0.1492	0.2529	7.4662	72.4160	0.1598	0.0169
10	1.6193	6.9071	0.1733	0.4195	7.0910	31.8611	0.1761	0.0056
Panel D: Factor Sorted - Portfolios of Equities								
1	1.2627	5.1770	0.1624	-0.6058	6.4515	1282.4224	0.2433	0.0581
2	1.3283	5.2712	0.1719	-0.6706	6.8126	245.2157	0.2211	0.1671
3	1.3866	5.3683	0.1796	-0.5302	6.8231	194.8802	0.2304	0.2330
4	1.3558	5.4700	0.1707	-0.4866	6.7290	182.5971	0.2118	0.2624
5	1.3385	5.6377	0.1625	-0.3869	6.6726	172.0896	0.1978	0.2731
6	1.3296	5.7471	0.1579	-0.3322	6.3432	161.9408	0.1874	0.2625
7	1.3424	5.9360	0.1550	-0.3007	6.4694	153.0223	0.1829	0.2366
8	1.3736	6.1653	0.1543	-0.1136	6.8134	133.0497	0.1771	0.1846
9	1.3747	6.5025	0.1465	0.0457	6.9675	97.8354	0.1571	0.1198
10	1.5893	7.3634	0.1585	0.4588	6.8935	51.3290	0.1639	0.0448

Table 31: Performance statistics for PROF portfolios

(a) Decile portfolios sorted in descending order by the ratio of gross profits to assets (PROF) and the factor loading from the following multivariable regression model. A rolling window of 120 months was used to estimate each factor loading in a multivariable regression model of the form: $r_{i,t} = \alpha + f_t\beta + \epsilon_{i,t}$. The data sample includes all stocks in CRSP spanning from July 1963 - December 2013, consisting of 606 months. To be included each month, a stock must have at least 60 months of historical data. The first out of sample return occurred in July 1973 yielding a total of 486 monthly out-of-sample returns.

Decile	Mean	Std	SR	Skewness	Kurtosis	AvgSize	IRSP	Turnover
Panel A: Characteristic Sorted - Individual Equities								
1	1.4651	4.9352	0.2113	-0.3868	7.2320	1451.2884	0.2880	0.0161
2	1.4779	5.0864	0.2075	-0.5376	6.2714	1476.1367	0.3092	0.0269
3	1.4014	5.2487	0.1866	-0.4599	6.2479	1375.1675	0.2667	0.0326
4	1.4354	5.3747	0.1885	-0.4059	7.3376	1376.1085	0.2524	0.0369
5	1.3155	5.3949	0.1656	-0.5005	7.3771	1524.8974	0.2170	0.0379
6	1.2957	5.5057	0.1586	-0.4989	6.9019	1646.4268	0.1957	0.0373
7	1.2650	4.9569	0.1700	-0.2850	7.3099	2021.6607	0.2131	0.0344
8	1.1795	3.9766	0.1904	-0.3320	5.2573	1818.8463	0.1694	0.0318
9	1.0577	4.3207	0.1471	-0.5474	6.2056	1365.5243	0.1124	0.0293
10	1.2216	4.7891	0.1669	-0.2413	6.5429	1425.2248	0.1571	0.0191
Panel B: Factor Sorted - Individual Equities								
1	1.3218	6.0991	0.1475	-0.5213	6.1592	1458.9437	0.1939	0.0999

2	1.2800	5.4308	0.1579	-0.4980	6.6487	1662.5159	0.2185	0.2332
3	1.3402	5.1304	0.1789	-0.3843	6.4041	1692.8206	0.2512	0.3009
4	1.3596	4.8751	0.1923	-0.4904	6.4245	1679.2967	0.2757	0.3362
5	1.3060	4.7026	0.1879	-0.4524	7.2981	1705.5378	0.2422	0.3479
6	1.4067	4.5311	0.2173	-0.4213	7.4265	1734.9664	0.2951	0.3402
7	1.3251	4.2034	0.2148	-0.6056	7.2813	1600.1667	0.2540	0.3178
8	1.2645	4.1359	0.2036	-0.3852	6.8854	1496.1358	0.2162	0.2809
9	1.2682	4.2834	0.1975	-0.2364	6.8791	1389.1088	0.2000	0.2232
10	1.2458	5.1176	0.1609	-0.5700	7.7181	1060.0129	0.1669	0.1001
				Panel C: Characteristic Sorted - Portfolios of Equities				
1	1.3907	6.3648	0.1522	-0.3294	5.8186	189.3396	0.1805	0.0086
2	1.3897	6.2085	0.1558	-0.1941	6.6292	259.5743	0.1810	0.0359
3	1.3288	6.3011	0.1439	-0.1772	6.7124	272.0949	0.1605	0.0691
4	1.3964	6.0755	0.1603	-0.1592	6.7637	268.5595	0.1899	0.0933
5	1.3772	5.7436	0.1663	-0.3262	6.5728	292.4726	0.2034	0.1138
6	1.3531	5.5190	0.1687	-0.3688	6.5254	312.3081	0.2072	0.1215
7	1.3707	5.3767	0.1764	-0.4212	6.8227	289.8493	0.2169	0.1025
8	1.3827	5.4334	0.1768	-0.4590	7.2475	341.6549	0.2178	0.0624
9	1.4860	5.7242	0.1858	-0.2191	6.9665	264.3746	0.2251	0.0269
10	1.2064	5.5313	0.1418	-0.1683	6.5138	184.1543	0.1544	0.0078
				Panel D: Factor Sorted - Portfolios of Equities				
1	1.4076	6.8592	0.1437	0.2326	7.8057	166.9053	0.1599	0.0910
2	1.3942	6.1886	0.1571	-0.1277	6.6955	235.8216	0.1835	0.2288
3	1.3864	6.0176	0.1602	-0.2582	6.5252	286.2536	0.1898	0.3033
4	1.3581	5.9166	0.1582	-0.2423	6.9696	370.2462	0.1902	0.3463
5	1.3584	5.7287	0.1634	-0.3312	6.6302	334.6335	0.1986	0.3698
6	1.3737	5.6790	0.1675	-0.3865	6.7272	294.8868	0.2063	0.3717
7	1.3482	5.5926	0.1656	-0.4024	6.8632	311.5909	0.2017	0.3534
8	1.3526	5.5162	0.1687	-0.4537	6.6643	291.1796	0.2034	0.3166
9	1.3522	5.4469	0.1707	-0.4219	6.4768	230.1525	0.2038	0.2376
10	1.3503	5.4087	0.1716	-0.3356	6.6237	152.7123	0.1941	0.0921

Table 33: Performance statistics for INV portfolios

(a) Decile portfolios sorted in descending order by the growth in total assets (INV) and the factor loading from the following multivariable regression model. A rolling window of 120 months was used to estimate each factor loading in a multivariable regression model of the form: $r_{i,t} = \alpha + f_t\beta + \epsilon_{i,t}$. The data sample includes all stocks in CRSP spanning from July 1963 - December 2013, consisting of 606 months. To be included each month, a stock must have at least 60 months of historical data. The first out of sample return occurred in July 1973 yielding a total of 486 monthly out-of-sample returns.

Decile	Mean	Std	SR	Skewness	Kurtosis	AvgSize	IRSP	Turnover
Panel A: Characteristic Sorted - Individual Equities								
1	1.1968	5.7132	0.1356	-0.3292	6.1187	1894.1748	0.1820	0.0370
2	1.1981	5.1805	0.1498	-0.4097	6.4542	1831.0808	0.2070	0.0760
3	1.2067	4.8514	0.1617	-0.4141	7.1375	1874.3697	0.2080	0.0994
4	1.2249	4.6771	0.1716	-0.6472	6.3774	1691.4812	0.2334	0.1159
5	1.3487	4.4736	0.2071	-0.5313	6.3955	1572.5631	0.2755	0.1261
6	1.3245	4.3987	0.2051	-0.2997	7.0100	1531.8470	0.2633	0.1263
7	1.2735	4.3405	0.1961	-0.5722	6.6042	1404.5103	0.2345	0.1206
8	1.3496	4.3544	0.2130	-0.4207	6.3165	1387.4106	0.2649	0.1081
9	1.4500	4.4521	0.2309	-0.6701	7.3976	1306.4763	0.2922	0.0820
10	1.5478	5.2199	0.2156	-0.3298	8.3856	983.2836	0.2719	0.0392
Panel B: Factor Sorted - Individual Equities								
1	1.2417	5.9046	0.1388	-0.4685	5.3733	1356.1277	0.1772	0.1109
2	1.2534	5.0398	0.1649	-0.6045	5.8711	1392.4742	0.2200	0.2608
3	1.3157	4.8435	0.1845	-0.4052	6.2160	1556.7439	0.2485	0.3380
4	1.2721	4.4618	0.1905	-0.4642	6.6395	1574.0473	0.2461	0.3788
5	1.3128	4.4233	0.2013	-0.1898	6.9725	1700.9240	0.2577	0.4050
6	1.3063	4.3813	0.2018	-0.4335	6.3827	1808.2396	0.2543	0.4034
7	1.3080	4.4636	0.1984	-0.3856	7.5235	1814.7782	0.2533	0.3920
8	1.3118	4.4975	0.1978	-0.7416	8.0595	1679.0401	0.2478	0.3534
9	1.3884	4.7837	0.2020	-0.4293	7.2102	1481.4823	0.2624	0.2754
10	1.4089	5.1753	0.1906	-0.2180	8.3139	1116.0338	0.2305	0.1209
Panel C: Characteristic Sorted - Portfolios of Equities								
1	1.0407	7.4335	0.0832	0.2028	7.3216	177.9602	0.0699	0.0129
2	1.3072	6.3713	0.1389	-0.1472	6.5024	245.6110	0.1566	0.0453
3	1.2580	6.0802	0.1375	-0.2453	6.6265	296.8456	0.1549	0.0758
4	1.3024	5.8445	0.1506	-0.3686	6.2362	278.5063	0.1787	0.0982
5	1.3591	5.6214	0.1667	-0.4500	6.6392	307.8645	0.2067	0.1159
6	1.3920	5.4128	0.1792	-0.4575	6.8690	275.9149	0.2242	0.1257
7	1.3647	5.2951	0.1780	-0.5528	6.7804	280.1670	0.2239	0.1136
8	1.4391	5.2700	0.1930	-0.5608	7.1011	343.1735	0.2469	0.0700
9	1.5337	5.3547	0.2076	-0.1919	7.3452	355.6432	0.2558	0.0266
10	1.6847	5.7419	0.2199	-0.0846	7.0263	112.6960	0.2608	0.0057
Panel D: Factor Sorted - Portfolios of Equities								

1	1.2501	7.0286	0.1178	0.3275	8.1306	124.2957	0.1196	0.0932
2	1.2764	6.0174	0.1419	-0.2677	6.3931	205.0109	0.1602	0.2376
3	1.3371	5.8308	0.1569	-0.3230	6.2839	247.8423	0.1864	0.3211
4	1.3651	5.7926	0.1628	-0.3617	6.5707	317.7527	0.1968	0.3719
5	1.3597	5.6977	0.1645	-0.3424	6.5704	326.4561	0.2008	0.3963
6	1.3754	5.6722	0.1680	-0.2762	6.9942	337.8435	0.2060	0.3971
7	1.4207	5.5923	0.1785	-0.4022	6.9402	343.3918	0.2213	0.3776
8	1.4121	5.4666	0.1811	-0.4075	7.0972	323.3564	0.2240	0.3393
9	1.4071	5.5371	0.1779	-0.3437	7.3097	258.7403	0.2170	0.2665
10	1.4780	5.6527	0.1868	-0.2480	7.2920	189.6925	0.2259	0.1127

Table 35: Performance statistics for hedged decile portfolios

Performance statistics for hedged decile portfolios sorted in descending order by their characteristics or factor. The factors were computed using a rolling window of 60 months in a multivariable regression model of the form: $r_{i,t} = \alpha + f_t\beta + \epsilon_{i,t}$. The data sample includes all stocks in CRSP spanning from July 1963 - December 2013, consisting of 606 months. To be included each month, a stock must have at least 60 months of historical data. The first out of sample return occurred in July 1973 yielding a total of 486 monthly out-of-sample returns. The 'P-' is the pvalue for the Sharpe ratios from a GMM estimation using Hansen-Hodrick standard errors testing whether the portfolios sorted by characteristics and factors are statistically different. I use the Newey-West estimator with five-lags for residual covariance matrix.

Variable	MeanDiff	Mean-P	SRDiff	SR-P
Panel A: Individual Equities				
BEME	-0.0130	0.9945	0.0198	0.8713
ME	-0.3672	0.0099	-0.0828	0.0277
Inv	-0.2834	0.1954	-0.1398	0.0681
Prof	0.2125	0.3507	0.0755	0.2164
Panel B: Portfolios of Equities				
BEME	0.1239	0.2984	0.0427	0.4376
ME	-0.0920	0.5595	-0.0305	0.4006
Inv	-0.4161	0.0009	-0.1412	0.0019
Prof	0.1269	0.5014	0.0798	0.2197

Table 36: Monthly traditional performance statistics individual stocks

Monthly traditional performance statistics for out of sample returns for portfolios formed by individual stocks sorted by expected return. To generate the expected returns for the characteristic based model (CHA), a model of the form $r_{i,t+1} = \alpha + \beta_1 ln(BE/ME)_{i,t} + \beta_2 ln(ME)_{i,t} + \beta_3 GPdat_{i,t} + \beta_6 INV_{i,t}$ was estimated and the estimates were averaged over an estimation window initialized with 60 months of data. All averaged estimates at time period t were generated using estimates prior to time period t. To generate the expected returns for the factor based model (FAC), a model of the form $r_{i,t+1} = \alpha + \beta_1 \lambda_{BEME,i,t} + \beta_2 \lambda_{ME,i,t} + \beta_3 \lambda_{GPdat,i,t} + \beta_6 \lambda_{INV,i,t}$ was estimated and the estimates were averaged over an estimation window initialized with 60 months of data. In the FAC models the λ estimates were generated from rolling regression factor loadings where the factors were constructed from 10 *minus* 1 decile spreads of the characteristics. The number of portfolios were either 100, 50, 25 or 10 as denoted in the Model column. Results are presented for the full sample. 'Turn' is the monthly turnover (actually 2× the turnover as calculated in the text), 'MeanTC' is the average return after transaction costs of 50 basis points, 'SRTC' is the Sharpe ratio after transactions costs of 50 basis points.

Models	Mean	Std	SR	Turn	MeanTC	SRTC	P-SR>0	P-SR$_{test}$ >SR$_{ben}$	AvgSize	AvgNum
Panel A: Top Portfolios formed from Equity Characteristics										
10 Ports	1.5228	4.8425	0.2297	0.0279	1.5089	0.2268	0.0000	0.4170	152.2470	95.2300
25 Ports	1.7202	4.9564	0.2642	0.0340	1.7033	0.2608	0.0000	0.0893	99.3934	38.0920
50 Ports	1.6568	5.3339	0.2336	0.0430	1.6353	0.2296	0.0000	0.6668	88.3556	19.0460
100 Ports	1.7731	5.9659	0.2284	0.0456	1.7503	0.2246	0.0000	0.7943	88.5130	9.5230
Panel B: Top Portfolios formed from Equity Estimated Factor Loadings										
10 Ports	1.3841	5.5770	0.1746	0.0992	1.3345	0.1657	0.0007	0.3722	520.3163	95.2300
25 Ports	1.4179	6.2243	0.1618	0.1193	1.3583	0.1523	0.0012	0.2601	334.3303	38.0920
50 Ports	1.3359	6.6123	0.1399	0.1387	1.2665	0.1294	0.0055	0.1254	234.9423	19.0460
100 Ports	1.3884	7.4257	0.1317	0.1542	1.3113	0.1213	0.0080	0.1533	203.7652	9.5230
Panel C: Benchmark Portfolio, All Stocks										
Benchmark	1.3114	4.4847	0.2009	0.0069	1.3080	0.2001	0.0002	1.0000	1706.7928	952.3005

All models are Equal Weighted.

Table 37: Monthly traditional performance statistics for portfolios of stocks

Monthly traditional performance statistics for out of sample returns for portfolios formed by portfolios of stocks sorted by expected return. To generate the expected returns for the characteristic based model (CHA), a model of the form $r_{i,t+1} = \alpha + \beta_1 ln(BE/ME)_{i,t} + \beta_2 ln(ME)_{i,t} + \beta_3 GPdat_{i,t} + \beta_6 INV_{i,t}$ was estimated and the estimates were averaged over an estimation window initialized with 60 months of data. All averaged estimates at time period t were generated using estimates prior to time period t. To generate the expected returns for the factor based model (FAC), a model of the form $r_{i,t+1} = \alpha + \beta_1 \lambda_{BEME,i,t} + \beta_2 \lambda_{ME,i,t} + \beta_3 \lambda_{GPdat,i,t} + \beta_6 \lambda_{INV,i,t}$ was estimated and the estimates were averaged over an estimation window initialized with 60 months of data. In the FAC models the λ estimates were generated from rolling regression factor loadings where the factors were constructed from 10 *minus* 1 decile spreads of the characteristics. The number of portfolios were either 100, 50, 25 or 10 as denoted in the Model column. Results are presented for the full sample. 'Turn' is the monthly turnover (actually 2× the turnover as calculated in the text), 'MeanTC' is the average return after transaction costs of 50 basis points, 'SRTC' is the Sharpe ratio after transactions costs of 50 basis points.

Models	Mean	Std	SR	Turn	MeanTC	SRTC	P-SR>0	P-SR$_{test}$ >SR$_{ben}$	AvgSize	AvgNum
Panel A: Top Portfolios formed from Equity Characteristics										
10 Ports	1.5844	6.1290	0.1915	0.0077	1.5805	0.1909	0.0002	0.0332	55.5234	60.0000
25 Ports	1.7872	6.8250	0.2017	0.0094	1.7825	0.2010	0.0000	0.0565	32.7393	24.0000
50 Ports	1.9359	7.1201	0.2142	0.0116	1.9301	0.2134	0.0000	0.0312	23.5585	12.0000
100 Ports	2.0898	7.1381	0.2352	0.0149	2.0824	0.2342	0.0000	0.0083	12.9506	6.0000
Panel B: Top Portfolios formed from Equity Estimated Factor Loadings										
10 Ports	1.4676	5.8982	0.1792	0.0815	1.4268	0.1723	0.0004	0.1147	73.1289	60.0000
25 Ports	1.5065	6.0176	0.1821	0.1021	1.4554	0.1736	0.0003	0.2340	47.0693	24.0000
50 Ports	1.5665	6.1886	0.1868	0.1235	1.5048	0.1768	0.0002	0.2552	34.3658	12.0000
100 Ports	1.6147	6.3994	0.1882	0.1353	1.5470	0.1776	0.0002	0.3321	25.5948	6.0000
Panel C: Benchmark Portfolio, Equal-Weight All Stocks										
Benchmark	1.2918	5.5499	0.1588	0.0000	1.2918	0.1588	0.0022	1.0000	294.6540	600.0000

All models are Equal Weighted.

CHAPTER 5: CONCLUSIONS

The chapters of my current body of work complement each other and present many potential extensions. To extend my first chapter, I am experimenting with more appropriate functional forms for the predictive regressions. By including interaction terms and non-linearities into the model, I hope to increase the forecasting power of the these regressions. These econometric issues are largely unexplored in the literature although Fama and French (2008) do tangle with the idea. Advances in the accuracy of return predictability regressions fits well into my studies on portfolio optimization and asset pricing models. As noted by Demiguel et al. (2009), using information about the cross-sectional characteristics of assets adds substantial value to their portfolio optimization experiments. Therefore as I improve on return predictability patterns I can extend my chapter 2 optimization analysis to include these new advances. Furthermore, as I observe and discover more accurate variables for identifying patterns in return predictability, I can augment my studies in chapter 3 and test these predictions in an asset pricing framework. To provide further insight on the use of signals from conditional asset pricing models, I have also compared sorts of stocks on the basis of *ex-ante* α and sorts of stocks on the basis of *ex-ante* expected return. I used estimation windows of 60 months, 36 months, and 24 months and found that decile portfolios formed from these two measures were statistically indistinguishable on the average and under several market regimes. Relating this to the momentum literature, “these two flavors of momentum taste the same,” and signals from factor models provide no value in forming tradeable and profitable portfolios. However, as my work stands, I think these separate sets of results provide the basis for another paper altogether. Therefore my current body of work provides ample avenues for future contributions to the field of empirical asset pricing.

REFERENCES

Andrea, F., & Pederson, L. (2014). Betting against beta. *Journal of Financial Economics*, *111*, 1–25.

Asness, C., Frazzini, A., & Pederson, L. (2012). Leverage aversion and risk parity. *Financial Analysts Journal*, *68*, 47–59.

Bailey, D., & de Prado, M. L. (2011). The sharpe ratio efficient frontier. *Unpublished Working paper*, –.

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

Black, F., Jensen, M., & Scholes, M. (1972). *Studies in the theory of capital markets*. New York: Praeger. (The Capital Asset Pricing Model: Some Empirical Tests)

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

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

Clarke, C. (2014a). The level, slope and curve factor model for stocks. *Unpublished Working paper*, –, –.

Clarke, C. (2014b). Return spreads in one-dimensional portfolio sorts across many anomalies. *Unpublished Working paper*, –, –.

Cochrane, J. (2011). Discount rates. *The Journal of Finance*, *66*, 1047–1108.

Daniel, K., & Titman, S. (1997). Evidence on the characteristics of cross sectional variation in stock returns. *The Journal of Finance*, *52*(1), 1–34.

Davis, J. L., Fama, E., & French, K. (2000). Characteristics, covariances, and average returns: 1929 to 1997. *The Journal of Finance*, *55*, 389–406.

Demiguel, V., Garlappi, L., & Uppal, R. (2009). Optimal versus naïve diversification: How inefficient is the $1/n$ portfolio strategy? *Review of Financial Studies*, *22*, 1915–1953.

DeMiguel, V., Martin-Utrera, A., & Nogales, F. J. (2013). Size matters: Optimal calibration of shrinkage estimators for portfolio selection. *Journal of Banking and Finance*, *37*, 3018–3034.

Dickson, M. (2015). Quantitative style investing. *Unpublished Working Paper*, –, 1–54.

Domowitz, I., Glen, J., & Madhavan, A. (2001). Liquidity, volatility, and equity trading costs across countries and over time. *International Finance*, *4*, 221–255.

Elton, J. E., Gruber, M. J., Brown, S. J., & Goetzmann, W. N. (2010). *Modern portfolio theory and investment analysis*. New Jersey: John Wiley and Sons.

Fama, E. (1976). *Foundations of finance*. New York, NY: Basic Books.

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

Fama, E., & French, K. (1993). Common risk factors in the returns on stocks and bonds. *The Journal of Finance*, *33*, 3–56.

Fama, E., & French, K. (1996). Multifactor explanations of asset pricing anomalies. *The Journal of Finance*, *51*, 55–84.

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

Fama, E., & French, K. (2008). Dissecting anomalies. *The Journal of Finance*, *63*, 1653–1678.

Fama, E., & French, K. (2014). Incremental variables and the investment opportunity set. *Unpublished Working paper*, –, –.

Fama, E., & French, K. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, *Forthcoming*, –.

Fama, E., & Macbeth, J. (1973). Risk, return, and equilibrium: Empirical tests. *The Journal of Finance*, *61*, 2163–2185.

Fleming, J., Kirby, C., & Ostdiek, B. (2001). The economic value of volatility timing. *The Journal of Finance*, *56*, 329-352.

Fleming, J., Kirby, C., & Ostdiek, B. (2003). The economic value of volatility timing using 'realized' volatility. *The Journal of Financial Economics*, *67*, 473–509.

Foster, D., & Nelson, D. (1996). Continuous record asymptotics for rolling sample variance estimators. *Econometrica*, *50*, 139–174.

Frahm, G., & Memmel, C. (2010). Dominating estimators for the global minimum variance portfolio. *Journal of Econometrics*, *159*, 289–302.

Garlappi, L., Uppal, R., & Wang, T. (2007). Portfolio selection with parameter and model uncertainty: A multi-prior approach. *Review of Financial Studies*, *20*, 41–81.

Green, R., & Hollifield, B. (1992). When will mean-variance efficient portfolios be well diversified? *The Journal of Finance*, *47*, 1785–1809.

Greene, W. H. (1997). *Econometric analysis*. New Jersey: Prentice Hall.

Han, Y., Zhou, G., & Zhu, Y. (2015). A trend factor: Any economic gains from using information over investment horizons? *Unpublished Working paper*, –, –.

Hanna, J. D., & Ready, M. J. (2005). Profitable predictability in the cross section of stock returns. *Journal of Financial Economics*, *78*, 463 - 505.

Hansen, L. (1982). Large sample properties of generalized method of moments estimators. *Econometrica*, *50*, 1029–1054.

Harvey, C. R., Liu, Y., & Zhu, H. (2013). ... and the cross-section of expected returns. *Unpublished working paper*, *Duke University*, –.

Hasbrouck, J. (2009). Trading costs and returns for us equities: Estimating effective costs from daily data. *Journal of Finance*, *64*, 1445-1477.

Haugen, R., & Nardin, L. (1996). Commonality in the determinants of expected stock returns. *Journal of Financial Economics*, *41*(3), 401 - 439.

Hou, K., Xue, C., & Zhang, L. (2012). Digesting anomalies: An investment approach. *Unpublished working paper*, *NBER Working Papers*, –.

Hou, K., Xue, C., & Zhang, L. (2014). A comparison of new factor models. *Unpublished Working Paper*, –, 1–91.

Jagannathan, R., & Ma, T. (2003). Risk reduction in large portfolios: Why imposing the wrong constraints helps. *The Journal of Finance*, *58*, 1651–1684.

Jegadessh, N. (1990). Evidence of predictable behavior of security returns. *The Journal of Finance*, *45*, 881–898.

Kan, R., & Z., G. (2007). Optimal portfolio choice with parameter uncertainty. *Journal Of Financial And Quantitative Analysis*, *42*, 621–656.

Kirby, C., & Cordis, A. (2015). Estimating the cost of equity capial using empirical asset pricing models. *Unpublished Working paper*, –, –.

Kirby, C., & Ostdiek, B. (2012). It's all in the timing: Simple active portfolio strategies that outperform naïve diversification. *Journal Of Financial And Quantitative Analysis*, *47*, 437–467.

Kirby, C., & Ostdiek, B. (2015). Optimizing the performance of sample mean-variance efficient portfolios. *Unpublished Working Paper*, –, –.

Ledoit, O., & Wolf, M. (2003). Improved estimation of the covariance matrix of stock retums with an application to portfolio selection. *Journal of Empirical Finance*, *10*, 603–621.

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

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

Merton, R. (1980). On estimating the expected return on the market: An exploratory investigation. *The Journal of Financial Economics*, *8*, 323–361.

Merton, R. C. (1973). An interremporal capital asset pricing model. *Econometrica*, *41*, 867–887.

Modigliani, F., & Merton, M. H. (1958). The cost of capital, corporation finance and the theory of investment. *American Economic Review*, *48*(3), 261.

Murtazashvili, I., & Vozlyublennaia, N. (2013). Diversication strategies: Do limited data constrain investors. *Journal of Financial Research*, *36*, 215–232.

Novy-Marx, R. (2013). The other side of value: The gross profitability premium. *The Journal of Financial Economics*, *108*, 1–28.

Novy-Marx, R., & Velikov, M. (2014). A taxonomy of anomalies and thier trading costs. *Unpublished Working paper*, –, –.

Opdyke, J. (2007). Comparing sharpe ratios: so where are the p-values? *Journal of Asset Management*, 308-336.

Pflug, G., Pichler, A., & Rendek, R. (2012). The 1=n investment strategy is optimal under high model ambiguity. *Journal of Banking and Finance*, *36*, 410–417.

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

Samuelson, P. (1967). General proof that diversification pays. *Journal of Financial and Quantiative Analsis*, *2*, 1–13.

Schwet, G. W. (2003). *Handbook of the economics of finance*. Amsterdam, Netherlands: Elsevier. (Anomalies and market efficiency)

Subrahmanyam, A. (2010). The cross-section of expected stock returns: what have we learnt from the past twenty-five years of reseach? *European Financial Management*, *16*, 27-42.

Tu, J., & Zhou, G. (2011). Markowitz meets talmud: A combination of sophisticated and naive diversification strategies. *The Journal of Financial Economics*, *99*, 204–205.