Charles University

Faculty of Social Sciences
Institute of Economic Studies

MASTER'S THESIS

Portfolio selection in factor investing

Author: **Bc. Martin Hronec**

Supervisor: **doc. PhDr. Jozef Baruník Ph.D.**

Academic Year: **2016/2017**

Declaration of Authorship

The author hereby declares that he compiled this thesis independently, using only the listed resources and literature, and the thesis has not been used to obtain a different or the same degree.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis document in whole or in part.

Prague, July 31, 2017

Signature

Acknowledgments

I would like to thank my supervisor PhDr. Jozef Baruník for his guidance, useful comments and future cooperation. I am very grateful to my girlfriend Dominika for the continuous support she offers. Last but not least, I would like to thank my good friend Katka for her willingness to help at any time.

Bibliographic note

Hronec Martin. *Portfolio selection in factor investing.* Prague 2017. pp. 79 Master's thesis Charles University, Faculty of Social Sciences, Institute of Economic Studies. Thesis supervisor doc. PhDr. Jozef Baruník Ph.D.

Abstract

This thesis empirically examines the role of advanced portfolio selection methods in factor investing. These methods provide more efficient exposure to underlying risk sources in factor portfolios. Their performance is evaluated across number of prominent factors and compared with more naive equal- and value- weighting, typically used in asset pricing literature as well commercial investment vehicles. The most diversified portfolio consistently achieves the highest returns, while having only moderate volatility and one of the lowest tail risk exposure. On the other hand, the diversified risk parity portfolio suffers high volatility as well as the greatest tail risk exposure, while achieving only comparable average returns with other strategies.

JEL Classification	G10, G11, G12
Keywords	factor investing, pricing anomaly,portfolio optimization, risk parity, risk budgeting, diversification

Author's e-mail	hroneq@gmail.com
Supervisor's e-mail	barunik@fsv.cuni.cz

Range of thesis: 132,790 characters with spaces

Abstrakt

Táto práca empiricky skúma úlohu pokročilých metód konštrukcií portfólia pri faktorovom investovaní. Tieto metódy umožňujú efektívnejšie zachytenie zdrojov rizika vo faktorových portfóliach. Ich výkonnosť je hodnotená naprieč mnohými faktormi a porovnáva sa s naivnejšími metódami, ktoré sa zvyčajne používajú v literatúre o oceňovacích anomáliach a faktorových indexoch. Najviac diverzifikované portfólio konzistentne dosahuje najvyššie výnosy, pričom má iba miernu volatilitu a jedno z najnižších rizík s nízkou pravdepodobnosťou výskytu. Na druhej strane, portfólio diferzifikovanej parity rizika trpí vysokou volatilitou a najväčšou expozíciou na riziká s nízkou pravdepdoobnosťou výskytu, pričom dosahuje len porovnateľne priemerné výnosy s inými stratégiami.

Klasifikace JEL	G10, G11, G12
Klíčová slova	faktorové investovanie, oceňovacie anomálie, optimalizácia portfolia, riziková parita,rozpočet rizika, diverzifikácia

E-mail autora	`hroneq@gmail.com`
E-mail vedoucího práce	`barunik@fsv.cuni.cz`

Master's Thesis Proposal

Institute of Economic Studies
Faculty of Social Sciences
Charles University in Prague

Author:	**Bc. Martin Hronec**	**Supervisor:**	**PhDr. Jozef Baruník Ph.D.**
E-mail:	hroneq@gmail.com	E-mail:	barunik@fsv.cuni.cz
Phone:	+420606681623	Phone:	+420(776)259273
Specialization:	FFM&B	Defense Planned:	June 2017

Proposed Topic:

Portfolio selection and risk management in equity value investing strategies.

Motivation:

Graham & Dodd (1934) were first to introduce a systematic accounting-based equity value investing strategy that significantly outperformed the market and spurred an intense debate about validity of this approach. The efficient markets side, the most prominently represented by Fama & French (1992), argues that the higher returns are only compensation for the higher inherent riskiness of such strategies. Also as with any finding in empirical research, the anomalous return relation could be the result of data snooping as presented by Lo & MacKinlay (1990) or more recently by Prado (2015). On the other hand, the behavioral finance side, the most prominently represented by Lakonishok, Shleifer & Vishny (1994), argues that systematic value strategies are not necessary more risky but they generate higher returns by being contrarian to the "naïve" strategies followed by other investors. The usefulness of relevant historical information for selection of firms with above average future prospects and elimination of those with poor prospects is empirically demonstrated by Piotroski (2001), followed by number of empirical studies, often based on long time ago proposed strategies, such as Greenblatt (2010), Carlisle, Mohanty & Oxman (2010), Carlisle & Gray (2013), Loughran and Wellman (2012), Novy-Marx (2013), Chingono & Rasmussen (2015) just to mention few.

Since most of these long or long-short strategies use different fundamental factors according to which they sort or restrict the tradable universe, the most profitable strategies are very often quite concentrated, not only with respect to the number of stocks in the portfolio, usually from 20 to 50, but mainly to other aspects of diversification such as exposure to risk factors. The portfolio weights are typically assigned almost exclusively as value weights, factor weights or equal weights. DeMiquel & Garlappi (2009) shows that equal weighted portfolio is probably the hardest to beat out-of-sample. This brings up an interesting question:
Can more advanced portfolio selection methods such as Hierarchical Risk Parity introduced by Prado (2016), risk parity based on the Effective Number of Bets proposed by Meucci, Santangelo & Deguest (2009) and risk management methods such as Dynamic Correlation or Tail Dependence Hedging suggested by Elkahmi & Stefanova (2010) or Principal Components as a Measure of Systematic Risk by Kritzman et al. (2010) bring long-promised out-of-sample supremacy over naïve 1/N portfolio strategy?

Hypotheses:

1. Hypothesis #1: Advanced portfolio selection and risk management methods improve returns of systematic accounting-based equity value investing strategies.
2. Hypothesis #2: Advanced portfolio selection and risk management methods improve risk-adjusted returns

of systematic accounting-based equity value investing strategies.
3. Hypothesis #3: Advanced portfolio selection and risk management methods decrease tail risk of systematic accounting-based equity value investing strategies.

Methodology:

The first-step consists of detailed literature review of advanced portfolio selection and risk management methods to identify other promising methods than those already mentioned in Motivation. In order to find out if methods we selected after literature review improve the performance of systematic accounting-based equity value investing strategies, I first need to replicate those strategies. I proceed to do so in Quantopian, an open-sourced hedge fund, which offers access to 14 years of point-in-time fundamental data from Morningstar database that can be directly used in backtesting environment in form of Python programming languate API. After implementation of original, naïve versions of investigated strategies I backtest them and collect the relevant information about results, e.g. excess return, Sharpe ratio, information ratio, maximum drawdown, etc. Subsequently I implement all chosen portfolio selection and risk management methods and incorporate them into systematic accounting-based equity value investing strategies,that also serve as performance benchmarks. Again I backest these upgraded versions and collect the relevant results. In order to test above mentioned hypotheses I compare the in and out-of-sample performance backtest results and present the results.

Expected Contribution:

To the best of my knowledge there is no comprehensive and systematic research about usefulness of advanced portfolio selection and risk management methods in accounting-based equity value investing strategies. Value equity portfolio managers tend to use simple naïve 1/N portfolio strategy because of its simplicity and strong out-of-sample performance. I will test if more advanced methods improve either returns, risk-adjusted returns or decrease tail risk specifically in accounting-based equity value investing strategies.

Outline:

1.) Introduction: motivation for and introduction to systematic accounting-based equity value investing strategy, advanced portfolio selection and risk management methods and reasons for their co-use.
2.) Literature review: divided in two main parts: systematic accounting-based equity value investing strategies and portfolio selection and risk management methods
3.) Methodology: theoretical background of applied portfolio selection and risk management methods
4.) Empirical results: results of backtests and comparison
5.) Discussion
6.) Conclusion

Core Bibliography:

Carlisle, T., Mohanty, S., & Oxman, J. (2010). Ben Graham's Net Nets: Seventy-Five Years Old and Outperforming.

DeMiguel, V., Garlappi, L., & Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy?. *Review of Financial Studies*, *22*(5), 1915-1953.

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

Graham, B., Dodd, D. L. F., & Cottle, S. (1934). *Security analysis* (p. 54). New York: McGraw-Hill.

Greenblatt, J. (2010). *The little book that still beats the market* (Vol. 29). John Wiley & Sons.

Gray, W. R., & Carlisle, T. E. (2012). *Quantitative Value,+ Web Site: A Practitioner's Guide to Automating Intelligent Investment and Eliminating Behavioral Errors* (Vol. 836). John Wiley & Sons.

Gray, W. R., & Vogel, J. (2012). Analyzing valuation measures: a performance horse race over the past 40 years.

Journal of Portfolio Management, *39*(1), 112.

Kritzman, M., Li, Y., Page, S., & Rigobon, R. (2010). Principal components as a measure of systemic risk.

Lakonishok, J., Shleifer, A., & Vishny, R. W. (1994). Contrarian investment, extrapolation, and risk. *The journal of finance*, *49*(5), 1541-1578.

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

Lopez de Prado, M. (2015). Backtesting. *Available at SSRN 2606462*.

Lopez de Prado, Marcos, Building Diversified Portfolios that Outperform Out-of-Sample (May 23, 2016). Journal of Portfolio Management, 2016, Forthcoming.

Loughran, T., & Wellman, J. W. (2012). New evidence on the relation between the enterprise multiple and average stock returns. *Journal of Financial and Quantitative Analysis*, *46*(06), 1629-1650.

Meucci, A., Deguest, R., & Santangelo, A. (2013). Measuring portfolio diversification based on optimized uncorrelated factors. *EDHEC-Risk Institute Publication (January)*.

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

Piotroski, J. D. (2000). Value investing: The use of historical financial statement information to separate winners from losers. *Journal of Accounting Research*, 1-41.

Rasmussen, D., & Chingono, B. K. (2015). Leveraged Small Value Equities. *Available at SSRN 2639647*..

Stefanova, D., & Elkamhi, R. (2011, March). Dynamic correlation or tail dependence hedging for portfolio selection. In *AFA 2012 Chicago Meetings Paper*.

Author **Supervisor**

Contents

List of Tables

Acronyms

ACCR Accruals

AG Asset growth

CAPM Capital asset pricing model

CEI Composite equity issues

CIa Cash-based operating profits to lagged assets

Cop Cash-based operating profitability

DIST Distress

dFin Changes in net financial assets

DR Diversification ratio

DRP Diversified risk parity

ENB Effective number of bets

Eprd Earnings predictability

ERC Equal risk contributions

EW Equal-weighting

FF3 Capital asset pricing model

FF3 Fama-French three factor model

FFC4 Fama-French-Carhart four factor model

GP Gross profitability

HRP Hierarchical risk parity

IA Investment-to-assets

Ivc Inventory changes

MISP Mispricing factor

MDP The most diversified portfolio

MOM Momentum

MV Minimum variance

NOA Net operating assets

Nop Net payout yield

NRP Naive risk parity

NSI Net stock issues

O O-score

Rdm R&D expenses to market

Rer Real estate ratio

VW Value-weighting

Chapter 1

Introduction

Factor investing, with its roots in arbitrage pricing theory, is best characterized as a way to capture systematic risk premia. It is based on the anomalies (factors) literature which serves as the scientific foundation for quantitative asset management. Number of factors were proposed to explain the so-called "cross-section" of returns, or the distribution of returns at a given point in time. The most famous examples are perhaps market, value, momentum and, quite recently, profitability. To display the power of factors, (Ang *et al.* 2009) show that risk factors represent 99.1% of the largest pension fund's[1] return variation.

Investor can then instead of allocating her portfolio between asset classes, such as bonds and equities, allocate among the risk factors in order to achieve desirable return or diversify herself.

The key question is which of the proposed factors actually matter. How is it possible that some relationship that has held for a long time in a variety of markets, and thus is very well known, continues to hold in the future. In essence, there are two reasons, risk or mispricing.

Argumentation for rational risk premium goes along the following lines. If some assets are riskier, and not just individually, which can be diversified away, but as a portfolio, then it's completely rational for them to be awarded a higher expected return. What does it mean to be risky? It means that investor has to lose sometimes and especially in the worst time possible, i.e. when almost everything crashes at once. If this would not be the case, than good and bad

[1] Government Pension Fund of Norway, holding 0.8% of global equity markets as of June 2017 according to Management (2017)

times could be diversified away, thus not deserving reward in terms of higher expected return.

On the other hand, mispricing simply means that investors make errors. Potential reasons for such errors are well documented in behavioral finance literature, e.g. Montier (2009). In this case, the assets have higher expected return not because they are riskier, but because investors make errors. In other words, assets are mispriced and they earn higher expected return when things return to rational prices, whether it means going up or down.

Risk and mispricing based explanations are not mutually exclusive, they can both be true and their relevance can vary through time. An illustrative example is the value factor. Some consider value strategy, general term for selecting portfolios with high exposure to the value factor, to be an aggressive strategy based on selecting distressed stocks, e.g. Fama (1998). Others consider that the purpose of value strategies in general is the selection of high quality stocks, e.g. Piotroski (2000). The value factor can thus be looked at not only from both of the risk and mispricing perspectives but also from their interaction.

Another concern is data-mining, i.e. extensively searching the data to find in-sample patterns in returns that are not real but random. Not surprisingly, random relationships doesn't tend to repeat and out-of-sample returns look anything but the ones promised by the in-sample discovery. Hou *et al.* (2017) conduct a largest replication in the field of pricing anomalies and obtain results suggesting the widespread p-hacking. On the similar note, McLean & Pontiff (2016) compare in-sample returns, post-sample returns and post-publication returns and find the latter returns to be significantly lower than in-sample returns, suggesting either data-mining or practitioners learning from the academic literature and arbitraging the premia away. In order to prevent the embarrassing number of false positive, Harvey *et al.* (2016) suggest multiple testing framework should be used when considering whether certain anomaly is statistically significant.

Above mentioned cases and other, in more details examined in Section 3.2, offer methods for distinguishing between real factors and false positives.

Answer to the question of why any, known or unknown, systematic factor offer a premia definitely helps in explaining which factors matter, however it does not directly provide the relevant factors. For a long time, the standard model asset pricing model used was the four factor model of Carhart (1997) based on the original Fama & French (1992) three factor model and momentum. Today, with hundreds of known factors, e.g. Harvey *et al.* (2016) documenting more than 300 in academic literature and Hou *et al.* (2017) implementing 447, investor faces "the factor zoo" (Cochrane (2011)).

With so many factors, the alpha puzzle of the cross-section is more critical than ever. If certain factor model is correct, this implies that portfolio returns can be explained by it, and thus that portfolio has no alpha.

However, if we consider that the correct model is some other factor model, then there is alpha in portfolio if we would use only the original model. With the proliferation of factors, the story is repeating itself and alpha always seems to come back.

Regardless of whether it is risk or mispricing that makes factors so powerful in explaining the cross-section of returns, holding the market alone is not optimal.

Even if one knows which factors are real, i.e. which are associated with higher expected returns going forward, in order to invest accordingly one must form the portfolio. However, one thing is to select the assets to be included into the portfolio, other is to actually assign them weights.

Typical weighting schemes used when constructing factor portfolios in the anomalies literature are value-weighting and equal-weighting. There is number of reasons why this is the case, while the most relevant is probably simplicity. Besides that, equal-weighting puts the same weight on every stock, thus matching the relevance of a given stock or observation in classical Fama & MacBeth (1973) cross-sectional regression. Value-weighting serves mainly as a remedy against the microcap stocks, with only limited economic relevance compared to the larger stocks, domination of the results simply because of their plentifulness.

My insight comes from the portfolio selection literature, according to which diversification pays off. Achieving diversification, given certain universe of assets, is however not as straightforward as putting equal weight on each asset.

Meucci (2010) proposes composing the portfolio's asset exposures by the portfolio's exposures to uncorrelated, in a sense statistical as opposed to fundamental, even though these are definitely tightly connected, sources of risk. The natural choice for such sources are the principal components, called principal portfolios. For a portfolio to be well diversified its overall risk should be evenly dispersed across these principal portfolios. Dispersion can be measured by Shannon entropy, while its exponential has intuitive meaning as the effective number of uncorrelated bets in the portfolio, i.e. the measure of diversification.

Earlier work and along the similar lines is the work of Choueifaty & Coignard (2008), who came up with the diversification ratio to be maximized. Choueifaty *et al.* (2013) show that square of the diversification ratio, in case of any portfolio, can be interpreted as the number of independent risk factors.

Portfolios based on equally-weighted risk contributions introduced by Qian and later developed further by (Maillard *et al.* 2010) are another attempt at diversification. They imitate the diversification effect of equally-weighted portfolios while taking into account single and joint risk contributions of the assets.

Risk contributions and sources of risk are not a mere mathematical decomposition of risk. They are quite good predictors of out-of-sample risks and therefore exhibit real and significant financial relevance.

Either because of risk or mispricing, factor portfolios are associated with higher expected returns. What happens when risk-based, diversification maximizing, portfolio selection methods are applied in the process of factor portfolios construction ?

From one perspective, if investors should not be compensated for diversifiable risk and advanced portfolio selection methods indeed provide more diversification than equal-weighting, then one could expect that the expected return should not change.

From another perspective, one could argue that advanced portfolio selection methods applied to restricted universe already exposed to some fundamental risk (or mispricing), provide more effective representation of factor portfolio and therefore should be able to capture alpha more effectively than more noise-contaminated methods such as equal-weighting.

Section 2 contains a short theoretical introduction into the Modern Portfolio

Theory, including its main shortcomings. Partial solution to such shortcomings is Arbitrage Pricing Theory, which is covered in Section 3 In Section 4 I review the state-of-the-art portfolio selection methods, which are later employed to answer the questions of interest. Methods used in the performance evaluation of individual strategies are introduced in Section 5. Finally, the main empirical results are presented in Section 6 alongside the disussion, while supplemental results are documented in B.2.

Chapter 2

Modern Portfolio Theory

2.1 Mean Variance Framework

Return maximization and risk minimization are conflicting objectives that need to be balanced according to the investor's preferences. The first rigorous asset allocation model to address this problem was developed by Markowitz (1952), known as an *mean-variance optimization*, and its framework serves as the foundation of modern portfolio theory. It allows the construction of optimal portfolio with respect to return maximization while constraining risk or risk minimization subject to minimum desired return.

Before outlining the process and main ideas of mean-variance analysis it is necessary to define relevant characteristics and statistics.

Discrete single-period return for the period $t-1$ until t can be calculated as

$$r_t = \frac{p_t - p_{t-1}}{p_{t-1}}$$

where p_t is the asset price at time t. When assuming continuous compounding of the capital is more convenient logarithmic returns are used

$$r_t = ln(\frac{p_t}{p_{t-1}}) = ln(p_t) - ln(p_{t-1})$$

Mean-variance optimization requires knowledge of expected (mean) returns, variance and covariance. The mean, variance and covariance are not observable ex-ante, thus they must be estimated typically based on historical returns

$$\hat{\mu_i} = \frac{1}{T}\sum_{t=1}^{T} r_{it}$$

$$\hat{\sigma}_i = \sqrt[2]{\frac{1}{T-1}\sum_{t=1}^{T}(r_{it} - \hat{\mu}_i)^2}$$

$$\hat{\sigma}_{ij} = \frac{1}{T-1}\sum_{t=1}^{T}(r_{it} - \hat{\mu}_i)(r_{jt} - \hat{\mu}_j)$$

$$\hat{\rho}_{ij} = \frac{\frac{1}{T-1}\sum_{t=1}^{T}(r_{it} - \hat{\mu}_i)(r_{jt} - \hat{\mu}_j)}{\hat{\sigma}_i\hat{\sigma}_j}$$

The portfolio weights are represented by $N \times 1$ vector $\boldsymbol{w}$ where w_i is the percentage holding of asset i and in case of no leverage condition, $\sum_{i=1}^{N} w_i = 1$ must hold. The expected portfolio return is simply the sum of the products of the asset returns and corresponding asset weights

$$\mu_p = \sum_{i=1}^{N} w_i \mu_i$$

in vector form

$$\mu_p = \boldsymbol{w'r}$$

The portfolio variance, σ_p^2 is calculated as

$$\sigma_p^2 = \sum_{i=1}^{N} w_i^2 \sigma_i^2 + \sum_{i=1}^{N}\sum_{\substack{i=1 \\ i \neq j}}^{N} w_i w_j \sigma_{ij}$$

in vector form

$$\sigma_p^2 = \boldsymbol{w'\Sigma w}$$

where $\boldsymbol{\Sigma}$ is $N \times N$ covariance matrix with the generic covariances σ_{ij} as elements.

From the formula for portfolio variance, the diversification effect, explicitly demonstrated in Markowitz (1959), can be observed. For clarity of illustration, assume a naive equal-weighted portfolio[1], where all n assets have the same

[1]Portfolio weighting method that beats a number of more advanced methods as I present later.

weight $1/n$. The portfolio variance can then be obtained as

$$\begin{aligned}\sigma_p^2 &= \sum_{i=1}^{N}\left(\frac{1}{n}\right)^2\sigma_i^2 + \sum_{i=1}^{N}\sum_{\substack{i=1\\ i\neq j}}^{N}\frac{1}{n}\frac{1}{n}\sigma_{ij}\\ &= \frac{1}{n}\left(\frac{1}{n}\sum_{i=1}^{N}\sigma_i^2\right) + \frac{n-1}{n}\left(\frac{1}{n(n-1)}\sum_{i=1}^{N}\sum_{\substack{i=1\\ i\neq j}}^{N}\sigma_{ij}\right)\end{aligned}$$

The mean of the variances and the mean of the covariances in the portfolio can be seen in the brackets. Distinguishing feature of the systematic (market) risk is that it cannot be diversified away through portfolio construction contrary to the asset-specific (unsystematic) risk. It can be clearly seen as the number of assets in the portfolio gets large, $n \to \infty$

$$\begin{aligned}&\frac{1}{n}\left(\frac{1}{n}\sum_{i=1}^{N}\sigma_i^2\right) \to 0,\\ &\frac{n-1}{n}\left(\frac{1}{n(n-1)}\sum_{i=1}^{N}\sum_{\substack{i=1\\ i\neq j}}^{N}\sigma_{ij}\right) \to \left(\frac{1}{n(n-1)}\sum_{i=1}^{N}\sum_{\substack{i=1\\ i\neq j}}^{N}\sigma_{ij}\right),\\ &\implies \sigma_p^2 \to \left(\frac{1}{n(n-1)}\sum_{i=1}^{N}\sum_{\substack{i=1\\ i\neq j}}^{N}\sigma_{ij}\right)\end{aligned}$$

This example shows why it is better to invest in a portfolio rather than hold individual assets.

Mean-variance optimization enables construction of mean-variance efficient portfolios by choosing the portfolio weights that offer the best return and risk pairs on the portfolio level. It is based on three main components: the quadratic objective function based on portfolio variance to be minimized; the set of the unknown variables representing the optimal portfolio weights to identify; and the set of constraints. Formally,

$$\min_{w}\left(\sum_{i=1}^{N}(w_i^2\sigma_i)^2+\sum_{i=1}^{N}\sum_{\substack{i=1\\ i\neq j}}^{N}w_iw_j\sigma_i\sigma_j\rho_{ij}\right)$$

$$\begin{aligned}&\text{s.t.}\\&\sum_{i=1}^{N}w_i\mu_i=\mu_p^*\\&\sum_{i=1}^{N}w_i=1\\&w_i\geq 0\end{aligned}$$

In a matrix form:

$$\begin{aligned}&\min_{w}\ \mathbf{w}'\boldsymbol{\Sigma}\mathbf{w}\\&\text{s.t.}\\&\mathbf{w}'\mu=\mu_{\mathbf{p}}^*\\&\mathbf{w}'\mathbf{e}=\mathbf{1}\\&w_i\geq 0\end{aligned}$$

where $\mathbf{e}$ is a $N\times 1$ vector of ones. This optimization problem is solved in an iterative manner, since restrictions on portfolio weights do not allow analytical solution. The set of all efficient portfolios are obtained by varying the target return r^* and they form the concave function called the mean-variance efficient frontier.

The choice of the efficient portfolios depends on the assumptions that investors care only about mean and standard deviation of an asset, and prefer a higher return over a lower return given the same level of risk, i.e. are risk averse. Two extremes are the minimum-variance portfolio, where the risk is minimized without considering the return and the maximum-return portfolio, where the return is maximized without considering the risk.

2.1.1 Drawbacks of Markowitz's portfolios

From theoretical perspective, mean-variance optimization is a clear, exact framework for facing the portfolio construction problem of investors. For its implementation, knowledge of expected returns, risks and correlations is necessary.

This is usually done using time series of historical returns and is called by Kan & Zhou (2007) the “plug-in” rule, since estimated parameters are simply treated as true parameters in the optimization process. Neglecting the parameters uncertainty renders the procedure to be totally deterministic. However, just because one does not see the estimation risk, it does not mean that it is not there. Estimation error is simply the difference between any parameter’s estimated value and that parameter’s true value. Ex-ante, it is very difficult to know how far estimates are from the true parameter. Using the mean-variance framework without acknowledging the threat of estimation error has severe consequences in a form of number of undesirable features as documented by number of authors, including Drobetz (2001), Herold and Maurer (2006), Jobson and Korkie (1981) and Michaud (1989).

These undesirable features are, in no particular order, poor out-of-sample performance, instability of weights and their extreme concentration, the lack of diversification, counter-intuitive nature, etc.

Regarding the above mentioned, the issue deserving attention is the lack of diversification of optimal portfolios which, paradoxically, contrasts with one of the main goals of Markowitz’s portfolio. Results of mean-variance optimization tend to have high weights on assets with high estimated returns compared to low standard deviations and negative correlations. Unfortunately, these assets are most likely to suffer from large estimation errors. A consequence of considering inputs into the optimization problem as very accurate is that, as Michaud (1989) wrote, “The unintuitive character of many optimized portfolios can be traced to the fact that MV optimizers are, in a fundamental sense, estimation-error maximizer”

The second major drawback is instability caused by high sensitivity of optimal portfolio weights to small changes in the estimated parameters’ instability. Best & Grauer (1991) suggested that this high sensitivity is especially pronounced in changes in expected returns.

The third issue is actual non-uniqueness of optimal portfolios. The uniqueness of the solution depends on the assumption that the inputs are without estimation error, which as discussed above is not the case.

The last but potentially the most serious limitation of mean variance efficient portfolios is the poor out-of-sample performance. Results of Klein & Bawa (1976) show that Markowitz’s portfolios suffer a significant deterioration of performance, given the expected performance and the same phenomenon holds in terms of risk-adjusted performance. Chopra (1993) obtain similar results and show that errors in mean are at least ten times as important as errors in variances, and errors in variances are about twice as important as errors in covariances.

Rather than discussing it now, advanced concept in this field are discussed later, in order to be presented together, in Section 4. Instead of them, this section is finished with an introduction of the capital asset pricing model as another pillar on which my thesis is based.

2.2 Capital Asset Pricing Model

The capital asset pricing model (CAPM) was introduced by Sharpe (1964), Lintner (1965b) and Mossin (1966). It is an equilibrium framework where investors make their portfolio decisions according to the mean-variance framework presented in Section 2.1. Its necessary assumption is a perfect market, thus number of conditions must hold. As mentioned, investors follow the mean variance framework. Unlimited short sales are allowed, unconstrained lending and borrowing at the risk-free rate is possible, no transaction costs and no taxes, infinitely divisible and marketable assets and importantly, homogeneous expectations regarding the mean, variance and correlations of assets. Consequentially, all investors hold the same risky portfolio, which in equilibrium is the market portfolio.

Tobin (1958) added the notion a risk-free asset into the model and pointed out that the efficient frontier is a straight line in its presence. Therefore there is single optimal portfolio of risky portfolio, dubbed the tangency portfolio. Further, under some assumptions, Sharpe (1964) proved that the tangency portfolio corresponds to the market-capitalization weighted[2] portfolio (or market portfolio). Relationship between the risk premium of asset i and the risk premium of the market portfolio is called the capital market line and mathematically it is represented according to the following formula:

$$E[R_i] - R_f = \beta_i^m (E[R_m] - R_f)$$

[2]Standardly referred to as value-weighted portfolio.

where R_i and R_m are the asset and market returns, R_f is the risk-free rate and the coefficient β_i^m is the beta of the asset i with respect to the market portfolio:

$$\beta_i^m = \frac{cov(R_i, R_m)}{\sigma^2(R_m)}$$

The capital asset pricing model highlights the role of beta, since the expected return of any asset or portfolio i depends only on the β. Contrary to idiosyncratic risks, β represents the asset's systematic risk that cannot be diversified away. Risk that cannot be eliminated, e.g. by holding large enough portfolio or more advanced and proper form of diversification, is compensated.

Even though testing the CAPM is difficult, because, as pointed out, it requires the knowledge of exact composition of the market portfolio, the model was extensively tested in number of empirical studies.

2.2.1 Empirical Evidence - Critique and Anomalies

The biggest challenge for the CAPM, as for any theory, is its confrontation with reality. CAPM has its fair share of problems. These problems have been known right from the start, when Lintner (1965a) did not find a statistically significant relationship between expected returns and market betas. Jensen *et al.* (1972) state that the bias of the regression coefficients arises through the so called errors-in-variables problem. Solution to this problem is offered in the same paper and subsequently in a famous work of Fama & MacBeth (1973), laying foundations of testing hypothesis in asset pricing literature. Method consists of sorting the assets into portfolios according to their estimated betas and from which the portfolio beta is calculated. Portfolio betas are a better measure, as the individual risks of the assets are smaller, thus there is lower residual variance in the model.

Even though Jensen *et al.* (1972) and Fama & MacBeth (1973) find supporting evidence for the relation of β to the excess return, it appeared to flat. Another concern was the instability of estimated coefficients in different subperiods. Further, assets with low (high) betas had a higher (lower) return than they should according to the CAPM, phenomenon known as low-beta anomaly. Some explanation for low-beta anomaly is offered by Frazzini & Pedersen (2014), who argue that in case of no leverage, high beta assets (i.e. riskier assets) should be over-weighted. As a consequence, there is a difference in high-beta and low-beta assets of correct or observed sign.

However, according to the CAPM, β is the only thing on which future returns depend. This starkly contrasts with many factors being found to be related to returns, and thus are not in line with the original theory. In general, factors that cannot be explained by some asset pricing model are called anomalies with respect to that model. Based on anomalies, new asset pricing models were proposed.

Chapter 3

Arbitrage Pricing Theory and Factor Models

In light of number of discovered anomalies, Ross (1976) came up with an alternative to CAPM, the arbitrage pricing theory (APT). The return of i-th asset is linearly dependent on number of factors:

$$R_i = \alpha_i + \sum_{j=1}^{m} \beta_i^j \mathcal{F}_j + \epsilon_i$$

where α_i is the intercept, β_i^j is the sensitivity of asset i to factor j, $\mathcal{F}_j$ is the (random) value of factor j and ϵ_i is the idiosyncratic risk of asset i. When the following conditions hold

$$\begin{aligned} &E[\epsilon_i] = 0 \\ &cov(\epsilon_i, \epsilon_k) = 0 \quad \text{for } i \neq j \\ &cov(\epsilon_i, \mathcal{F}_j) = 0 \end{aligned}$$

then the risk premium of asset i is a linear function of the risk premia of the factors:

$$\pi_i = E[R_i] - R_f = \sum_{j=1}^{m} \beta_i^j \pi(\mathcal{F}_j)$$

where $\pi(\mathcal{F}_j) = \mu(\mathcal{F}_j) - Rf$ and $\mu(\mathcal{F}_j) = E[\mathcal{F}_j]$.

A necessary assumption to make is that there is enough assets to build a portfolio which is sufficiently diversified with respect to the risk of individual assets, i.e. idiosyncratic risk. Another core assumption is a well-functioning security market, not allowing persistent arbitrage opportunities. Arbitrage means that risk-free profits can be made without a net investment by investors, simply exploiting security mispricing. If security prices allow for arbitrage opportunities, the market is not in equilibrium, hence there will be pressures on prices to adjust and eliminate these risk-free profits.

The basic idea of APT is that systematic risks are not captured by single market risk but a number of risk factors. In general, there are three types of risk factors with distinct characteristics:

- Statistical factors (e.g. determined by principal component analysis)
- Macroeconomic factors (e.g. surprises in inflation)
- Market factors (e.g. value or momentum)

In this thesis, I deal with two of the above mentioned factors, however, in a hierarchical or in a nested way.

It is important to mention the difference between factors and firm-specific characteristics, even though they are used interchangeably throughout this thesis, since there should be no confusion. Firm-specific characteristics are variables that can be computed using individual-firm data. Factors, on the other hand, are variables that proxy for a common source of risk, e.g., the market return. Firm-specific characteristics are related to factors because the return of a long-short portfolio based on a characteristic can be used as a proxy for an underlying unknown risk factor. The relation between characteristics and risk factors, however, is not always clear.

Standard long-short factor portfolio in asset pricing literature is obtained by eliminating the stocks around the median and selecting stocks in the bottom and top quantiles based on some firm-specific characteristics.

Let R_i be the rank of stock i according to the factor characteristic. Factor portfolios are value-weighted (VW) and equally-weighted (EW), i.e.

$$w_i^{VW} = \begin{cases} \frac{-ME_i}{\sum_{i=1}^{N} ME_i} & \text{if } R_i < q_{c_1}(R_i) \\ \frac{ME_i}{\sum_{i=1}^{N} ME_i} & \text{if } R_i > q_{c_2}(R_i) \end{cases}$$

and

$$w_i^{EW} = \begin{cases} -1/N & \text{if } R_i < q_{c_1}(R_i) \\ 1/N & \text{if } R_i > q_{c_2}(R_i) \end{cases}$$

where q_{c_1} and q_{c_2} are desired cut-off quantiles, usually 30% or 20% and N is the number of stocks in long and short leg of the portfolio together.

3.1 Classical Asset Pricing Models

There is an extensive literature regarding identification of risk factors. Some of the risk factors are long established and widely known, such as the size factor discovered by Banz (1981) or the value factor discovered by Rosenberg *et al.* (1985), others are more recent, such as the quality factor discovered by Novy-Marx (2013).

3.1.1 Fama-French Three Factor Model

One of the first asset pricing models based on multiple factors, and perhaps still the most famous one, is the Fama-French three factor model (FF3) of Fama & French (1993). It is based on their prior work focused on examination of individual factors, such as size, earnings-to-price, leverage and book-to-market equity.

$$E[R_i] - R_f = \beta_i^m(E[R_m] - R_f) + \beta_i^{smb}E[R_{smb}] + \beta_i^{hml}E[R_{hml}] + \epsilon_i^{FF3}$$

where R_{smb} is the return of small stocks minus the return of large stocks, and R_hml is the return of stocks with high book-to-market values minus the return of stocks with low book-to-market values.

In order to test validity of their model, two linear regressions, corresponding to the FF3 model and the CAPM, were performed on 6 portfolios constructed using a two-way grouping based on size and B/M, following the Chan *et al.* (1991). The first regression corresponds to CAPM:

$$E[R_i] - R_f = \alpha_i^{CAPM} + \beta_i^m(E[R_m] - R_f) + \epsilon_i^{CAPM}$$

whereas the second regression uses the three-factor model:

$$E[R_i] - R_f = \alpha_i^{FF3} + \beta_i^m(E[R_m] - R_f) + \beta_i^{smb}E[R_{smb}] + \beta_i^{hml}E[R_{hml}] + \epsilon_i^{FF3}$$

If the FF3 model is valid, the null hypothesis of $\mathcal{H}_0 : \alpha_i^{CAPM} = 0$ must be rejected in favor of the alternative hypothesis $\mathcal{H}_1 : \alpha_i^{FF} = 0$, the β_i^{smb} and β_i^{hml} estimates must be significant and R^2 should improve, i.e. $R^2_{FF} \geqslant R^2_{CAPM}$.

Further, regarding the validity of the FF3 factors, there is a strong international evidence, including Chan *et al.* (1991), Drew *et al.* (2003) and Bauer *et al.* (2010). The Fama-French model has since become a benchmark framework for the asset pricing along with the model presented next.

3.1.2 Fama-French-Carhart Factor Model

From the number of anomalies being investigated by early researchers in the field, anomalies related to the past returns were of a special interest. Bondt & Thaler (1985) found that long-horizon past loser stocks outperform long-horizon past winner stocks. On the other side of the horizon spectrum, Jegadeesh (1990) and Lehmann (1990) report other evidence of return reversals, but with a short-term horizon, i.e. stocks that have poorly performed in the previous week or month perform better in the next month. The most prominent form of the past returns based anomaly is the momentum strategy of Jegadeesh & Titman (1993), who found that buying stocks that have performed well over the past three to twelve months and selling stocks that have performed poorly produces abnormal positive returns.

Using this form of momentum, Carhart (1997) adds it to the FF3 model and uses this four-factor model to evaluate the performance of equity mutual funds.

$$E[R_i - R_f = \alpha_i^{FFC4} + \beta_i^m(E[R_m] - R_f) + \beta_i^{smb}E[R_{smb}] + \beta_i^{hml}E[R_{hml}] + \beta_i^{wml}E[R_{wml}] + \epsilon_i^{FFC4}$$

where R_{wml} is the return difference between winner and loser stocks over the past twelve months. Alongside, and perhaps exceeding the FF3 model in a role of a standard model in most studies on equity funds, it is called the Fama-French-Carhart four-factor model (FFC4).

3.2 Which Factors Really Matter?

Since the publication of Fama-French-Carhart risk factor model, hundreds of studies have tried to identify asset pricing anomalies that explain the cross-section of expected returns. Relevance of these anomalies and their relationships to expected returns is however not in the historical insight they offer. Their relevance depends on the extent to which these relationships hold outside a study's original sample. Out-of-sample performance is not only important as a proof that documented relation is not spurious but can make it clearer why cross-sectional return predictability is observed in the first place.

To this end, McLean & Pontiff (2016) compare in-sample returns, post-sample returns, and post-publication returns for a large sample of predictors. If return predictability in published studies results solely from statistical biases, then predictability should disappear out of sample. The term "statistical biases" is used to describe a broad array of biases that are consequences of poor research design (e.g. data mining), and are to some extent responsible for such a decline in out-of-sample return predictability. For example, Lo & MacKinlay (1990) argue that almost no empirical studies are free of data mining, which becomes more severe as the number of published studies performed on a single data set increases, Fama (1998) shows that value-weighting shrinks anomalies alphas, Schwert (2003) documents that anomalies return predictability often seem to disappear, reverse, or weaken after publication. Apart from statistical biases, the extent to which investors learn from the publication can also be responsible for difference between in-sample and post-publication returns. The concept is nicely explained by McLean & Pontiff (2016): "If return predictability reflects only rational expectations, then publication will not convey information that induces a rational agent to behave differently. Thus, once the impact of statistical bias is removed, pre- and post-publication return predictability should be equal. If return predictability reflects mispricing and publication leads sophisticated investors to learn about and trade against the mispricing, then we expect the returns associated with a predictor should disappear or at least decay after the paper is published. Decay, as opposed to disappearance, will occur if frictions prevent arbitrage from fully eliminating mispricing." Examples of such frictions include systematic noise trader risk (De Long *et al.* (1990)) and idiosyncratic risk and transaction costs (Pontiff (1996) and Pontiff (2006)). These effects can be magnified by principal-agent problems between investors and investment professionals (Shleifer & Vishny (1997)).

Based on 26% lower out-of-sample and 58% lower post-publication portfolio returns, greater post-publication declines for anomalies with higher in-sample returns and returns higher for portfolios concentrated in stocks with high idiosyncratic risk and low liquidity, McLean & Pontiff (2016) conclude that investors learn about mispricing from academic publications.

Another attempt at answering which factors really matter is Linnainmaa & Roberts (2016). They amass comprehensive accounting data from Moody's manuals from 1918 through 1963 when the standard finance research database Compustat becomes free of backfill bias and thus is the typical starting point of the time periods examined in anomalies literature. Merging these data with the Compustat and CRSP records, its coverage and the quality of publicly traded firms is similar before and after 1963. Subsequently they examine premiums for 38 anomalies and investigate three potential explanations for them: risk, mispricing, and data-snooping. Each of these explanations correspond to different testable hypotheses across three eras in their merged dataset: (1) pre-sample data existing before the discovery of the anomaly, (2) in-sample data used to identify the anomaly, and (3) post-sample data accumulating after identification of the anomaly. Results from pre-1963 data, only seven out of the 38 anomalies earn average returns that are positive and statistically significant at the 5% level, are consistent with data-snooping as the anomalies are clearly sensitive to the choice of sample period. As explained in McLean & Pontiff (2016), if the anomalies are risk-based, then we would have expected them to be similar across periods, absent structural breaks in the risks that matter to investors. If the anomalies are mispricing-based, then we would have expected them to be at least equal if not larger during the pre-discovery sample period[1]. The data-mining as the most probable explanation for most of the anomalies is also supported by the fact that the average anomaly becomes less profitable and more volatile either before or after the original study's sample period. This does not mean that all anomalies are spurious, however there is an ex-ante uncertainty of not knowing which anomalies are real and which are spurious. Out-of-sample testing is the cleanest way to determine which factors are true predictors of cross-sectional returns, however it cannot be used in the real time. A genuine out-of-sample test needs data in the future.

[1] Larger because of limits to arbitrage, e.g. higher transaction costs (Hasbrouck (2009))

In order to account for data mining in the anomalies literature, without necessary out-of-sample data available, Harvey *et al.* (2016) conduct a meta-study based on 312 papers concerned with cross-sectional return patterns published in the top journals in finance, economics and accounting or presented at top conferences. They employ four multiple testing frameworks, to obtain true for multiple testing adjusted threshold statistical significance level from the first empirical tests in 1967 through to present day. The threshold cutoff increases over time as even more anomalies are tried and newly discovered factor today should have a t-statistic over three. After going through 296 significant anomalies, they report that 80-158 (27%-53%) are false discoveries, depending on the specific methods of adjusting for multiple testing. The estimates are likely conservative because many factors have been tried by researchers but never reported, given they were unsuccessful. According to the authors, the main suspects among publication biases mainly responsible for the false discoveries in anomalies literature are difficulty to publish a negative result in top academic journals and difficulty to publish replication studies in finance and economics[2]. Consequently, there is strong incentive for publication of new factors rather than necessary verification of known factors. Because of this incentives, it is virtually impossible to publish weak or negative results.

Although one can relatively easily identify published variables, one cannot observe the variables that have been tried, but not published. Yan & Zheng (2017) solve this problem by examination of anomalies that are based on financial statements. By permuting 240 accounting variables with 15 base variables they form more than 18,000 fundamental signals, a universe a data snooper might face, construct long-short portfolios based on each fundamental signal and bootstrap the alphas of long-short returns and its significance. Results, robust to alternative universe of fundamental signals as well as alternative sampling procedure, suggest that the superior performance of the top fundamental signals cannot be attributed to pure chance. Following these results, they investigate whether they are consistent with mispricing- or risk-based explanations. They find that fundamental-based trading strategies are more pronounced among stocks with greater limits to arbitrage[3] as explained by Lakonishok *et al.* (1994), following high-sentiment periods as argued by Stambaugh *et al.*

[2]In other scientific fields, replication studies are common in top journals.

[3]E.g. small, low-institutional ownership, high-idiosyncratic volatility, and low-analyst coverage stocks

(2012), and have higher returns during recessions than during expansions. Overall, their results suggest that relatively large number of fundamental signals have true predictive ability and this predictive ability is more consistent with mispricing-based explanation.

Looking for factors that matter in a more conventional way, Green *et al.* (2017) focus on simultaneously including 94 firm characteristics as explanatory variables, while avoiding overweighting microcap stocks. Since the mean absolute correlation between these characteristics is low, it allows identification of characteristics that provide independent information about stock returns. They find that only 12 of 94 of characteristics provide independent information in non-microcaps during 1980-2014 and that just two characteristics matter after 2003. These results support conclusions of Harvey, Liu and Zhou's (2016) data snooping critique and McLean and Pontiff's (2016) finding of post-publication decay. On the other hand, significant differences in returns predictability in pre- versus post-2003 as well as in the case of different results based on size indicate that data snooping is not a complete explanation.

Further, Feng *et al.* (2017) study which out of the 114 factors in their analysis contain useful new information. In order to do so, every time a new factor is introduced, authors test whether its risk price is nonzero, controlling for all factors existing up to that point, and whether in future, when this factor will be part of potential controls, will be selected as part of the best model. Although majority do not contain useful new information, authors find that some factors (e.g., corporate investment) tend to appear throughout the 20 years considered, being selected almost all the time. Others appear at the beginning but are substituted with more modern ones. Results suggest that studying the marginal contribution of new factors relative to the vast set of existing ones is a conservative and productive way to screen new factors as they are proposed.

Very concerning is the case documented by McLean & Pontiff (2016) where 10 out of 82 studied anomalies in-sample performance of original studies could not be replicated. Until very recently, finance literature did not put enough attention to the replication of empirical results. The most famous general study is conducted by Ioannidis (2005) who argues that most research findings are false for most designs and for most fields. Echoing his troubling claims, Chang & Li (2015) fail to replicate more than half of 67 published papers from 13 economic journals. On the similar note, Brodeur *et al.* (2016) document a concerning "two-humped" pattern of test statistics from 50 000 tests in prestigious economic journals. First hump is associated with high p-values and the second hump with p-values slightly below 5%, suggesting that researchers look for specification that deliver just-significant results and ignore those that give just-insignificant results, thus commiting p-hacking.

To address the lack of comprehensive replication in anomalies literature, Hou *et al.* (2017) create the largest factor library of 447 pricing anomalies with definitions as well as description for transparent implementation. Covered anomalies span the categories from the momentum, value-versus-growth, investment, profitability, intangibles, to trading frictions. One of the reasons for diverse results often concerning the same characteristic is the wide range of research designs used in the pricing literature. To put anomalies on the equal footing, authors follow a common set of replicating procedures. Since whenever possible I follow these procedures in my empirical section I leave details, with exception of the key size cutoff and value-weighting, to that part. In replication, authors use value-weighting based on number of arguments. According to Fama (1998), value-weights more accurately reflect the wealth effect experienced by investors. Fama & French (2008) document that microcaps are influential in equal-weighted returns. Microcaps are stocks with the market equity below the 20th percentile of NYSE stocks. Microcaps are on average only 3% of the market value of the NYSE-Amex-NASDAQ universe, but represent around 60% of the total number of stocks.

There is a number of reasons for eliminating microcaps. One of the assumptions of APT is functioning markets preventing arbitrage opportunities. This is unlikely to be the case for microcaps because of their transaction costs and illiquidity. Asparouhova *et al.* (2013) document serious microstructure frictions, e.g. bid-ask spreads, nonsynchronous trading, discrete prices which can cause bias in cross-sectional equal-weighted returns.

Typically, researchers use NYSE-Amex-NASDAQ cutoffs for determining which stocks are microcaps. Fama & French (2008) show that using this procedure populates the extreme deciles in number of anomalies by microcaps, that represent more than 60% of them.

On the contraty, Hou *et al.* (2017) use NYSE breakpoints. This smoothes the impact that microcaps have in extreme deciles. They show on data from

1967 to 2014, and I verify this claim, that there are 2,406 microcaps on average, which account for 61% of the total number of firms, 3,938. This is in contrast to microcaps representing only 3.28% of the total market capitalization, small stocks 6.77%, and big stocks 90%. With equal-weights, microcaps earn on average 1.32% per month relative to 1.03% for big stocks. In contrast, the value-weighted market return of 0.93% is close to 0.92% for big stocks.

Further, authors use portfolio sorts instead of Fama & MacBeth (1973) cross-sectional regressions. Their reasoning serves as a nice summary for potential pitfalls when using cross-sectional regression without proper diligence and I refer reader to their paper for more details.

They treat an anomaly as a replication failure if the average return of its high-minus-low decile is insignificant at the 5% level (t ¡ 1.96). Results are not good for anomalies literature, 286 out of 447 anomaly variables (64%) are insignificant.

However, none of these single mispricing factors (i.e. HML, SMB, MOM) can accommodate a large set of anomalies very well. Clearly, when a larger set of anomalies is considered, old models do not have a strong ability to explain the larger number of anomalies analyzed. Of course, this may not be too surprising since anomalies are typically only deemed "anomalous" once it is determined that e.g. FF 3-factor model canâ€™t explain the excess returns associated with the strategy that generates the anomaly. Indeed, it is a sign of a good asset pricing model to be able to explain number of anomalies, so they are not "anomalous" anymore. In this spirit, authors use Hou *et al.* (2015) q-factor model, motivated from investment-based asset pricing, to explain the successfully replicated anomalies.

The q-factor model is based on the long line of literature concerned about relationship between the production side and asset prices, including Cochrane (1991), Berk *et al.* (1999) and Liu *et al.* (2009). The main idea is quite simple, investment has predictive power with respect to returns because given expected cash flows, high costs of capital mean low net present values of new capital and low investment, whereas low costs of capital mean high net present values of new capital and high investment. Reasoning for relation of return on equity to returns is based on the discount rates. In their model, high discount rates are equivalent to high expected return on equity (ROE) associated with low investment. This is because the only reason for low investment, when there is high expected ROE, is to have high discount rates that prohibit more investment. The same goes for low ROE with high investment.

In q-factor model the expected return of an asset in excess of the riskfree rate is described by the sensitivity of its return to four factors. They are the market excess return (MKT), the difference between the return on a portfolio of small-market equity stocks and the return on a portfolio of big-market equity stocks (r_{ME}), the difference between the return on a portfolio of low-investment stocks and the return on a portfolio of high-investment stocks ($r_{\Delta A/A}$), and the difference between the return on a portfolio of high return-on-equity stocks and the return on a portfolio of low return-on-equity stocks (r_{ROE}).

$$E[R_i]-R_f = \beta_i^m(E[R_m]-R_f)+\beta_i^{ME}E[R_{ME}]+\beta_i^{\Delta A/A}E[R_{\Delta A/A}]+\beta_i^{ROE}E[R_{ROE}]$$

The size factor's main role is to lower the average magnitude of the alphas across size-related portfolios. According to the authors, compared to the ROE and the investment, the size factor plays only a secondary role in the q-factor model.

The q-factor model is able to explain, in an internally consistent and economically meaningful way, number of anomalies that neither Fama-French model nor Fama-French-Carhart model can. Anomalies that cannot be explained, like earnings surprise, idiosyncratic volatility, financial distress, net stock issues, composite issuance, etc. have lower alpha when using q-factor model.

Given the relative success of q-factor model in explaining number of significant anomalies in the greatest up-to-date replication study of Hou *et al.* (2017), it is valid to ask if there is no better asset pricing model.

There are two well-known and promising models in current literature. Fama & French (2015) five factor model and Stambaugh & Yuan (2016) four factor model based on the so-called mispricing factors.

Fama & French (2015) propose a five-factor model (FF5) by adding two factors into their 3-factor model. Its factors being the market excess return, size (SMB), value (HML), operating profit (RMW), and asset growth (CMA).

$$E[R_i]-R_f = \beta_i^m(E[R_m]-R_f)+\beta_i^{ME}E[R_{ME}]+\beta_i^{\Delta A/A}E[R_{\Delta A/A}]+\beta_i^{ROE}E[R_{ROE}]$$

Factors operating profit and asset growth are somewhat different versions than those in q-factor model. They provide theoretical motivation in a form of comparative statics of a present-value relation.

More interestingly, Stambaugh & Yuan (2016) construct a four-factor model consisting of market, size and two "mispricing" factors. The main idea behind the model is that some anomalies reflect, at least partially, mispricing. This mispricing should be manifested in a common way across stocks. The two mispricing factors are based on 11 anomalies, examined earlier by Stambaugh *et al.* (2012) and Stambaugh *et al.* (2015). Out of studies originally documenting underlying anomalies, studies containing mispricing interpretations include Ritter (1991) for net stock issues (NSI), Daniel & Titman (2006) for composite equity issues (CEI), Sloan (1996) for accruals (ACCR), Hirshleifer *et al.* (2004) for net operating assets (NOA), Cooper *et al.* (2008) for asset growth (AG) , Titman *et al.* (2004) for investment-to-assets (IA), Campbell *et al.* (2008) for financial distress (DIST), Jegadeesh & Titman (1993) for momentum, and Wang & Yu (2013) for profitability anomalies including return on assets (ROA) and gross profitability (GP). The underlying 11 anomalies are first classified into two clusters using a correlation-based distance measurement and a clustering method.

The first cluster of anomalies includes NSI, CEI, ACCR, NOA, AG, and IA. Since these anomalies are more or less directly impacted by managers' actions, factor is called by authors as MGMT-characteristic.

The second cluster includes DIST, O, MOM, GP and NOA. Since these anomalies are related more to performance and less directly or immediately impacted by managers' actions, it is called by authors as PERF-characteristic. It is important to note that even though authors assign names to the clusters, they do not claim that cluster must reflect only one and only behavioral story.

Following Stambaugh *et al.* (2015), who average across all 11 anomalies constructing a single composite mispricing measure, individual stock's rankings are average within each of the two clusters. Stock's rankings across anomalies are equal weighted. By combining information across anomalies, authors strive to construct factors capturing common elements of mispricing. Finally, the two mispricing factors are finished by applying a standard $2x3$ sorting procedure.

In detail, each month NYSE, AMEX, and NASDAQ stocks, eliminating stocks with prices smaller than $5, are sorted by size and split into big and small groups based on the NYSE median size. Subsequently, all stocks are sorted independently according to PERF-characteristic and assigned to three groups using the 20th and 80th percentiles of the combined NYSE, AMEX, and NASDAQ universe as breakpoints. The same procedure is applied in the case of MGMT-characteristic. In the end, value-weighted returns are calculated for each of the portfolios formed by the intersection of the two size categories and extreme categories for both clusters. The value of factor for a given month is the simple average of the returns on the two low PERF-characteristics portfolios minus the average of the returns of the two high PERF-characteristics.

Additionally, they construct a three-factor model with a single mispricing factor, following already mentioned Stambaugh *et al.* (2015).

As authors, Hirshleifer & Jiang (2010) and Kozak *et al.* (2017) admit though, mispricing factors can capture systematic risks for which investors require compensation as well, and to make things more complicated, there does not have to be a strict distinction between mispricing and risk compensation. Nevertheless, a mispricing interpretation pushed forward is also partially supported by the evidence of McLean & Pontiff (2016), given the fact that following an anomaly's academic publication, there is an increase in trading volume and "arbitrage" profits tend to decline.

Fama & French (2016) study the ability of the FF5 model of Fama & French (2015) to explain a small set of return anomalies. Hou *et al.* (2015) compare the FF5 model with their q-factor model in explaining a range of anomalies. Finally, Stambaugh & Yuan (2016) compare their mispricing factor model, both 4-factor version as well as 3-factor version with only one mispricing factor, with all of the above and the FF3 model of Fama & French (1993).

Models are compared in number of dimensions, including explaining 73 anomalies covered by Hou *et al.* (2015), evaluating each model's ability to explain factors from the another and using Bayesian posterior model probabilities, developed by Barillas & Shanken (2015).

The mispricing four-factor model outperforms other models in explaining 73 anomalies considered. Specifically, the Gibbons *et al.* (1989) test, concerned whether all the anomalies' alphas equal zero, produces a p-value of 0.10 for the mispricing model compared to 0.003 or less for its competitors. It also outperforms other 4 models in terms of explaining each other's factors. The three-factor mispricing model, from which the sole mispricing factor MISP is used in this thesis, performs on par with the four-factor version.

Perhaps this isn't surprising, since the factor model used to "control" for the anomalies is built using the anomalies. However, the same critique can be said for the other factor models, e.g. value is not significant when controled for HML factor in the FF3 model.

At a minimum, for researchers looking to identify "new" ideas, this factor model is a great tool. If one identifies a significant "alpha" after controlling for the mispricing factor model, there is a good chance there may be something special associated with the strategy under review.

Chapter 4

Risk-based Approaches

As presented in Section 2.1.1 there are severe problems with applying mean-variance framework as a proper portfolio selection method. The main reason in the estimation error for the input parameters μ and Σ making direct application of the mean-variance approach towards portfolio optimization prohibitive from the practitioner's as well as from the scientific point of view. In light of this, risk-based portfolio selection methods proceed by giving less room to estimation errors by not requiring any explicit or implicit effort to model, measure or predict returns. As a consequence, it is not possible to form a set of recommended portfolios for different desired levels of risk. This problem is often solved through leverage, in case of multi-asset class allocation problem. Given the fact that in my research setting, portfolio selection methods are applied not only within one asset class, but specifically in rather narrow universe of factor portfolios, usage of leverage is not necessary and probably ill-advised given the strong exposure towards underlying factor risk.

Different risk-based solutions have been suggested in finance literature over the year, with the most prominent being investigated in my thesis:

- the risk parity strategy
- the equally-weighted strategy
- the minimum-variance strategy
- the naive risk parity strategy
- the optimal risk parity or equal risk contribution strategy
- the most diversified portfolio strategy

- the diversified risk parity strategy
- the hierarchical risk parity strategy

When saying that some portfolio is well-diversified, it is expected that it is immune against shocks created by a single or a few assets. Given the portfolio of risky assets, a reliable measurement of the degree of diversification can only be achieved by incorporating the dependence structure among these assets, i.e. the information of how the portfolio constituents interact. Typically, this is done by evaluating the covariance matrix Σ.

In the following sections, individual portfolio selection methods, later used in the empirical part, are introduced. The Section 4.2 contains summary, comparison and expectations of these methods going forward.

4.0.1 The Equally-Weighted Approach

The most elementary and naive approach to measure the diversification of a portfolio of risky assets is to count the number of its constituents. It is build upon the well-documented fact that the return variance of an equally-weighted portfolio declines with the number of its constituents. As a diversification measure however, it completely ignores the heterogeneity of the assets as well as different weighting schemes.

The most straightforward risk-based strategy based on this concept, where portfolio allocations have only to be driven by the number of assets in the investment universe it the equal-weighting. Portfolio selection problem is solved using simple $1/N$, formally :

$$w_i = w_j = \frac{1}{N} \quad \forall i, j$$

Hence, when the equally-weighted approach is used, all assets are given an identical and static weight. It is obvious that since there is no optimization problem to solve, the equally-weighted portfolio allocation is determined regardless of any statistical estimate for returns, risks and correlations. A conditions for ex-ante optimality for equally-weighted portfolio, in a mean-variance framework, are very unrealistic: equal expected returns, equal volatilities and uniform correlations among assets. Given these unrealistic conditions, why would investors use the equally-weighted strategy? First of all, DeMiguel *et al.*

(2009) find using several asset allocation models and different datasets, that none of the theoretically more robust asset allocation models was consistently better out-of sample than the naive $1/N$ rule. Second, according to the behavioral finance the equally-weighted portfolio can be a reasonable way to avoid regret that would occur in case of more concentrated but unfavorable bet.

Nevertheless, the assumption that using this simple and heuristic method provides diversification can be misleading: it depends on the characteristics of assets that the strategy does not consider, e.g. if their risk is very different, the equally-weighted scheme can lead to concentrated risk loadings as will be clear in Section 4.1.2

4.0.2 The Value-weighted Portfolio

According to the CAPM theory, there is a linear relationship between portfolio return and risk, thus the existence of a market portfolio. The market portfolio consists of all risky assets and their weights are their share of the total financial market's value. The market portfolio is said to be completely diversified and the risk the risk associated with it is called systematic or non-diversifiable. As Roll (1977) points out, the market portfolio is a theoretical, unobservable object. As such, it cannot be used as benchmark of diversification, only it very limited proxies, e.g. index, could. The reason I include it in my empirical tests is twofold. First, it is a standard weighting scheme in anomalies literature. Second, from the CAPM side, it is connected to diversification based on the fact, that if my factor portfolio would be the market portfolio (which it clearly is not) then in equilibrium, every investor is expected to hold some value-weighted fraction of it.

Even though the turnover of this strategy is zero in case of stable universe, its turnover is on par with other strategies given the fact that universe changes as different stocks fall into the extreme factor deciles.

The above described framework constitutes the theoretical definition of diversification implied by the CAPM but this does not mean that there is no other portfolio with smaller return variance. By its very definition, the so-called Global Minimum-Variance Portfolio always allows for lower return variance.

4.0.3 The Global Minimum-Variance Approach

Global minimum-variance portfolio or simply minimum-variance portfolio (MV) lies at the start of the efficient frontier. It is expected to have the lowest possible volatility out of every portfolio in a given universe. As with the Markowit's portfolio, model parameters are not observable ex-ante, thus they must be estimated. Using the estimates of correlation matrix, weights determining optimization can be performed. If MV portfolio is supposed to be both fully invested and restricted from short-selling, as is usually the required case, the optimization have to contain the two traditional constraints. Optimization problem then takes the following form:

$$\min_{w} \left(\sum_{i=1}^{N} (w_i^2 \sigma_i)^2 + \sum_{i=1}^{N} \sum_{\substack{i=1 \\ i \neq j}}^{N} w_i w_j \sigma_i \sigma_j \rho_{ij} \right)$$

$$\text{s.t.}$$

$$\sum_{i=1}^{N} w_i = 1$$

$$w_i \geq 0$$

or in the matrix form:

$$\min_{w} \mathbf{w}'\mathbf{\Sigma}\mathbf{w}$$

$$\text{s.t.}$$

$$\mathbf{w}'\mathbf{e} = \mathbf{1}$$

$$w_i \geq 0$$

where $\boldsymbol{e}$ is a $Nx1$ vector of ones.

Given the inequality constraint in the optimization setting, analytical solution is not possible and iterative procedures have to be used in order to obtain the optimal weights. Few comments regarding the properties of minimum variance portfolios are appropriate.

First, the MV portfolio does not have to be invested in all assets in the universe. Second, the marginal risk contribution[1] is the same for all assets

[1] As defined in the next Section 4.1

belonging to the optimal portfolio. Formally:

$$\frac{\partial \sigma_p}{\partial w_i} = \frac{\partial \sigma_p}{\partial w_j} \quad \forall i, j$$

This has to hold, otherwise increasing weight of one asset and at the same time decreasing weight of the other would mean lower risk, leading to contradiction.

Third, since assets have different weights, then even though marginal risks are equal, it does not mean that total risk contributions are equal.

Fourth, there is a difference between the portfolio weight for an asset in the MV portfolio and the percentage contribution to the overall risk from that asset. Therefore the MV portfolio is not truly diversified, even though it is supposed to have the lowest variance.

4.1 Risk Parity Strategies

Based on the design of the optimization problem, the MV portfolios tend to load on low-volatility assets.

In contrast, risk parity approach recommends to build the portfolio in a way that its risk is equally distributed among assets. Focus in on the allocation or budgeting of risk. Question to answer is then how much does specific asset contribute to overall risk?

4.1.1 The Naive Risk Parity Strategy

The simplest and the easiest to solve form of risk parity strategy is commonly referred to as naïve risk parity (NRP). Optimal weights can be calculated without taking into account the correlation information between assets.

$$w_i = \frac{\sigma_i^{-1}}{\sum_{j=1}^{N} \sigma_j^{-1}}$$

Consequently, the higher the volatility of an asset, the lower its weight in the portfolio. However, contributions to the portfolio volatility are equal only in two cases; when there are only two assets in the portfolio and when all correlations in the portfolio are equal. It is clear that in reality, naïve risk parity is not an authentic risk parity strategy defined in Section 4.1.2.

4.1.2 The Optimal Risk Parity Strategy

The authentic risk parity strategy, as opposed to the NRP strategy, is called optimal risk parity or equally weighted risk contribution (ERC) strategy. It suggests looking at portfolio through risk contributions instead of weights.

Consider a portfolio of N assets, let w_i be the weight of the asset i and $R(w_1, \dots, w_n)$ be a risk measure for the portfolio $\boldsymbol{w} = (w_1, \dots, w_n)$. In the following introduction into the topic I follow Maillard *et al.* (2010) and rely on foundations of risk measures provided by Artzner *et al.* (1999).

If the risk measure is coherent and convex, it satisfies the Euler decomposition

$$R(w_1, \dots, w_n) = \sum_{i=1}^{n} w_i \frac{\partial R(w_1, \dots, w_n)}{\partial w_i}$$

Marginal risk of asset i is weighted by its portfolio weight. Summed across the portfolio it gives the risk measure. The risk contribution of the i-th asset is then:

$$RC_i(w_1, \dots, w_n) = w_i \frac{\partial R(w_1, \dots, w_n)}{\partial w_i}$$

The risk budgeting portfolio should satisfy the following conditions:

$$\begin{cases} RC_1(w_1, \dots, w_n) = b_1 \\ \quad \vdots \\ RC_2(w_1, \dots, w_n) = b_2 \\ \quad \vdots \\ RC_n(w_1, \dots, w_n) = b_n \end{cases}$$

where $\{b_1, \cdots, b_n\}$ are risk budgets, e.g. amount of risk measured in absolute terms. In order to assign weights and obtain the risk budgeting portfolio, the above mentioned conditions in the form of system of nonlinear equations have to be solved.

The core risk measure in general, as well as in my thesis, is the volatility risk measure,

$$R(\boldsymbol{w}) = \sigma(\boldsymbol{w}) = \sqrt[2]{\boldsymbol{w^T \Sigma w}}$$

In case of volatility, the marginal risk and the risk contribution of asset i are then defined as:

$$\frac{\partial R(\boldsymbol{w})}{\partial w_i} = \frac{(\boldsymbol{\Sigma w})_i}{\sqrt[2]{\boldsymbol{w^T \Sigma w}}}$$
$$RC_i(\boldsymbol{w}) = w_i \frac{(\Sigma w_i)}{\sqrt[2]{w^T \Sigma w}}$$

In general, a proper long-only risk budgeting portfolio is defined as:

$$\begin{cases} w_i(\boldsymbol{\Sigma w})_i = b_i(\boldsymbol{w^T \Sigma w}) \\ b_i \geq 0 \\ w_i \geq 0 \\ \sum_{i=1}^{N} b_i = 1 \\ \sum_{i=1}^{N} w_i = 1 \end{cases}$$

The core idea of ERC strategy is preventing one or few assets from driving portfolio risk. Consequently, the strategy aims at equalizing risk contributions from the different assets. Risk contributions going forward are unknown, therefore need to be estimated again on ex-ante basis. The ERC strategy wants the portfolio to be equally weighted in terms of risk allocations. Optimization in the ERC strategy consists of searching for such a portfolio weights such that risk budgets correspond to weights assigned to each asset. Formally:

$$\begin{cases} w_i \frac{\partial \sigma_p}{\partial w_i} = b_i \frac{w_i(\boldsymbol{\Sigma w})_i}{\sqrt[2]{\boldsymbol{w^T \Sigma w}}} \\ b_i = \frac{1}{N} \\ \sum_{i=1}^{N} w_i = 1 \\ w_i \geq 0 \end{cases}$$

In order to obtain the optimal weights, nonlinear problem has to be put into a constrained optimization framework. Reasonable objective function is, however, required. The main goal of the strategy is to have the same risk contribution from each component, therefore the condition, from which the objective function can be derived, is formally defined as:

$$w_i \frac{\partial \sigma_p}{\partial w_i} = w_j \frac{\partial \sigma_p}{\partial w_j}; \qquad \forall i, j$$

From this equation, the optimization problem to solve for the ERC portfolio can be fully stated as follows:

$$\min_{w} \sum_{i=1}^{N}\sum_{j=1}^{N}\left(w_i\frac{\partial\sigma_p}{\partial w_i} - w_j\frac{\partial\sigma_p}{\partial w_j}\right)^2$$

$$\text{s.t.}$$

$$\sum_{i=1}^{N} w_i = 1$$

$$0 \le w_i \le 1$$

This is a constrained nonlinear programming problem for which analytical solutions are not available. To find them, it is necessary to use a numerical algorithm of an iterative nature. Maillard *et al.* (2010) and Roncalli (2014) suggest the use of the Sequential Quadratic Programming Algorithm (SQP) and I follow their steps in the empirical part of this thesis. This method generates approximated solutions that allow convergence to the minimum of the nonlinear optimization problem. The tentative solutions are obtained using the current iterate simultaneously to replace the original objective function with a quadratic problem approximation and to approximate the constraint functions after linearizing them.

4.1.3 The Most Diversified Portfolio Approach

Choueifaty [2006] proposed a measure of portfolio diversification, called the Diversification Ratio (DR), which he defined as the ratio of the portfolio's weighted average volatility to its overall volatility. This measure embodies the very nature of diversification whereby the volatility of a long-only portfolio of assets is less than or equal to the weighted sum of the assets' volatilities.

$$DR_p = \frac{\sum_{i=1}^{N} w_i\sigma_i}{\sqrt{\sum_{i=1}^{N} w_i^2\sigma_i^2 + \sum_{i=1}^{N}\sum_{\substack{i=1\\ i\neq j}}^{N} w_i w_j \rho_{ij}}} = \frac{\sum_{i=1}^{N} w_i\sigma_i}{\sigma_p}$$

Using the matrix notation, it is written alternatively as:

$$DR_p = \frac{w'\sigma}{\sqrt{w'\Sigma w}} = \frac{w'\sqrt{diag(\Sigma)}}{\sqrt{w'\Sigma w}}$$

The DR of a long-only portfolio is greater than or equal to one. It is equal to one in case when portfolio consists of only one asset or when it consists of perfectly positively correlated assets, because in this cases numerator and denominator would be the same. The DR close to 1 means the portfolio is poorly diversified. On the other side, the DR does note have an upper bound. The higher the DR of the portfolio, the more diversified the portfolio is. In essense, the DR of a portfolio measures the benefit from holding assets that are not perfectly correlated.

This is clear from the fact that the numerator of the diversification ratio computes the volatility of a given portfolio if all pair-wise correlations were equal to one, which corresponds to the maximum value theoretically admissible for portfolio standard deviation. Meucci (2010) points out that since the DR emphasizes how far two measures of volatility are for the same portfolio, it is a relative measure of diversification, not an absolute one.

The most diversified approach consists in the maximization of the diversification ratio in an objective function. Result of this optimization is the most diversified portfolio (MDP). Corresponding optimization problem looks like this:

$$\max_w \frac{\sum_{i=1}^N w_i\sigma_i}{\sqrt{\sum_{i=1}^N w_i^2\sigma_i^2 + \sum_{i=1}^N \sum_{\substack{i=1\\ i\neq j}}^N w_i w_j \rho_{ij}}}$$

$$\text{s.t.}$$

$$\sum_{i=1}^N w_i = 1$$

$$0 \leq w_i \leq 1$$

In a matrix notation:

$$\max_w \frac{w'\sqrt{diag(\Sigma)}}{\sqrt{w'\Sigma w}}$$

$$\text{s.t.}$$

$$\sum_{i=1}^N w_i = 1$$

$$0 \leq w_i \leq 1$$

As shown by Choueifaty *et al.* (2013), the long-only MDP always exists and is unique when the covariance matrix is definite. They also provide an extremely insightful interpretation among other mathematical properties of the DR and MDP. Assuming an universe of F independent risk factors and a portfolio with each risk factor exposure inversely proportional to the underlying factor's variance, allocates the risk budget equally across all risk factors. Therefore, its DR squared (DR^2) equals F, the number of independent risk factors or degrees of freedom represented in the portfolio.

An equivalent definition of the MDP is also offered by Choueifaty *et al.* (2013): "Any stock not held by the MDP is more correlated to the MDP than any of the stocks that belong to it. Furthermore, all stocks belonging to the MDP have the same correlation to it." and called the core property of the MDP. It means that even though portfolio does not contain some of the assets from the universe, it effectively represents the whole underlying universe. From an another perspective, it's also shown that "The long-only MDP is the long-only portfolio such that the correlation between any other long-only portfolio and itself is greater than or equal to the ratio of their DRs."

For example, an MDP portfolio constructed using some index of stocks may hold just a few percent of it. That does not mean, however, that this portfolio is not diversified, as the remaining stocks that are not part of the portfolio are more correlated to the MDP compared with the few percent it actually holds. This is consistent with the notion that the MDP is the undiversifiable portfolio.

Theoretical as well as empirical results of Choueifaty & Coignard (2008) suggest that the MDP is more efficient ex post than the value-weighted benchmark, minimum-variance portfolio, and equal-weight portfolio, moreover its clear that the MDP is a strong candidate for being the un-diversifiable portfolio, and as such should offer the full exposure to any underlying factor premium.

4.1.4 Diversified Risk Parity

Another approach towards assessing the degree of a portfolio's diversification stems from information theory. Loosely speaking, information theory is concerned with the quantification of the disorder of a random variable, with its most prominent measure being the Shannon entropy. Woerheide & Persson (1992) introduce measures from information theory as well as measures of economic concentration to portfolio theory in order to assess the concentration of weights on individual assets. Thus, their approach of measuring the diversification of a portfolio depends not only on the number of assets, but also on the fractions of wealth invested into the assets. However, given the portfolio of risky assets, a reliable measure of a level of diversification must take into the account sources of risk. To achieve this, Meucci (2010) builds on principal component analysis of the portfolio assets to extract the main drivers of its risk.

In case of a portfolio consisting only of uncorrelated constituents, its variance is a sum of variances of individual securities.

$$Var(R_w) = \sum_{n=1}^{N} Var(w_n R_n)$$

In a portfolio with correlated constituents, this relationship does not hold. Nevertheless, even if constituents are correlated, it is always possible to identify uncorellated sources of risk. The most straightforward method is the principal component decomposition of the covariance matrix. According to the spectral decomposition theorem,Σ can be expressed as a product

$$\Sigma = E \Lambda E'$$

where $\Lambda = diag(\lambda_1, \cdots, \lambda_n)$ is a diagonal matrix consisting of Σ's eigenvalues that are assembled in descending order, $\lambda_1 \geq \cdots \geq \lambda_n$. These eigenvectors define a set of N uncorrelated principal portfolios with variance λ_i for $i = 1, \cdots, N$ and returns $\tilde{R} = E'R$. This decomposition holds for any universe with a well-defined covariance, and not necessarily for normal markets only. Based on Partovi *et al.* (2004), it is easy to see that portfolio can be considered to be either a combination of its constituents with corresponding weights or a combination of the uncorrelated principal portfolios with weights $\tilde{w} = E'w$.

From the definition, the principal portfolios are uncorrelated, thus the portfolio variance is simply a variance weighted average of the principal portfolios' variances λ_i:

$$Var(R_w) = \sum_{i=1}^{N} \tilde{w}_i^2 \lambda_i$$

To get to the diversification distribution, concept introduced by Meucci (2010), the variance is simly normalized:

$$p_i = \frac{\sum_{i=1}^{N} \tilde{w}_i^2 \lambda_i}{Var(R_w)}, \quad i = 1, \cdots N$$

The diversification distribution p is always positive and such that sum of all p_i is always unity.

In the spirit of equal risk budgeting, a portfolio is considered to be well-diversified when all p_i are of similar magnitude and the diversification distribution has a uniform shape. On the other hand, portfolios with an exposure to a single source of risk, i.e. a single principal portfolio, have more peaked distribution. In order to measure the degree of dispersion in the diversification distribution, the Shannon entropy can be used. Specifically, because of ease of interpretation, it is better to use exponential of the Shannon entropy:

$$\mathcal{N}_{Ent} = exp(-\sum_{i=1}^{N} p_i ln(p_i))$$

where $\mathcal{N}_{Ent}$ can be interpreted as the number of uncorrelated bets. As an example, a completely concentrated portfolio is characterized by $p_i = 1$ for one i and $p_j \neq 0$ for $i = j$ resulting in an entropy of 0 which implies $\mathcal{N}_{Ent} = 1$. On the other side of the spectrum, $\mathcal{N}_{Ent} = N$ holds for a portfolio that is completely homogenous in terms of uncorrelated risk sources. In this case, $p_i = p_j = 1/N$ holds for all i, j implying an entropy equal to $ln(N)$ and $\mathcal{N}_{Ent} = N$.

The number of effective bets as a measure of diversification can be used both in terms of ex-post measurement of actual diversification achieved and as a portfolio selection method. The latter is manifested in searching for the maximum diversification portfolio, or by Lohre *et al.* (2012) called diversified risk parity (DRP) portfolio. The DRP long-only portfolio weights w_{DRP} can be obtained by solving the following optimization problem:

$$\begin{aligned}
&\max_{w} \mathcal{N}_{Ent}(w) \\
&\text{s.t.} \\
&\sum_{i=1}^{N} w_i = 1 \\
&0 \leq w_i \leq 1
\end{aligned}$$

4.1.5 Hierarchical Risk Parity Strategy

In order to deal with insufficient accuracy of returns forecasts, modern portfolio selection approaches tend to focus on the risk side. It is still necessary to rely on quadratic optimizers and thus on the inversion of a positive-definite covariance. DeMiguel *et al.* (2009) show, that even after keeping forecasted returns out of the picture, many of the best known quadratic optimizers underperform the equal-weighting benchmark. López de Prado (2016) clarifies that the main reason is the inversion of a positive-definite covariance matrix. Stability of the inverse matrix depends on the condition number of a covariance matrix, i.e. the ratio between the highest and smallest eigenvalues. The more correlated the assets, the greater the condition number. The more correlated the assets, the greater the need for diversification, and yet the more unstable the matrix inverse. One of the main reasons for the instability of quadratic optimizers is that the vector space is represented as a complete, fully connected, graph. In these terms, the matrix inversion consists of evaluating the rates of substitution across the complete graph. This has an unfortunate side effect that small estimation errors over several edges lead to wrong inversions.

An attempt to provide a solution is to bring the concept of hierarchy into the picture.

López de Prado (2016) introduced hierarchical portfolio construction (HRP) method addressing the absence of concept of hierarchy in correlation matrices. Using graph theory and machine learning, it is no longer necessary for correlation matrix to be invertible or positive-definite in order to leverage information it contains, as was the case with quadratic optimizers. HRP algorithm consists of three parts:

1. **Tree clustering**: Grouping similar investments into clusters.
2. **Quasi-diagonalization**: Reorganizing the rows and columns of the covariance matrix, so that the largest values lie along the diagonal.
3. **Recursive bisection**: Splitting allocations through recursive bisection of the reordered covariance matrix.

As previously, assume N variables over T periods, i.e. $T \times N$ data matrix X consisting of returns series and $N \times N$ correlation matrix with elements $\rho = \{\rho_{i,j}\}_{i,j=1,\dots,N}$

Tree clustering Using the distance measure $d : (X_i, X_j) \subset B \to \mathbb{R}, d_{i,j} = \sqrt{\frac{1}{2}(1-\rho_{i,j})}$, where B is the Cartesian product of items in {1,...,N}, we compute from correlation matrix a $N \times N$ distance matrix $D = \{d_{i,j}\}_{i,j=1,\dots,N}$.

Matrix D is a proper metric space[2].
Next, I compute the Euclidean distance of distances matrix with entries $\tilde{d}$: $(D_i, D_j) \subset B \rightarrow \mathbb{R} \in [0, \sqrt{N}], \tilde{d}_{i,j} = \sqrt{\sum_{n=1}^{N}(d_{n,i} - d_{n,j})^2}$ Each $\tilde{d}_{i,j}$ is a function of the entire correlation matrix[3], therefore $\tilde{d}$ is defined over the entire metric space D.
Having matrix $\tilde{D}$ I can proceed with clustering together two collumns $(i*, j*) = argmin_{i,j}\{\tilde{d}_{i,j}\}$ with resulting cluster $u[1]$.
In order to incorporate new cluster $u[1]$ into the distance of distances matrix $\tilde{D}$ I need to perform a linkage creation, i.e. define the distance between a new cluster and the unclustered (original) items: $\dot{d}_{i,u[1]} = min[\{\tilde{d}_{i,j}\}_{j \in u[1]}]$
By dropping the clustered columns and rows and appending $\dot{d}_{i,u[1]}$ I update matrix $\{\tilde{d}_{i,j}\}$.
Repeating the last three steps recursively gives us $N-1$ clusters appended to matrix D.

The last four steps can be described by a $(N-1) \times 4$ linkage matrix Y, i.e. for each cluster one row of 4 elements, where first two report the merged constituents, third reports distance between them and the fourh shows the number of original items in the cluster.

Quasi-diagonalization In this part of the algorithm correlated items are placed close to each other and uncorrelated far apart. This quasi-diagonalization is achieved by replacing clusters with their components recursively, until no clusters remain. This is done inside of the linkage matrix, preserves order and returns a sorted list of original (unclustered) items.

Recursive bisection From the Quasi-diagonalization part of the algorithm I have a quasi-diagonal matrix. For a diagonal covariance matrix the inverse-variance allocation is optimal[4], therefore I can define the variance of a continuous subset as the variance of an inverse-variance allocation and split allocations between adjacent subsets in inverse proportion to their aggregated variances. The algorithm consists of the following steps:

1. Initialization

[2] Satisfying non-negativity, coincidence, symmetry and sub-additivity.
[3] $d_{i,j}$ is only function of a particular correlation pair.
[4] For the proof see López de Prado (2016).

Set the list of items: $L = \{L_0\}$, with $L_0 = \{n\}_{n=1,...,N}$ and assign a unit weight to them: $w_n = 1, \forall n = 1, ..., N$

2. If $|L_i| = 1, \forall L_i \in L$, then stop

3. For each $L_i \in L$, s.t. $|L_i| > 1$:

   (a) bisect L_i into two subsets, $L_i^{(1)}$ and $L_i^{(2)}$, where $|L_i^{(1)}| = int[\frac{1}{2}|L_i|]$[5]

   (b) for $j = 1, 2$, define $var(L_i^{(J)}) = \tilde{V}_i^{(j)} = \tilde{w}_i^{(j)} V_i^{(j)} \tilde{w}_i^{(j)}$, where $V_i^{(j)}$ is the covariance matrix between the constituents of the $L_i^{(j)}$ bisection and $\tilde{w}_i^{(j)} = diag[V_i^{(j)}]^{-1} \frac{1}{tr[diag[V_i^{(j)}]^{-1}]}$

   (c) compute the split factor: $\alpha_i = 1 - \frac{\tilde{v_i}^{(1)}}{\tilde{v_i}^{(1)} + \tilde{v_i}^{(2)}}$, so that $0 \leq \alpha_i \leq 1$

   (d) re-scale allocations w_n by a factor of $\alpha_i, \forall n \in$

   (e) re-scale allocations w_n by a factor of $(1 - \alpha_i), \forall n \in L_i^{(2)}$

4. Go back to step two.

The variance of the partition L_i^j is calculated using inverse-variance weightings $\tilde{w}_i^{(j)}$ in step 3.2. leveraging quasi-diagonalization in a bottom-up way. Step 3.3. in algorithm is a top-down weight allocation of variance in inverse proportion to the cluster's variance.
In the algorithm described above, one is fully invested and shorting is not allowed, i.e. $0 \leq w_i \leq 1, \forall i = 1, ..., N$ and $\sum_{i=1}^{N} w_i = 1$. This holds, because we are only splitting weights received from higher hierarchical levels. If we would like to change the constraints, e.g. allow shorting, leverage or require part of the portfolio to be in cash we would simply change the equations in 3.3.-3.5.

The algorithm's complexity is $T(n) = O(log_2 n)$. For more detailed description of algorithm I refer the reader to the original paper López de Prado (2016).

4.2 Comparison

Risk-based strategies enjoy ever-growing interest from academics as well as from practitioners. Research done in the area is generally going in two distinct, but very often interconnected and together presented, directions. The first approach is focused on the empirical examination of historical performance

[5] Order is preserved.

of different risk-based strategies, i.e. comparison of various portfolio selection methods according to the number of performance evaluation measures. The second approach consists of analytical exploration like theoretical properties of individual strategies or dynamics of weights turnover.

In the empirical category, to which this thesis belongs as well, most of the studies are done in multi asset class setting, making the comparison to my work of limited relevance, given the staggering difference of underlying universe on which portfolio selection is performed. The relevance of such studies for our work is with respect to the verification of usefulness of underlying portfolio selection methods in a general way. Choueifaty *et al.* (2013) extend previous work of Choueifaty & Coignard (2008) concerning the MDP strategy both from theoretical as well as empirical perspective. Using index MSCI World as the reference universe, they show consistent outperformance of the MDP strategy over number of strategies, including those examined in this thesis, such as EW, ERC and MV. Linzmeier (2011) obtain consistent results with a different dataset, showing that MDP and surprisingly MV strategies achieve the best Sharpe ratios. Unfortunately for the MDP strategy, Chow et al. (2011), Leote et al. (2012), Chaves et al. (2011) and Lohre *et al.* (2012) obtain results suggesting that the "non-optimized" allocations (EW, RP) are not dominated by the "optimized" ones (MV and MD). All in all, with the exception of the fact that risk-based approaches seem to consistently dominate market-capitalization portfolios, empirical results can be deemed contradictory and seem highly contingent on the universe and the period of study, which perfectly illustrate the limitations of such empirical back-tests.

In that perspective, by being analytical, the second strand of the literature offers a more general understanding of the individual characteristics of popular risk-based portfolios and of their differences. Since it is clearly beyond the scope of this work, I present them only for a reference purposes , with the exception of few, which insights are used elsewhere in the text, for interested reader. In general, these studies provide either exact or approximate analytical solutions for specific cases or compute the sensitivities of the portfolios to number of characteristics. They include Maillard *et al.* (2010), Kaya *et al.* (2012), Kaya & Lee (2012), Clarke *et al.* (2011) and Scherer (2011).

Chapter 5

Performance Evaluation

Performance evaluation criteria usually used in the literature are based on the investor's preferences and cover the main attributes such as financial efficiency, level of diversification and asset allocation stability. To match the rebalancing frequency considered in this study, I use annualized version of measures whenever possible.

As a measure of attractiveness based on profitability of strategy I, simply use arithmetic mean.

In the case of risk, there is a number of options to consider. I work with standard deviation of returns, even though it means assuming that investors treat returns that are greater than mean in the same way as those that are lower. Higher moments of returns distribution are also considered by reporting skewness and kurtosis.

In order to put return and risk into perspective and to be able measure risk-adjusted performance, reward-to-variability ratios are considered. Based on the work of Sortino *et al.* (1999) and Nawrocki (1999), the most prominent measure is the Sharpe ratio defined as the difference between the mean strategy return and the mean risk-free return divided by the standard deviation of the strategy returns. Formally:

$$SR_k = \frac{E[R_k - R_f]}{\sigma_k}$$

To address the weak points of standard deviation as a measure of risk, included in the Sharpe ratio as well, the downside risk measures that focus

only on the variability of returns lower than a minimum acceptable return (MAR), are considered.
Regarding the above, I include the Sortino Ratio, with MAR represented by the risk free return, defined as:

$$\text{Sortino}_k = \frac{E[R_k - MAR]}{\sqrt{\frac{1}{T}\sum_{t=1}^{T} min(R_{t,k} - MAR; 0)^2}}$$

Sortino ratio does not allow for a different emphasis on favorable events and unfavorable ones, e.g. outperforming or underperforming the benchmark. In order to accommodate such concerns, I use the upside potential ratio (UPR) by Sortino *et al.* (1999) defined as :

$$UPR = \frac{\frac{1}{T}\sum_{t=1}^{T} max(R_{t,k} - MAR; 0)}{\sqrt{\frac{1}{T}\sum_{t=1}^{T} min(R_{t,k} - MAR; 0)^2}}$$

and Ω-ratio by Cascon *et al.* (2002) defined as:

$$\Omega = \frac{\frac{1}{T}\sum_{t=1}^{T} max(R_{t,k} - MAR; 0)}{\frac{1}{T}\sum_{t=1}^{T} |min(R_{t,k} - MAR; 0)^2|}$$

These ratios put different levels of importance towards out- and underformance of MAR.

Even though up-to now introduced performance measures cover wide range of interesting aspect, one important aspect is not captured explicitly, the tail risk. Therefore, I report the Value at Risk, Expected Shortfall and Maximum Drawdown.

The Value at Risk (Var) risk metric summarizes the distribution of possible losses by a quantile, a point with a specified probability of greater losses. It is the maximum loss which will not be exceeded with a probability/confidence level $\alpha \in (0, 1)$. Following Follmer & Schied (2004), let X be value the of the portfolio, then VaR is formally defined as:

$$\text{VaR}_\alpha(X) = inf\{x \in \mathcal{R} : P(X + x < 0) \leq 1 - \alpha\}.$$

However, the VaR is not a coherent risk measure since it fails to hold the subadditivity axiom of coherence[1] if the returns are not normally distributed. Because of these features of VaR, I also report the Conditional Value at Risk.

The Conditional Value at Risk (CVAR) CVaR is also called Expected Shortfall and it is a coherent risk measure. Again following, Follmer & Schied (2004), it is defined as:

$$\text{CVaR}_\alpha = \frac{1}{\alpha}\int VaR_\gamma(X)d\gamma$$

By definition, the CVaR is always greater than the VaR for the same confidence level and it increases with decreasing α.

A drawdown is the drop in the portfolio value comparing to the maximum achieved in the past. Following Chekhlov *et al.* (2000), assume $R_p(w_1, ..., w_n, t)$ is the cumulative portfolio return over the preceding portfolio holding time, then the drawdown function is defined as:

$$D(\boldsymbol{w}, t) = \max_{0\le\tau\le t}\{R_p(\boldsymbol{w}, \tau)\} - R_p(\boldsymbol{w}, t)$$

where $\max\limits_{0\le\tau\le t}\{R_p(\boldsymbol{w}, \tau)\}$ is the maximum of the cumulative portfolio return over the history preceding time t. The maximum drawdown is then defined as:

$$\text{MDD}(\boldsymbol{w}) = \max_{0\le\tau\le T}\{D(\boldsymbol{w}, t)\}$$

After listing the tools used in measuring financial efficiency and tail risk, another important attribute of strategy is introduced. The level of diversification or said differently, the lack of concentration.

As the concept of diversification was already introduced in the Section 4, I simply present the formal definitions of two measures used. Both of considered measures could be used either with portfolio weights or total risk contributions. I choose the effective number of bets (ENB), based on the Shannon entropy, to go with the total risk contributions in form of principal portfolios as described in Section 4.1.4 and Gini coefficient (G) to go with weights.

The concept of effective number of bets is already defined in Section 4.1.4,

[1]If $X, Y \in \mathcal{L}$ then $\rho(X+Y) \le \rho(X) + \rho(Y)$

therefore I just recommend reader to that part.

Gini coefficient is computed according to Chaves *et al.* (2011). First, portfolio weights are sorted in ascending order and only then is Gini coefficient calculated:

$$G_k(w) = \frac{2}{N}\sum_{i=1}^{N} i(w_{i,k}\overline{w}_k)$$

The Gini coefficient has range between zero and one. It has an opposite interpretation compared to the SE, the lower it is, the more diversified in terms of weights the portfolio is. Specifically, zero means perfect equality of weights, as is the case , while one means perfect inequality that is maximum weight concentration, i.e. being invested in the only one asset.

Further, in order to be able to tell if the average returns and volatilities of considered strategies differ significantly, I employ Wilcoxon signed-rank test and Levene's test. These tests are alternatives for the normality requiring Student's t-test and Bartlett's test. For specific details, I refer reader to Wilcoxon (1945) and Levene *et al.* (1960), respectively.

Even though, performance of individual strategies is compared between each other, the naive equal-weighting portfolio serves as a main benchmark, following DeMiguel *et al.* (2009).

Chapter 6

Empirical results

Factor portfolios under investigation are formed using methods presented in Section 4. Description of datasets on which backtests are conducted as well as details about backtesting procedure is provided below in Section 6.1. Performance of individual portfolio selection methods is evaluated according to the number of measures introduced in Section 5 with the main empirical results reported in Section 6.2.

Analysis is performed using Python programming language. Optimization tasks are implemented using Scipy library.

6.1 Data

Selection of asset universe is one of the most important and necessary decision to make in an asset allocation study. Mean-variance framework as well as risk-based approaches are generally applied in allocation between asset classes, however there is no reason for not using it in the equity-only domain. Actually, as argued in the introduction, given the question of efficiency in capturing risk factor premia, it might be a wanted endeavour. Accordingly, this thesis is focused on the U.S. stock market, the most developed stock market in the world.

I use standard datasets used in asset pricing literature:

- Center for Research in Security Prices (CRSP); from where I obtain prices, Standard Industrial Classification (SIC) code and number of outstanding shares.

- COMPUSTAT industrial files; from where I obtain annual fundamental data.

I integrate data across COMPUSTAT and CRSP through PERMNO/LPERMNO link. Time period under investigation is factor specific and ranges from June 1965 (the earliest) to June 2015 (the latest). Starting date is restricted by the availability of necessary accounting data for factor construction. For years before time period under consideration COMPUSTAT data have a selection bias. My sample consists of all firms with common stock on the NYSE, AMEX or NASDAQ that have a month-end market value on CRSP. I exclude firms with negative book equity as well as financial firms, i.e. those with SIC code between 6000 and 6799. Fama & French (1992) argue for exclusion because of difference in use of leverage between financial and non-financial firms. However, the main reason is to stay in line with common factor construction procedures and be sure that skewed financials fundamental data are not the main reason for my results. Based on the argumentation of Hou *et al.* (2017), already described in literature review, I use NYSE size cutoffs when defining and subsequently eliminating microcaps. This procedure is important in my research design also because of another reason. Since advanced portfolio selection methods often allow relatively high weight concentration in the portfolio, I want to eliminate a chance that portfolios would be highly concentrated on small stocks without necessary capacity for investment. Finally I require all stocks included in the portfolios to have 3 years of historical returns for estimation purposes. In order to have equal setting for each portfolio selection method, this requirement holds for equal-weighting as well as value-weighting, even though no estimation is necessary in those cases.

Transforming a theoretical, non-investable risk factor into an investable one is not straightforward. The implementation decisions, like short sale constraint, rebalancing frequency, asset universe and weighting scheme, must be made.

In the academic literature, risk factors are designed as long/short portfolios. As mentioned in Section 3, this construction technique dates back to Fama & MacBeth (1973). In contrast, the commercial investment vehicles designed to offer exposure to risk factors are mainly long-only. It is still unanswered question which approach is better, with relevant studies being, e.g. Bender *et al.* (2010), Ilmanen & Kizer (2012) and Huij *et al.* (2014).

Regarding setting for my empirical tests, there are two reasons I operate under long-only or no short-sale condition. First, Huij *et al.* (2014), as the most recent study regarding the topic of interest, compare long-only versus long-short approaches to factor investing in an empirical study. Main findings are that although the long-short approach is relatively better in theory, the long-only approach seems to be a better alternative in most of the scenarios that account for practical issues such as benchmark restrictions, implementation costs and factor decay. As mentioned above, practitioners seem to agree. Second reason is that it is not obvious how to translate most of the advanced portfolio selection methods introduced in Section 4 into a long-short framework given the structure of factor portfolio being split into two extreme parts of the characteristic distribution.

The choice of rebalancing frequency is another important topic, not only when it comes to the design of my empirical research, but also in case of practical implementation of factor investing. Factor portfolios are usually rebalanced on monthly, quarterly, or yearly basis. Reason for choosing these frequencies is given by the frequency of new information being available (quarterly and annual financial statements) and availability of prices (monthly CRSP). In general, increasing the frequency may have a considerable impact on the portfolio turnover and also on trading costs.

Since most of the portfolio selection methods employed in my thesis solve non-trivial optimization problem, testing monthly rebalanced strategies is computationally very expensive, nevertheless absolutely feasible. Given the time constraint, I focus on the annual rebalancing, while monthly rebalancing is used just once for testing of robustness of main results.

For annual rebalancing, I rank all stocks in the universe[1], at the end of June of each year t according to some factor, e.g. MISP, computed using fundamental data from the fiscal year ending in calendar year $t-1$. A general factor portfolio consists of stocks that are above (below) 80% (20%) factor rank quantiles.

As a workhorse factor I choose mispricing factor (MISP) from Stambaugh & Yuan (2016). It is a composite factor based on 11 prominent anomalies, net stock issues (NSI), composite equity issues (CEI), accruals (ACCR), net operating assets (NOA), asset growth (AG), investment-to-assets (IA), distress (DSTR), O-score (O), momentum (MOM), gross profitability (GP), and return on assets (ROA). In the first step all stocks in my universe are ranked according to each anomaly, e.g. firm with the lowest net operating assets

[1]Does not contain stocks eliminated in the previous steps, e.g. microcaps or financials.

is given the rank one. Next, average rank across all characteristics is computed. Apart from using the MISP factor, I use individual factors as well. Further, I study factors that cannot be explained by currently most powerful asset pricing model of Stambaugh & Yuan (2016) and work on the annually rebalanced basis. These are, in no particular order, industry-adjusted real estate ratio (RER), earnings predictability (EPRD), cash-based operating profits-to-lagged assets (CLA), cash-based operating profitability (COP), changes in net financial assets (dFIN), net payout yield (NOP), inventory changes (IVC) and R&D Expense-to-market (RDM). Definitions, corresponding to the construction manuals, taken from Stambaugh *et al.* (2012) and Hou *et al.* (2017), are presented in the Appendix.

After having a general portfolio, i.e. constituents of the portfolio, I form 8 portfolios corresponding to 8 portfolio selection methods presented in Section 4, i.e. equal-weighting (EW), value-weighting portfolio (VW), minimum variance portfolio (MV), naive risk parity portfolio (NRP), optimal risk parity portfolio (ERC), the most diversified portfolio (MDP), hierarchical risk parity portfolio (HRP) and diversified risk parity portfolio (DRP).

Garlappi *et al.* (2007) argue that equal-weighting should serve as a benchmark when comparing different portfolio selection methods. Equal-weighting, along with value-weighting, is also used in anomalies literature as a simple portfolio selection method related to the Fama & MacBeth (1973) cross-sectional regression. In light of this, equal-weighting is selected as a benchmark for the rest of the strategies.

Subsequently, for each of the eight portfolios I calculate returns from July of year t to June of $t+1$ and save weights of individual stocks in formed portfolios as well as returns of individual stocks in the portfolios for performance evaluation purposes.

6.2 Results

This section contains the main results of my empirical investigation. Results for the mispricing factor, it's constituents and other significant anomalies with respect to the mispricing model of Stambaugh & Yuan (2016) are reported in Sections 6.2.1, 6.2.2 and 6.2.3, respectively. The robustness of results is considered in Section 6.2.4 and final discussion of results is provided in Section 6.3.

6.2.1 Mispricing Factor

Table 6.1 shows the basic statistics of the returns of mispricing factor (MISP) porfolio formed based on eight portfolio selection methods, i.e. strategies.

MISP	EW	VW	MV	NRP	ERC	MDP	HRP	DRP
Mean p.a. (%)	17.80	17.10	14.90	16.45	18.38	22.13	16.31	18.59
Volatity p.a. (%)	15.56	16.00	12.49	14.36	15.41	14.13	13.02	22.95
Skewness	-0.53	0.06	-0.11	-0.21	-0.22	-0.26	-0.59	0.13
Kurtosis	11.60	10.14	7.41	7.97	7.19	13.40	14.09	7.19

Table 6.1: Return characteristics for different portfolio selection methods applied on mispricing factor portfolio. All values displayed in percent, except the skewness and kurtosis.

In following discussions, unless stated otherwise, *significant* means at 99% significance level given the the Wicoxon signed-rank test for comparison of means and Levene's test for comparison of variances between two return series.

Naive equal-weighting strategy (EW), serving as benchmark returned 17.8% annualy on average. According to the well established stylized fact, value-weighting (VW) performed slightly worse, however, surprisingly with moderately higher volatility. Mean is significantly lower only at 95% level, while volatility is significantly higher.

As expected, minimum variance strategy (MV) achieves the lowest volatility not only in the in-sample estimation but also out-of-sample. This comes at price of the lowest return as well.

Naive risk parity strategy (NRP) has significantly lower volatility than EW as well as optimal risk parity (ERC). However, as in the case of the MV strategy, this is accompanied by significantly lower average return.

The ERC strategy significantly outperforms the benchmark with comparable volatility. In a sense, this contrasts with expectations based on findings of Maillard *et al.* (2010), who find that ERC strategy lies between the MV and EW strategy. While this is indeed the case for volatility, even though more tilted towards EW side, it is not the case for returns.

The Hierarchical Risk Parity strategy (HRP) has the second lowest volatility among all strategies but also the second lowest return. Yet again, return goes hand in hand with volatility as in the previous two cases. Another notable aspect is the highest kurtosis as well as the lowest skewness.

Diversified Risk Parity strategy (DRP) is the most volatile strategy with significantly higher volatility than even the second most volatile one, i.e. EW. This result is extremely surprising, given the prominent role of DRP among diversification maximizing strategies. Its abnormal volatility dominates even the fact that it has the second highest return.

The strategy with the highest return is the most diversified portfolio strategy (MDP), returning 22.13% p.a. on average, i.e. 3.33% p.a. more than benchmark. This difference is highly significant both statistically and economically. What is, however, more impressive is that this excessive return doesn't come at the cost of higher volatility, which is significantly lower than benchmark and is the third lowest among all strategies.

Concerning skewness, only VW and DRP are positively skewed but relatively close to 0. Strategies with the most negative skewness are EW and HRP, indicating the longer or fatter tail on the left side of the probability density function compared to the right side. All strategies are very leptokurtic, with MDP's and HRP's kurtosis higher than benchmark.

According to return and volatility, strategies are ranked in the following order:

$$\mu_{MV} < \mu_{HRP} < \mu_{NRP} < \mu_{VW} < \mu_{EW} < \mu_{ERC} < \mu_{DRP} < \mu_{MDP}$$

$$\sigma_{MV} < \sigma_{HRP} < \sigma_{MDP} < \sigma_{NRP} < \sigma_{ERC} < \sigma_{EW} < \sigma_{VW} < \sigma_{DRP}$$

In Figure 6.1, the hypothetic wealth evolution under different portfolio selection strategies, assuming 1$ initial capital investment, is depicted.

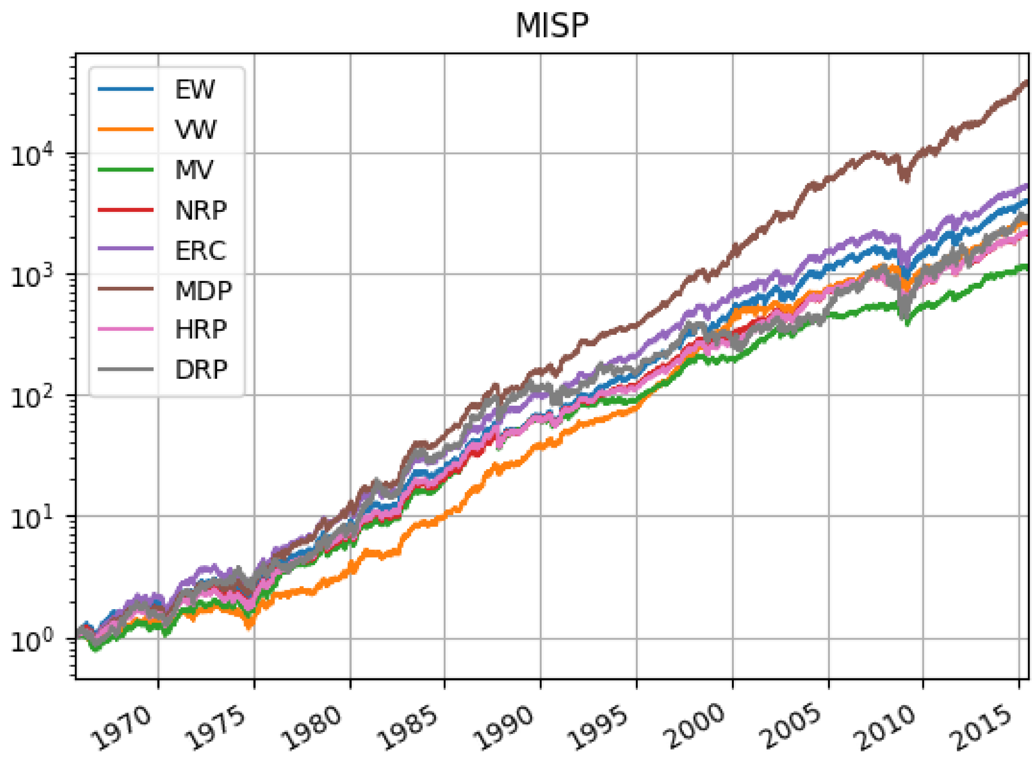

Figure 6.1: Cumulative performance of eight performance selection methods applied on mispricing factor portfolio from 1965 to 2015.

Even though distributional characteristics of returns generated by different portfolio selection strategies allow their comparison, this is certainly only one evaluation dimension. It is more insightful to look at risk-adjusted performance in the form of reward-to-variability measures, presented in Table 6.2.

MISP	EW	VW	MV	NRP	ERC	MDP	HRP	DRP
Sharpe	1.13	1.04	1.35	1.11	1.18	1.61	1.21	0.81
Sortino	1.46	1.40	1.83	1.46	1.54	2.09	1.58	1.08
UPR	2.81	2.97	2.84	2.84	2.86	2.71	2.72	3.00
Omega	3.60	3.28	4.44	3.56	3.77	5.48	4.01	2.52

Table 6.2: Risk-adjusted performance for different portfolio selection methods applied on mispricing factor portfolio. All measures computed on the annual basis. In Sortino ratio, upside potential ratio (UPR) and Omega ratio, the threshold level or minimum acceptable return is represented through the risk free return.

As was expected based on the Table 6.1, strategy with the highest Sharpe ratio is MDP and strategy with the lowest one is DRP. Interesting comparison is between rank of strategies based on Sharpe ratio versus rank based on Sortino ratio, which is the same. This means that no strategy was penalized by the Sharpe ratio for the lower partial moments relative to the other strategy.

Above presented risk-adjusted statistics look at risk based on the standard deviation or downside standard deviation, i.e. capturing variability around a central tendency. Table 6.3 offers an explicit insight into the tail risk inherent to each portfolio selection method.

MISP	EW	VW	MV	NRP	ERC	MDP	HRP	DRP
VaR 5%	1.49	1.50	1.00	1.37	1.49	1.29	1.23	2.04
CVaR 5%	2.27	2.23	1.74	2.10	2.24	2.06	1.92	3.21
MDD	-48.86	-41.52	-34.85	-48.45	-49.09	-42.74	-46.26	-63.83

Table 6.3: Tail risk measures for different portfolio selection methods applied on mispricing factor portfolio. VaR 5% and CVar 5% calculated on the daily basis. All values displayed in percent.

Maximum drawdown (MDD) is comparable across the strategies with two notable exceptions. The MV strategy, as expected based on the literature review, has substantially smaller maximum drawdown than other strategies. On the other hand, quite surprisingly, the DRP has the highest drawdown of -63.83% implying very little tail-risk protection, even compared to the benchmark.

As in the case of MDD, Value at risk (VaR) and Expected shortfall (CVaR) with 5% confidence level are comparable across the strategies with the exception of the MV and DRP strategies. The MV strategy, as expected based on the literature review, has substantially smaller VaR and CVaR.

For MISP factor, it seems that more advanced portfolio selection methods, with the exception of MDP, offer very little protection against the tail risk. In terms of tail-risk, the MDP strategy is beaten only by MV and VW, while in case of VW only slightly.

The last dimension of performance evaluation is the level of diversification achieved, both in terms of portfolio weights as well as sources of risk represented as principal components or principal portfolios.

MISP	EW	VW	MV	NRP	ERC	MDP	HRP	DRP
ENB	1.17	1.32	2.46	1.68	1.75	2.82	2.03	4.29
Gini(w)	0.00	0.54	0.62	0.49	0.39	0.61	0.69	0.51

Table 6.4: Average diversification measures across time for different portfolio selection methods applied on mispricing factor portfolio.

The level of concentration of portfolio weights, as given by the Gini coefficient, is 0 for EW from the definition. The highest weight concentration occurs in case of HRP, followed by MV and MDP. On the other hand, the level of diversification, when measured as effective number of bets (ENB) based on the actual principal components (sources of risk) , looks different. By great margin, the highest ENB is achieved by DRP. This is not entirely surprising, since it is being maximized in-sample, however the persistence of this attribute towards out-of-sample is impressive. Documented relationship between concentration in terms of weights and in sources of risk can be observed as well in Table 6.4. Specifically, strategies with higher weights concentration have less risk concentration. To put it differently, in order to achieve higher risk diversification it is necessary to allow higher weight concentration.

Even though mispricing factor is a composite factor, it is still only one factor and therefore stocks included in the portfolio according to it create just one universe (in time). In order to see if the results from the previous section hold more generally in factor investing, it is necessary to apply them in a number of factor portfolios and I proceed to do so in Sections 6.2.2 and 6.2.3.

6.2.2 Constituents of Mispricing Factor

Natural candidate factors for further tests are the constituents of the mispricing factor (MISP), since they are based on the same line of literature as mispring factor and I have already implemented them during the construction of mispricing factor. I disregard ROA factor, for which monthly rebalancing is necessary, and work with the remaining 10 factors: net stock issues (NSI), composite equity issues (CEI), accruals (ACCR), net operating assets (NOA), asset growth (AG), investment- to-assets (IA), distress (DIST), O-score (O), momentum (MOM) and gross profitability (GP). Definitions of all factors can be found in Appendix A. Results of backtests for individual factors are shown in a compact way in Appendix. All results in this section are averages of results for those underlying individual factors. Same performance evaluation criteria as in the case of single mispricing factor are used, starting with basic statistics of the returns.

Constituents	EW	VW	MV	NRP	ERC	MDP	HRP	DRP
Mean p.a. (%)	24.35	21.40	17.69	21.66	24.40	30.80	21.89	23.00
Stdev p.a. (%)	20.24	19.92	14.16	18.03	19.83	18.91	16.04	29.08
Skewness	-0.30	0.04	-0.26	-0.33	-0.30	-0.27	-0.48	0.06
Kurtosis	6.49	7.05	5.93	7.09	6.39	10.79	10.66	7.92

Table 6.5: Average results across 10 constituents of mispricing factor for different portfolio selection methods. All values displayed in percent, except the skewness and kurtosis.

Stambaugh & Yuan (2016) shows that, with equal weighting as a portfolio selection method, the spread between the alphas for portfolios of stocks in the extreme deciles of the average ranking across the 11 anomalies is nearly twice the average across those anomalies of the spread between the top- and bottom-decile alphas of portfolios formed using an individual anomalies. This observation is in contrast with my results, average mean return across constituents is 6.55% higher for EW and 8.67% for MDP strategy. The same effect can be seen in case of volatility, which is higher 4.56% for EW and 3.98% for MDP strategy. The main reason is that portfolios formed in my thesis are long-only while portfolios generating returns in mentioned study are long-short. There are also another minor differences. I take averages across 10 constituents while the original paper averages across 11 and I also require stocks included to have 3 year history of returns for the covariance matrix estimation purposes.

Apart from the difference in absolute returns and volatility, there is little change in terms of relative order.

$$\mu_{MV} < \mu_{VW} < \mu_{NRP} < \mu_{HRP} < \mu_{DRP} < \mu_{EW} < \mu_{ERC} < \mu_{MDP}$$

$$\sigma_{MV} < \sigma_{HRP} < \sigma_{NRP} < \sigma_{MDP} < \sigma_{ERC} < \sigma_{VW} < \sigma_{EW} < \sigma_{DRP}$$

The most notable change is that volatility of EW is higher than VW, which is expected given the stylized facts from asset pricing literature. Other than that, the MDP strategy is the one with the highest return, while the MV strategy has the lowest volatility as well as returns.

Given the fact that both average mean returns and average volatility are higher, it is useful to look again at the risk-adjusted measures when looking at constituents.

Constituents	EW	VW	MV	NRP	ERC	MDP	HRP	DRP
Sharpe	1.24	1.02	1.20	1.15	1.27	1.73	1.35	0.79
Sortino	1.59	1.39	1.58	1.52	1.58	2.21	1.78	1.04
UPR	2.76	2.91	2.86	2.80	2.80	2.63	2.69	2.85
Omega	4.09	3.39	3.91	3.94	4.13	6.21	4.63	2.62

Table 6.6: Average risk-adjusted performance across 10 constituents for different portfolio selection methods. All measures computed on the annual basis. In Sortino ratio, upside potential ratio (UPR) and Omega ratio, the threshold level or minimum acceptable return is represented through the risk free return.

Results in Table 6.6 are very similar to the results from Table 6.2 that are describing mispricing factor. Therefore, even though mean returns are higher on average for constituents than single mispricing factor, it is also accompanied by higher volatility, making almost no difference on risk adjusted basis. This means that the conversion rate between risk, defined as standard or downside deviation, and return is basically the same as for the mispricing factor.

Constituents	EW	VW	MV	NRP	ERC	MDP	HRP	DRP
VaR 5%	1.74	1.74	1.21	1.59	1.76	1.55	1.39	2.46
CVaR 5%	2.74	2.70	1.99	2.53	2.72	2.55	2.23	3.97
MDD	-53.61	-49.14	-41.49	-52.25	-51.96	-49.07	-49.49	-69.67

Table 6.7: Average tail risk measures across 10 constituents of mispricing factor for different portfolio selection methods. VaR 5% and CVar 5% calculated on the daily basis. All values displayed in percent.

Looking at Table 6.7, the tail risk characteristics seem slightly worse than for the mispricing factor across the board.

The MV strategy has the lowest maximum drawdown as well as 5% confidence level VaR and CVaR. A very concerning fact for the DRP strategy, especially when taken together with previous results, is that extreme drawdown of −63.83% and CVaR of 3.21% were not a simple outliars specific to the mispricing factor, but look like more pervasive attribute of the DRP strategy reaching on average almost −70% and 3.97% .

	EW	VW	MV	NRP	ERC	MDP	HRP	DRP
ENB	1.12	1.31	2.43	1.56	1.67	2.6	2.02	4.00
Gini(w)	0.00	0.58	0.64	0.56	0.41	0.69	0.70	0.61

Table 6.8: Average diversification measures across time for 10 constituents of mispricing factor.

Table 6.8 shows measures of diversification for weights as well as risk sources and paints the similar picture as Table 6.4 that describes the same attributes for mispricing factor. DRP is again the strategy with the highest number of effective number of bets, thus being the most diversified with respect to the risk sources. This is achieved through relatively high weight concentration. Results for the DRP strategy are thoroughly surprising. On one hand, strategy seems to achieve the highest level of diversification among all strategies and yet clearly is exposed to the high amount of statistical risk, either in terms of volatility or tail risk.

In order to verify my up-to-now results as well as obtain truly broad insight into the portfolio selection in factor investing, I further look at the next list of interesting factors. Inspired by the results of Hou *et al.* (2017), I consider factors which can be considered anomalies with respect to the most powerful asset pricing models today, q-factor model from Hou *et al.* (2015) as well as the mispricing model from Stambaugh & Yuan (2016).

6.2.3 Other Significant Anomalies

Annually rebalanced factors that cannot be explained by the mispricing model or q-factor model are considered in this section, including industry-adjusted real estate ratio (Rer), earnings predictability (Eprd), cash-based operating profits to lagged assets (Cla), cash-based operating profitability (Cop), changes in net financial assets (dFin), net payout yield (Nop), inventory changes (Ivc), and R&D expenses to market (Rdm). They are selected based on the results presented in Hou *et al.* (2017) and Stambaugh & Yuan (2016). As in the previous section, definitions of all factors can be found in Appendix A and the results of backtests for individual factors are shown in a compact way in Appendix B.2. All results in Tables 6.9, 6.10, 6.11 and 6.12 are averages of results for those underlying individual factors. Same performance evaluation criteria are used as in the case of single mispricing factor, starting again with the basic statistics of the returns in Table 6.9.

Other Factors	EW	VW	MV	NRP	ERC	MDP	HRP	DRP
Mean	21.78	20.90	16.26	19.39	21.70	27.99	19.37	21.89
Stdev	18.91	18.90	13.17	17.24	18.85	17.83	15.32	28.60
Skewness	-0.23	0.01	-0.19	-0.29	-0.25	-0.26	-0.47	0.21
Kurtosis	6.09	6.98	7.75	7.56	6.31	10.63	11.86	7.33

Table 6.9: Average results across 8 other significant factors for different portfolio selection methods. All values displayed in percent, except the skewness and kurtosis.

Both return-wise as well as volatility-wise results for the other 8 factors fall between the original mispricing factor (lower bound) and the averages of its constituents. EW outperforms VW in terms of returns as expected, while again having comparable volatility.

The best strategy according to the raw returns remains to be MDP with reasonable amount of volatility. The ERC strategy performs similarly, both in terms of returns as well in terms of volatility as benchmark, confirming doubt about its place between EW and MV strategies. HRP again ranks second in terms of volatility.

The MV strategy is again the one with the lowest volatility as well as the lowest average returns.

The observed skewness and kurtosis are comparable with skewness and kurtosis in case of MISP components. Remarkably, the relative order based on return or volatility is unchanged as shown below:

$$\mu_{MV} < \mu_{VW} < \mu_{NRP} < \mu_{HRP} < \mu_{DRP} < \mu_{EW} < \mu_{ERC} < \mu_{MDP}$$

$$\sigma_{MV} < \sigma_{HRP} < \sigma_{NRP} < \sigma_{MDP} < \sigma_{ERC} < \sigma_{VW} < \sigma_{EW} < \sigma_{DRP}$$

Since both volatilities and returns moved in tandem compared to the MISP constituents, it is reasonable to expect similar risk-adjusted results.

Other Factors	EW	VW	MV	NRP	ERC	MDP	HRP	DRP
Sharpe	1.18	1.12	1.19	1.12	1.18	1.70	1.26	0.79
Sortino	1.51	1.48	1.59	1.46	1.51	2.14	1.63	1.05
UPR	2.90	2.97	2.90	2.87	2.89	2.73	2.77	2.96
Omega	3.72	3.52	3.81	3.56	3.72	5.76	4.13	2.52

Table 6.10: Average risk-adjusted performance across 8 other significant factors for different portfolio selection methods. All measures computed on the annual basis. In Sortino ratio, upside potential ratio (UPR) and Omega ratio, the threshold level or minimum acceptable return is represented through the risk free return.

This is indeed the case, not only with regard to the relative order but to the magnitude of the ratios as well.

Other Factors	EW	VW	MV	NRP	ERC	MDP	HRP	DRP
VaR 5%	1.82	1.76	1.20	1.65	1.81	1.60	1.42	2.55
CVaR 5%	2.75	2.65	1.87	2.51	2.74	2.58	2.23	4.05
MDD	-50.66	-45.87	-47.21	-49.24	-50.39	-45.07	-46.65	-71.55

Table 6.11: Average tail risk measures across 8 other significant factors for different portfolio selection methods. VaR 5% and CVar 5% calculated on the daily basis. All values displayed in percent.

Looking at Table 6.11, the tail risk characteristics fall in the middle between the mispricing factor and characteristics of its constituents.

Interestingly, the MDP strategy has the lowest maximum drawdown as opposed to the MV strategy. 5% confidence level VaR and CVaR is lowest for the MV strategy as was in the previous cases. Regarding the DRP strategy, as previous results suggested, extreme exposure to the tail risk can be seen.

Other Factors	EW	VW	MV	NRP	ERC	MDP	HRP	DRP
ENB	1.23	1.27	2.30	1.60	1.69	2.91	1.99	4.09
Gini(w)	0.00	0.60	0.66	0.53	0.44	0.63	0.73	0.58

Table 6.12: Average diversification measures across time for 8 other significant factors.

Results regarding diversification, both in terms of weights as well as in terms of risk sources, are very similar to the previous cases.

6.2.4 Other Robustness Tests

In the preceding Sections 6.2.1, 6.2.2 and 6.2.3, results for different portfolio selection methods, across number of significant factors are presented. Similarity of results across the factors suggests certain robustness. However, over the whole research process I made the number of methodological choices and thus thorough testing for robustness is required. This is done along a variety of dimensions. Results in this section are aggregated over every factor analyzed previously, i.e. MISP, its constituents and other significant factors, totaling 19 factors. Results are described without presentation of all results and tables, given the space constraint.

The cutoff percentile for each factor is either 20% or 80% depending whether stocks to long are in the bottom two deciles or in the top two deciles. This is a standard procedure in anomalies literature, however sometimes 30% and 70% is used. Splitting universe in deciles is another common procedure. I replicate the whole analysis for factor portfolios constructed using both bottom and top 30% and 10% of stocks from the universe.

When including 30% of the universe in the portfolio selection process, during certain years when number of stocks in universe and thus in the portfolio is particularly high, I encounter the problem with correlation matrix not being positive definite. When sample size is small, a sample covariance or correlation matrix may not be positive definite due to mere sampling fluctuation. This condition is however necessary for every portfolio selection method relying on optimization. In order to solve this problem, when working with extreme 30%, I use five year long estimation window instead of three.

Results are as expected, given the stylized facts from the literature. Absolute returns, volatility and tail risk are on average lower when using 30% cutoff and on average higher when using 10% cutoff. Consequently, risk-adjusted performance matrices are very similar.

In terms of weights concentration, I obtain comparable results across all three cutoff methods. On average, the effective number of bets is marginally higher when using more broad universe, i.e. 30% and marginally lower when using narrower universe, i.e. 30%. These differences look economically insignificant.

To be sure that results for 30% cutoff are not primarily driven by longer estimation window used, I use 5 year estimation window for 10% and 20% cut-offs as well. In general, I obtain slightly better results on a risk-adjusted basis and slightly higher effective number of bets on average.

I also split every backtest into two sub-periods, the first half and the second half of the examined time period, and evaluate performance on those subperiods. With the exception of the MDP strategy, mean returns are lower in the second subperiod with comparable volatility to the first subperiod. Consequently, results are weaker on the risk-adjusted basis as well. The MDP strategy achieves similar results in both subperiods. Concentration in terms of weights is marginally higher in the first subperiod, this difference seems economically meaningless. Effective number of bets are comparable across subperiods for all strategies.

Further, I repeat the default testing procedure for all factors examined in this thesis and obtain very similar risk-adjusted results, tail risk measures as well as diversification measures.

In Section 6.3, summarizing discussion as well as suggestions for further research are presented.

6.3 Discussion

Eight portfolio selection methods are used in the portfolio formation based on mispricing factor, its constituents and other factors that are significant anomalies with respect to the mispricing model of Stambaugh & Yuan (2016) as well as q-factor model of Hou *et al.* (2015). In aggregate, the main results are based on 152 strategy backtests. Robustness tests consist of additional 616 backtests.

When looking at all portfolio selection methods together, obtained results differ between mispricing factor, its constituents and other factors in terms of returns and volatilities. On average the highest return and volatility is achieved by constituents of the mispricing factor and the lowest return as well as volatility is achieved in case of single mispricing factor. Since risk and reward move more or less in synchronization, risk-adjusted measures are similar across the groups.

Tail risk measures paint the similar picture, where mispricing factor has on average the smallest out-of-sample tail risk and its constituents the highest. Nevertheless, there is a little bit more heterogeneity in case of other significant factors.

In contrast to different volatilities and tail risk measures, weight concentrations as well as diversification regarding statistical sources of risk are very homogenous across the groups (for a given strategy), suggesting that differences between individual groups of factors in terms of risk are not caused by different level of diversification but because of different level of riskiness of each underlying factor group. The same conclusion can be reached when individual factors are considered instead of groups of factors.

More interesting and related to the questions directly asked in this thesis is looking at differences between the portfolio selection methods. Here, relative results between groups of factors as well as individual factors are surprisingly stable.

First of all, in terms of returns and volatility, strategies can be ranked in the following order:

$$\mu_{MV} < \mu_{VW} < \mu_{NRP} < \mu_{HRP} < \mu_{DRP} < \mu_{EW} \approx \mu_{ERC} < \mu_{MDP}$$

$$\sigma_{MV} < \sigma_{HRP} < \sigma_{NRP} < \sigma_{MDP} < \sigma_{ERC} < \sigma_{VW} < \sigma_{EW} < \sigma_{DRP}$$

Consequentially, when looking at Sharpe ratio, not as the best risk-adjusted measure in literate but as a standard, rank of strategies can be described in a following way:

$$\mathrm{SR}_{DRP} < \mathrm{SR}_{VW} < \mathrm{SR}_{NRP} < \mathrm{SR}_{MV} < \mathrm{SR}_{EW} \approx \mathrm{SR}_{ERC} < \mathrm{SR}_{HRP} < \mathrm{SR}_{MDP}$$

Concerning tail risk, there is a small difference in heterogeneity between the constituents of mispricing factor and other significant factor. With the exception of MV and DRP all strategies are similar for group of constituents, while in case of the other significant factors there are small differences even among the non-outlying strategies. Looking at maximum drawdown (MDD) as the representative measure of the tail risk, the riskier half of the strategies can be ranked in the following order:

$$\text{MDD}_{NRP} < \text{MDD}_{ERC} \approx \text{MDD}_{EW} < \text{MDD}_{DRP}$$

while the less riskier half is different for the groups of factors. Noteworthy, the MDP strategy has the smallest and the second smallest maximum drawdown among its competitors.

Given the fact that investigated portfolios do not differ in original asset universe from which they are picked, the difference is of course in the weight of individual assets. It is insightful to look at the comparison of weights concentration, especially since its relative order is very stable across different factors as well as across time.

$$\text{G(w)}_{EW} < \text{G(w)}_{ERC} < \text{G(w)}_{NRP} < \text{G(w)}_{VW} < \text{G(w)}_{DRP} \approx \text{G(w)}_{MDP} < \text{G(w)}_{MV} < \text{G(w)}_{HRP}$$

Concentration of portfolio weights is highest for the HRP strategy, relatively closely followed by MDP and MV. From the definition, the EW strategy has zero concentration based on weights. The lowest, non-zero weight concentration is consistently achieved by the ERC strategy.

As argued throughout the thesis, weight concentration contrast strongly with the diversification achieved with respect to the risk sources, defined through Meucci (2010) framework. According to the effective number of bets (ENB), individual strategies can be ranked in following order:

$$\text{ENB}_{EW} < \text{ENB}_{VW} < \text{ENB}_{NRP} < \text{ENB}_{ERC} < \text{ENB}_{HRP} < \text{ENB}_{MV} < \text{ENB}_{MDP} < \text{ENB}_{DRP}$$

The DRP strategy has the highest number of effective bets and by a wide margin. Strategy with the second highest number of effective bets is the MDP, the best-performing strategy based on the number of measures, closely followed by MV.

The worst diversified across sources of risk are the EW and VW strategies, the very strategies used in the discovery of factor premia.

All these results alone offer a valuable insight for practitioners of factor investing who cannot skip the portfolio construction part of the whole process. Contribution consists not only in showing what works significantly better than standardly used equal-weighting and value-weighting methods, i.e. the MDP strategy, but also in what does not add any value depending on what aspects of performance matter, and in what should factor investors be extremely careful about, i.e. the DRP strategy.

However, as argued in the introduction, there is a deeper question.

Either because of risk or mispricing, original equal-weighted and value-weighted factor portfolios are associated with higher expected returns. Ex-ante, it is not obvious what should happen with the performance of factor portfolios when advanced portfolio selection methods are used instead of original non-optimal methods.

From one perspective, if investors should not be compensated for diversifiable risk and advanced portfolio selection methods indeed provide more diversification than equal-weighting, then one could expect that the expected returns should not change compared to the benchmark of equal-weighting. This is simply because by using equal-weighting unnecessary diversifiable risk is undertaken, thus not deserving compensation in form of higher expected returns.

From another perspective, factor portfolios are constructed from very restricted universe exposed to some fundamental risk (or mispricing). Portfolios constructed using advanced portfolio selection methods, more or less directly based on the idea of diversification across independent risk sources, see Sections 4.1.4, 4.1.2, 4.1.3 and 4.1.5, provide more effective representation of underlying universe from which they are built. Since in case of factor portfolios the underlying restricted universe is already exposed to some fundamental risk (or mispricing), it could be argued that these portfolio selection methods offer less noise-contaminated exposure to underlying risks (mispricing) and therefore shoule be able to capture associated premia in a superior way.

To answer this question, it is important to look at the actual, out-of-sample achieved diversification with respect to the risk sources and compare it with the returns.

Looking at the relative order of all strategies according to the effective number of bets, as well as individual factor results, it is clear that the naive

EW and VW strategies are prime examples of not providing full exposure to all risk sources explaining the return variation.

Out of the strategies based on diversification across risk sources, the ERC and HRP strategies are weaker in terms of achieving the same properties as in-sample than the other two. In terms of returns, ERC has consistently comparable return with the EW strategy and HRP has consistently lower return in terms of returns, but with the lower volatility.

The most diversified out-of-sample are the MDP and DRP strategies, with the highest number of effective bets. Based on the argumentation above, they should be considered effective representations of underlying factors, having exposure to more sources of risk not only from theoretical perspective but also in reality as shown empirically. Surprisingly, their performances differ significantly almost according to every single metric.

The MDP strategy has clearly the highest return with comparable volatility and one of the lowest tail risk among other strategies. These results would suggest that constructing portfolio by maximizing the diversification ratio, i.e. degrees of freedom or independent sources of risk for given portfolio, is a more precise way to fully capture the premia associated with underlying factors.

Despite the benchmark beating results of MDP, results for DRP, the most, head above the rest, out-of-sample diversified strategy, are anything but similar. Its returns are more comparable with EW and advanced ERC and HRP strategies. However, opposed to these strategies, it suffers abnormal levels of volatility as well as tail risk.

Overall out of four diversification maximizing strategies, all achieve higher out-of-sample diversification across sources of risk than equal-weighting and value-weighting. The level of achieved diversification differs across them, where the DRP stands head above the rest, followed by the MDP. The ERC, HRP, DRP strategies have similar average returns to EW. These results would support the explanation of not compensating the unnecessary, undiversifiable risk undertaken by EW. However, the MDP achieves consistently significantly higher returns, undermining the conclusion above and giving evidence more in favor of offering less noise-contaminated exposure to underlying risks.

To summarize, these are conflicting but provide fertile ground for further research. Few suggestions.

First of all, even though the MDP, DRP and ERC strategies are clearly related, empirical results suggest that there are significant differences deserving thorough analytical investigation.

Second, I work with long-only portfolios, but factors are constructed being long-short. Modifying advanced portfolio selection methods into long-short framework, while retaining the desirable properties of their long-only counterparts, in order to construct the factor portfolios more closely to their nature looks like a logical next step.

Third, all advanced methods rely heavily on the estimates of covariance matrix, therefore more robust methods, e.g. shrinkage can bring additional benefits.

Finally, recent advances in the field of spectral portfolio theory, see Chaudhuri & Lo (2016), are potentially worthy of combining with methods used in this thesis.

Conclusion

Factor investing aims at capturing systematic risk or mispricing premia. This thesis argues that using simple and, widely used equal- and value-weighting in the equity portfolio construction process is not an obvious choice when more advanced alternatives are considered.

In order to identify which factors are real, as opposed to being only random results of data mining, I cover number of recent prominent studies resulting in the selection of factors that are components of the currently most powerful asset pricing model or cannot be explained by it. Eight risk-based portfolio selection methods (strategies) are then applied in the construction of corresponding long-only factor portfolios. Using 50 years of fundamental and pricing data from CRSP and COMPUSTAT databases, performance of individual strategies is compared based on the financial efficiency, risk-adjusted performance, tail risk and the level of diversification achieved. Specific attention is given to the advanced methods focused on the idea of diversification across independent risk sources, thus being more effective representations of the underlying factor than their naive counterparts. From them, the most diversified portfolio proves to be the best way to form a factor portfolio. It consistently earns significantly higher returns than other strategies, while having comparable volatility and one of the smallest tail risk exposure comparable with the minimum variance strategy. Further it has the second largest out-of-sample diversification across risk sources, measured as entropy based effective number of bets. This contrasts heavily with the diversified risk parity portfolio, which has the largest out-of-sample diversification across risk sources, however its returns are only comparable to the naive equal-weighting and it has very high tail risk exposure as well as volatility, making it the least attractive method from a risk-adjusted performance perspective. Other two advanced methods, i.e. equal risk contributions and the hierarchical risk parity, have significantly lower out-of-sample diversification and their average returns as well as volatilities differ. The equal risk contribution method has average return and volatility comparable with the equal-weighting, while the hierarchical risk parity method has lower average return but lower volatility as well.

All these results alone offer a valuable insight for practitioners of factor investing. Contribution consists not only in showing what works significantly better than standardly used equal-weighting and value-weighting methods, but also in what does not add any value and in what should factor investors be extremely careful about.

Finally there is a mixed evidence, regarding the question of whether advanced portfolio selection methods can capture factor alpha more effectively than the naive equal-weighting. The performance of the most diversified portfolio would suggest the positive answer, however performances of other three advanced portfolio selection methods point more in favor of "no compesation for diversifiable risk" explanation. These differences provide a fertile ground for further research.

Bibliography

Ang, A., W. N. Goetzmann, & S. Schaefer (2009): "Evaluation of active management of the norwegian government pension fund–global." *report to the Norwegian Ministry of Finance* .

Artzner, P., F. Delbaen, J.-M. Eber, & D. Heath (1999): "Coherent measures of risk." *Mathematical finance* **9(3)**: pp. 203–228.

Asparouhova, E., H. Bessembinder, & I. Kalcheva (2013): "Noisy prices and inference regarding returns." *The Journal of Finance* **68(2)**: pp. 665–714.

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

Barillas, F. & J. Shanken (2015): "Comparing asset pricing models." *Technical report*, National Bureau of Economic Research.

Bauer, R., M. Cosemans, & P. C. Schotman (2010): "Conditional asset pricing and stock market anomalies in europe." *European Financial Management* **16(2)**: pp. 165–190.

Bender, J., R. Briand, F. Nielsen, & D. Stefek (2010): "Portfolio of risk premia: A new approach to diversification." *The Journal of Portfolio Management* **36(2)**: pp. 17–25.

Berk, J. B., R. C. Green, & V. Naik (1999): "Optimal investment, growth options, and security returns." *The Journal of Finance* **54(5)**: pp. 1553–1607.

Best, M. J. & R. R. Grauer (1991): "On the sensitivity of mean-variance-efficient portfolios to changes in asset means: some analytical and computational results." *Review of Financial Studies* **4(2)**: pp. 315–342.

Bondt, W. F. & R. Thaler (1985): "Does the stock market overreact?" *The Journal of finance* **40(3)**: pp. 793–805.

Brodeur, A., M. Lé, M. Sangnier, & Y. Zylberberg (2016): "Star wars: The empirics strike back." *American Economic Journal: Applied Economics* **8(1)**: pp. 1–32.

Campbell, J. Y., J. Hilscher, & J. Szilagyi (2008): "In search of distress risk." *The Journal of Finance* **63(6)**: pp. 2899–2939.

Carhart, M. M. (1997): "On persistence in mutual fund performance." *The Journal of finance* **52(1)**: pp. 57–82.

Cascon, A., C. Keating, & W. Shadwick (2002): "An introduction to omega." *The Finance Development Centre, Fuqua-Duke University* .

Chan, L. K., Y. Hamao, & J. Lakonishok (1991): "Fundamentals and stock returns in japan." *The Journal of Finance* **46(5)**: pp. 1739–1764.

Chang, A. C. & P. Li (2015): "Is economics research replicable? sixty published papers from thirteen journals say'usually not'." .

Chaudhuri, S. E. & A. W. Lo (2016): "Spectral portfolio theory." .

Chaves, D., J. Hsu, F. Li, & O. Shakernia (2011): "Risk parity portfolio vs. other asset allocation heuristic portfolios." *The Journal of Investing* **20(1)**: pp. 108–118.

Chekhlov, A., S. P. Uryasev, & M. Zabarankin (2000): "Portfolio optimization with drawdown constraints." .

Chopra, V. K. (1993): "Improving optimization." *The Journal of Investing* **2(3)**: pp. 51–59.

Choueifaty, Y. & Y. Coignard (2008): "Toward maximum diversification." *The Journal of Portfolio Management* **35(1)**: pp. 40–51.

Choueifaty, Y., T. Froidure, & J. Reynier (2013): "Properties of the most diversified portfolio." *Journal of Investment Strategies* **2(2)**: pp. 49–70.

Clarke, R., H. De Silva, & S. Thorley (2011): "Minimum-variance portfolio composition." *The Journal of Portfolio Management* **37(2)**: pp. 31–45.

Cochrane, J. H. (1991): "Production-based asset pricing and the link between stock returns and economic fluctuations." *The Journal of Finance* **46(1)**: pp. 209–237.

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

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

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

De Long, J. B., A. Shleifer, L. H. Summers, & R. J. Waldmann (1990): "Noise trader risk in financial markets." *Journal of political Economy* **98(4)**: pp. 703–738.

DeMiguel, V., L. Garlappi, & R. Uppal (2009): "Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy?" *Review of Financial Studies* **22(5)**: pp. 1915–1953.

Drew, M. E., T. Naughton, & M. Veeraraghavan (2003): "Firm size, book-to-market equity and security returns: Evidence from the shanghai stock exchange." *Australian Journal of Management* **28(2)**: p. 119.

Fama, E. F. (1998): "Market efficiency, long-term returns, and behavioral finance." *Journal of financial economics* **49(3)**: pp. 283–306.

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

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

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

Fama, E. F. & K. R. French (2015): "Incremental variables and the investment opportunity set." *Journal of Financial Economics* **117(3)**: pp. 470–488.

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

Fama, E. F. & J. D. MacBeth (1973): "Risk, return, and equilibrium: Empirical tests." *Journal of political economy* **81(3)**: pp. 607–636.

Feng, G., S. Giglio, & D. Xiu (2017): "Taming the factor zoo." .

Follmer, H. & A. Schied (2004): "Stochastic finance: An introduction in discrete time walter de gruyter." *Berlin, New York* .

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

Garlappi, L., R. Uppal, & T. Wang (2007): "Portfolio selection with parameter and model uncertainty: A multi-prior approach." *Review of Financial Studies* **20(1)**: pp. 41–81.

Gibbons, M. R., S. A. Ross, & J. Shanken (1989): "A test of the efficiency of a given portfolio." *Econometrica: Journal of the Econometric Society* pp. 1121–1152.

Green, J., J. R. Hand, & X. F. Zhang (2017): "The characteristics that provide independent information about average us monthly stock returns." *The Review of Financial Studies* p. hhx019.

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

Hasbrouck, J. (2009): "Trading costs and returns for us equities: Estimating effective costs from daily data." *The Journal of Finance* **64(3)**: pp. 1445–1477.

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

Hirshleifer, D. & D. Jiang (2010): "A financing-based misvaluation factor and the cross-section of expected returns." *The Review of Financial Studies* **23(9)**: pp. 3401–3436.

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

HOU, K., C. XUE, & L. ZHANG (2017): "Replicating anomalies." *Technical report*, National Bureau of Economic Research.

HUIJ, J., S. D. LANSDORP, D. BLITZ, & P. VAN VLIET (2014): "Factor investing: Long-only versus long-short." .

ILMANEN, A. & J. KIZER (2012): "The death of diversification has been greatly exaggerated." *The Journal of Portfolio Management* **38(3)**: pp. 15–27.

IOANNIDIS, J. P. (2005): "Why most published research findings are false." *PLoS medicine* **2(8)**: p. e124.

JEGADEESH, N. (1990): "Evidence of predictable behavior of security returns." *The Journal of finance* **45(3)**: pp. 881–898.

JEGADEESH, N. & S. TITMAN (1993): "Returns to buying winners and selling losers: Implications for stock market efficiency." *The Journal of finance* **48(1)**: pp. 65–91.

JENSEN, M. C., F. BLACK, & M. S. SCHOLES (1972): "The capital asset pricing model: Some empirical tests." .

KAN, R. & G. ZHOU (2007): "Optimal portfolio choice with parameter uncertainty." *Journal of Financial and Quantitative Analysis* **42(03)**: pp. 621–656.

KAYA, H. & W. LEE (2012): "Demystifying risk parity." .

KAYA, H., W. LEE, & Y. WAN (2012): "Risk budgeting with asset class and risk class approaches." *The Journal of Investing* **21(1)**: pp. 109–115.

KLEIN, R. W. & V. S. BAWA (1976): "The effect of estimation risk on optimal portfolio choice." *Journal of Financial Economics* **3(3)**: pp. 215–231.

KOZAK, S., S. NAGEL, & S. SANTOSH (2017): "Interpreting factor models." .

LAKONISHOK, J., A. SHLEIFER, & R. W. VISHNY (1994): "Contrarian investment, extrapolation, and risk." *The journal of finance* **49(5)**: pp. 1541–1578.

LEHMANN, B. N. (1990): "Fads, martingales, and market efficiency." *The Quarterly Journal of Economics* **105(1)**: pp. 1–28.

LEVENE, H. *et al.* (1960): "Robust tests for equality of variances." *Contributions to probability and statistics* **1**: pp. 278–292.

LINNAINMAA, J. T. & M. R. ROBERTS (2016): "The history of the cross section of stock returns." *Technical report*, National Bureau of Economic Research.

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

LINTNER, J. (1965b): "The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets." *The review of economics and statistics* pp. 13–37.

LINZMEIER, D. (2011): "Risk balanced portfolio construction." *Technical report*, Working paper.

LIU, L. X., T. M. WHITED, & L. ZHANG (2009): "Investment-based expected stock returns." *Journal of Political Economy* **117(6)**: pp. 1105–1139.

LO, A. W. & A. C. MACKINLAY (1990): "When are contrarian profits due to stock market overreaction?" *Review of Financial studies* **3(2)**: pp. 175–205.

LOHRE, H., U. NEUGEBAUER, & C. ZIMMER (2012): "Diversified risk parity strategies for equity portfolio selection." *The Journal of Investing* **21(3)**: pp. 111–128.

MAILLARD, S., T. RONCALLI, & J. TEÏLETCHE (2010): "The properties of equally weighted risk contribution portfolios." *The Journal of Portfolio Management* **36(4)**: pp. 60–70.

MANAGEMENT, N. B. I. (2017): "Government pension fund of norway."

MARKOWITZ, H. (1952): "Portfolio selection." *The journal of finance* **7(1)**: pp. 77–91.

MARKOWITZ, H. (1959): *Portfolio Selection, Efficent Diversification of Investments*. J. Wiley.

MCLEAN, R. D. & J. PONTIFF (2016): "Does academic research destroy stock return predictability?" *The Journal of Finance* **71(1)**: pp. 5–32.

MEUCCI, A. (2010): "Managing diversification." .

MICHAUD, R. O. (1989): "The markowitz optimization enigma: is' optimized'optimal?" *Financial Analysts Journal* **45(1)**: pp. 31–42.

MONTIER, J. (2009): *Behavioural Investing: a practitioners guide to applying behavioural finance*. John Wiley & Sons.

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

NAWROCKI, D. N. (1999): "A brief history of downside risk measures." *The Journal of Investing* **8(3)**: pp. 9–25.

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

PARTOVI, M. H., M. CAPUTO *et al.* (2004): "Principal portfolios: recasting the efficient frontier." *Economics Bulletin* **7(3)**: pp. 1–10.

PIOTROSKI, J. D. (2000): "Value investing: The use of historical financial statement information to separate winners from losers." *Journal of Accounting Research* pp. 1–41.

PONTIFF, J. (1996): "Costly arbitrage: Evidence from closed-end funds." *The Quarterly Journal of Economics* **111(4)**: pp. 1135–1151.

PONTIFF, J. (2006): "Costly arbitrage and the myth of idiosyncratic risk." *Journal of Accounting and Economics* **42(1)**: pp. 35–52.

LÓPEZ DE PRADO, M. (2016): "Building diversified portfolios that outperform out of sample." *The Journal of Portfolio Management* **42(4)**: pp. 59–69.

QIAN, E. (????): "Risk parity portfolios: Efficient portfolios through true diversification, 2005." *Panagora Research Pape* .

RITTER, J. R. (1991): "The long-run performance of initial public offerings." *The journal of finance* **46(1)**: pp. 3–27.

ROLL, R. (1977): "A critique of the asset pricing theory's tests part i: On past and potential testability of the theory." *Journal of financial economics* **4(2)**: pp. 129–176.

RONCALLI, T. (2014): "Introducing expected returns into risk parity portfolios: A new framework for asset allocation." .

ROSENBERG, B., K. REID, & R. LANSTEIN (1985): "Persuasive evidence of market inefficiency." *The Journal of Portfolio Management* **11(3)**: pp. 9–16.

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

Scherer, B. (2011): "A note on the returns from minimum variance investing." *Journal of Empirical Finance* **18(4)**: pp. 652–660.

Schwert, G. W. (2003): "Anomalies and market efficiency." *Handbook of the Economics of Finance* **1**: pp. 939–974.

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

Shleifer, A. & R. W. Vishny (1997): "The limits of arbitrage." *The Journal of Finance* **52(1)**: pp. 35–55.

Sloan, R. (1996): "Do stock prices fully reflect information in accruals and cash flows about future earnings?(digest summary)." *Accounting review* **71(3)**: pp. 289–315.

Sortino, F. A., R. v. d. Meer, & A. Plantinga (1999): "The dutch triangle." *The Journal of Portfolio Management* **26(1)**: pp. 50–57.

Stambaugh, R. F., J. Yu, & Y. Yuan (2012): "The short of it: Investor sentiment and anomalies." *Journal of Financial Economics* **104(2)**: pp. 288–302.

Stambaugh, R. F., J. Yu, & Y. Yuan (2015): "Arbitrage asymmetry and the idiosyncratic volatility puzzle." *The Journal of Finance* **70(5)**: pp. 1903–1948.

Stambaugh, R. F. & Y. Yuan (2016): "Mispricing factors." *The Review of Financial Studies* **30(4)**: pp. 1270–1315.

Titman, S., K. J. Wei, & F. Xie (2004): "Capital investments and stock returns." *Journal of financial and Quantitative Analysis* **39(04)**: pp. 677–700.

Tobin, J. (1958): "Liquidity preference as behavior towards risk." *The review of economic studies* **25(2)**: pp. 65–86.

Wang, H. & J. Yu (2013): "Dissecting the profitability premium." .

WILCOXON, F. (1945): "Individual comparisons by ranking methods." *Biometrics bulletin* **1(6)**: pp. 80–83.

WOERHEIDE, W. & D. PERSSON (1992): "An index of portfolio diversification." *Financial services review* **2(2)**: pp. 73–85.

YAN, X. & L. ZHENG (2017): "Fundamental analysis and the cross-section of stock returns: A data-mining approach." *The Review of Financial Studies* **30(4)**: pp. 1382–1423.

Appendix A

Definition of Anomalies

Components of MISP Factor

Net Stock Issues (NSI) The stock issuing market has long been viewed as producing an anomaly arising from sentiment-driven mispricing: smart managers issue shares when sentiment-driven traders push prices to overvalued levels. Ritter (1991) and Loughran and Ritter (1995) show that, in post-issue years, equity issuers underperform matching nonissuers with similar characteristics. Motivated by this evidence, Fama and French (2008) show that net stock issues and subsequent returns are negatively correlated At the end of June of year t, net stock issues, Nsi, are measured as the natural log of the ratio of the split-adjusted shares outstanding at the fiscal year ending in calendar year t-1 to the split-adjusted shares outstanding at the fiscal year ending in t-2. The split-adjusted shares outstanding is shares outstanding (Compustat annual item CSHO) times the adjustment factor (item AJEX). At the end of June of each year t, stocks with negative Nsi are sorted into two portfolios (1 and 2), stocks with zero Nsi into one portfolio (3), and stocks with positive Nsi into seven portfolios (4 to 10). Monthly decile returns are from July of year t to June of t + 1, and the deciles are rebalanced in June of t + 1.

Composite Equity Issues (CEI) Daniel and Titman (2006) find that issuers underperform nonissuers using a measure they denote as composite equity issuance. At the end of June of each year t, stocks are sorted into deciles based on composite equity issuance, Cei, which is the log growth rate in the market equity not attributable to stock return, log (ME t /ME t-5) - r(t - 5, t). r(t

- 5, t) is the cumulative log stock return from the last trading day of June in year t - 5 to the last trading day of June in year t, and ME t is the market equity (from CRSP) on the last trading day of June in year t. Monthly decile returns are from July of year t to June of t + 1, and the deciles are rebalanced in June of t + 1.

Accruals (ACCR) Sloan (1996) shows that firms with high accruals earn abnormally lower average returns than firms with low accruals, and he suggests that investors overestimate the persistence of the accrual component of earnings when forming earnings expectations. Prior to 1988, the balance sheet approach in Richardson, Sloan, Soliman, and Tuna (2005) is used to measure total accruals, Ta, as dWc + dNco + dFin. dWc is the change in net non-cash working capital. Net non-cash working capital is current operating asset (Coa) minus current operating liabilities (Col), with Coa = current assets (Compustat annual item ACT) - cash and short-term investments (item CHE) and Col = current liabilities (item LCT) - debt in current liabilities (item DLC). dNco is the change in net non-current operating assets. Net non-current operating assets are non-current operating assets (Nca) minus non-current operating liabilities (Ncl), with Nca = total assets (item AT) - current assets - long-term investments (item IVAO), and Ncl = total liabilities (item LT) - current liabilities - long-term debt (item DLTT). dFin is the change in net financial assets. Net financial assets are financial assets (Fna) minus financial liabilities (Fnl), with Fna = short-term investments (item IVST) + long-term investments, and Fnl = long-term debt + debt in current liabilities + preferred stocks (item PSTK). Missing changes in debt in current liabilities, long-term investments, long-term debt, short-term investments, and preferred stocks are set to zero. Starting from 1988, the cash flow approach is used to measure Ta as net income (item NI) minus total operating, investing, and financing cash flows (items OANCF, IVNCF, and FINCF) plus sales of stocks (item SSTK, zero if missing) minus stock repurchases and dividends (items PRSTKC and DV, zero if missing). Data from the statement of cash flows are only available since 1988. At the end of June of each year t, stocks are sorted into deciles based on Ta for the fiscal year ending in calendar year t - 1 scaled by total assets for the fiscal year ending in t - 2. Monthly decile returns are calculated from July of year t to June of t + 1, and the deciles are rebalanced in June of t + 1.

Net Operating Assets (NOA) Hirshleifer, Hou, Teoh, and Zhang (2004) find that net operating assets, defined as the difference on the balance sheet between all operating assets and all operating liabilities, scaled by total assets, is a strong negative predictor of long-run stock returns. The authors suggest that investors with limited attention tend to focus on accounting profitability, neglecting information about cash profitability, in which case net operating assets (equivalently measured as the cumulative difference between operating income and free cash flow) captures such a bias. Noa is operating assets minus operating liabilities. Operating assets are total assets (item AT) minus cash and short-term investment (item CHE), and minus other investment and advances (item IVAO, zero if missing). Operating liabilities are total assets minus debt in current liabilities (item DLC, zero if missing), minus long-term debt (item DLTT, zero if missing), minus minority interests (item MIB, zero if missing), minus preferred stocks (item PSTK, zero if missing), and minus common equity (item CEQ).

Asset Growth (AG) Cooper, Gulen, and Schill (2008) find that companies that grow their total assets more earn lower subsequent returns. They suggest that this phenomenon is due to investors' initial overreaction to changes in future business prospects implied by asset expansions. At the end of June of each year t, stocks are sorted into deciles based on investment-to-assets, I/A, which is measured as total assets (Compustat annual item AT) for the fiscal year ending in calendar year t-1 divided by total assets for the fiscal year ending in t-2 minus one. Monthly decile returns are computed from July of year t to June of t + 1, and the deciles are rebalanced in June of t + 1.

Distress (DIST) Financial distress is often invoked to explain otherwise anomalous patterns in the cross-section of stock returns. However, Campbell, Hilscher, and Szilagyi (2008) find that firms with high failure probability have lower rather than higher subsequent returns. The authors suggest that their finding is a challenge to standard models of rational asset pricing. Failure probability is estimated with a dynamic logit model that uses several equity market variables, such as stock price, book-to-market, stock volatility, size relative to the S&P500, and cumulative excess return relative to the S&P500. Specifically, using the above study's equations (2) and (3) along with its Table IV (12-month column), the distress anomaly measure - failure probability - is computed as

$$\begin{aligned}\pi = &-20.26\ NIMTAAVG + 1.42\ TLMTA - 7.13\ EXRETAVG + 1.41\ SIGMA \\ &- 0.045\ RSIZE - 2.13\ CASHMTA + 0.075\ MB - 0.058\ PRICE - 9.16,\end{aligned}$$

where

$$NIMTAAVG_{t-1,t-12} = \frac{1-\Phi^3}{1-\Phi^{12}}\left(NIMTA_{t-1,t-3} + \cdots + \Phi^9 NIMTA_{t-10,t-12}\right)$$

$$EXRETAVG_{t-1,t-12} = \frac{1-\Phi}{1-\Phi^{12}}\left(EXRET_{t-1} + \cdots + \Phi^{11} EXRET_{t-12}\right)$$

and $\Phi = 2^{-1/3}$. $NIMTA$ is net income (Compustat quarterly item NIQ) divided by firm scale, where the latter is computed as the sum of total liabilities (item LTQ) and market equity capitalization (data from CRSP). $EXRET_s$ is the stock's monthly log return in month s minus the log return on the S&P500 index. Missing values for $NIMTA$ and $EXRET$ are replaced by those quantities' cross-sectional means. $TLMTA$ equals total liabilities divided by firm scale. $SIGMA$ is the stock's daily standard deviation for the most recent three months, expressed on an annualized basis. At least five non-zero daily returns are required. $RSIZE$ is the log of the ratio of the stock's market capitalization to that of the S&P500 index. $CASHMTA$ equals cash and short-term investment (item CHEQ) divided by firm scale. MB is the market-to-book ratio. Following Campbell, Hilscher, and Szilagyi (2008), book equity is increased by 10% of the difference between market equity and book equity. If the resulting value of book equity is negative, then book equity is set to \$1. $PRICE$ is the log of the share price, truncated above at \$15. All explanatory variables except $PRICE$ are winsorized above and below at the 5% level in the cross section. CRSP based variables, $EXRETAVG$, $SIGMA$, $RSIZE$ and $PRICE$ are for month $t-1$. NIQ is for the most recent quarter for which the reporting date provided by Compustat (item RDQ) precedes the end of month $t-1$, whereas the items requiring information from the balance sheet (LTQ, $CHEQ$ and MB) are for the prior quarter.

O-score (O) This distress measure, from Ohlson (1980), predicts return s in a manner similar to the measure above. It is the probability of bankruptcy estimated in a static model using accounting variables. Following Ohlson (1980), it is constructed as:

$$
\begin{aligned}
O = &-0.407\ SIZE + 6.03\ TLTA - 1.43\ WCTA + 0.076\ CLCA - 1.72\ OENEG \\
&= -2.37\ NITA - 1.83\ FUTL + 0.285\ INTWO - 0.521\ CHIN - 1.32,
\end{aligned}
$$

where $SIZE$ is the log of total assets (Compustat annual item AT), $TLTA$ is the book value of debt (item DLC plus item DLTT) divided by total assets, $WCTA$ is working capital (item ACT minus item LCT) divided by total assets, CLCA is current liabilities (item LCT) divided by current assets (item ACT), $ONEEG$ is 1 if total liabilities (item LT) exceed total assets and is zero otherwise, $NITA$ is net income (item NI) divided by total assets, $FUTL$ is funds provided by operations (item PI) divided by total liabilities, $INTWO$ is equal to 1 if net income (item NI) is negative for the last 2 years and zero otherwise, $CHIN$ is $(NI_j - NI_{j-1})/(|NI_j| - |NI_{j-1}|)$, in which NI_j is the income (item NI) for year j, which is the most recent reporting year that ends (according to item DATADATE) at least four months before the end of month $t-1$.

Investment-to-Assets (IA) Titman, Wei, and Xie (2004) and Xing (2008) show that higher past investment predicts abnormally lower future returns. Titman, Wei, and Xie (2004) attribute this anomaly to investors' initial underreaction to overinvestment caused by managers' empire-building behavior. Following the above studies, investment-to-assets is computed as the changes in gross property, plant, and equipment (Compustat annual item PPEGT) plus changes in inventory (item INVT), divided by lagged total assets (item AT). The most re-cent reporting year used is the one that ends (according to item DATADATE) at least four months before the end of month $t-1$.

Momentum (MOM) The momentum effect, discovered by Jegadeesh and Titman (1993), is one of the most robust anomalies in asset pricing. It refers to the phenomenon whereby high (low) past recent recent returns forecast high (low) future returns. The momentum ranking at the end of month $t-1$ uses the cumulative returns from month $t-12$ to month $t-2$. This is the choice of ranking variable used by Carhart (1997) to construct the widely used momentum factor.

Gross Profitability (GP): Novy-Marx (2013) shows that sorting on the ratio of gross profit to assets creates abnormal benchmark-adjusted returns, with more profitable firms having higher returns than less profitable ones. He argues that gross profit is the cleanest accounting measure of true economic profitability. The farther down the income statement one goes, the more polluted profitability measures become, and the less related they are to true economic profitability. Following Novy-Marx (2013), gross profits-to-assets, Gpa, is measured as total revenue (Compustat annual item REVT) minus cost of goods sold (item COGS) divided by total assets (item AT, the denominator is current, not lagged, total assets). At the end of June of each year t, stocks are sorted into deciles based on Gpa for the fiscal year ending in calendar year $t-1$. Monthly decile returns are calculated from July of year t to June of $t+1$, and the deciles are rebalanced in June of $t+1$.

Significant Anomalies

Rer, Industry-adjusted Real Estate Ratio: Following Tuzel (2010), the real estate ratio is measured as the sum of buildings (Compustat annual item PPENB) and capital leases (item PPENLS) divided by net property, plant, and equipment (item PPENT) prior to 1983. From 1984 onward, the real estate ratio is the sum of buildings at cost (item FATB) and leases at cost (item FATL) divided by gross property, plant, and equipment (item PPEGT). Industry-adjusted real estate ratio, Rer, is the real estate ratio minus its industry aver- age. Industries are defined by two-digit SIC codes. To alleviate the impact of outliers, the real estate ratio at the 1st and 99th percentiles of its distribution are winsorized each year before computing Rer. Following Tuzel (2010), industries with fewer than five firms are excluded. At the end of June of each year t, stocks are sorted into deciles based on Rer for the fiscal year ending in calendar year t - 1. Monthly decile returns are calculated from July of year t to June of t + 1, and the deciles are rebal- anced in June of t+1. Because the real estate data start in 1969, the Rer portfolios start in July 1970.

Eprd: Following Francis, Lafond, Olsson, and Schipper (2004), earnings persistence, Eper, and earnings predictability, Eprd, is estimated from a first-order autoregressive model for annual split-adjusted earnings per share (Compustat annual item EPSPX divided by item AJEX). At the end of June of each year t, the autoregressive model is estimated in the ten-year rolling window up to the fiscal year ending in calendar year t - 1. Only firms with a complete ten-year history are included. Eper is measured as the slope coefficient and Eprd is measured as the residual volatility.

Cla, Cash-based Operating Profits-to-lagged Assets: Cash-based operating profits-to-lagged assets, Cla, is total revenue (Compustat annual item REVT) minus cost of goods sold (item COGS), minus selling, general, and administrative expenses (item XSGA), plus research and development expenditures (item XRD, zero if missing), minus change in accounts receivable (item RECT), minus change in inventory (item INVT), minus change in prepaid expenses (item XPP), plus change in deferred revenue (item DRC plus item DRLT), plus change in trade accounts payable (item AP), and plus change in accrued expenses (item XACC), all scaled by one-year-lagged book assets (item AT). All changes are annual changes in balance sheet items and missing changes are set to zero. At the end of June of each year t, stocks are sorted into deciles based on Cla for the fiscal year ending in calendar year t - 1. Monthly decile returns are calculated from July of year t to June of t + 1, and the deciles are rebalanced in June of t + 1.

Cop, Cash-based Operating Profitability: Following Ball, Gerakos, Linnainmaa, and Nikolaev (2016), cash-based operating prof- itability, Cop, is measured as total revenue (Compustat annual item REVT) minus cost of goods sold (item COGS), minus selling, general, and administrative expenses (item XSGA), plus research and de- velopment expenditures (item XRD, zero if missing), minus change in accounts receivable (item RECT), minus change in inventory (item INVT), minus change in prepaid expenses (item XPP), plus change in deferred revenue (item DRC plus item DRLT), plus change in trade accounts payable (item AP), and plus change in accrued expenses (item XACC), all scaled by book assets (item AT, the denominator is current, not lagged, total assets). All changes are annual changes in balance sheet items and missing changes are set to zero. At the end of June of each year t, stocks are sorted into deciles based on Cop for the fiscal year ending in calendar year t - 1. Monthly decile returns are calculated from July of year t to June of t + 1, and the deciles are rebalanced in June of t + 1.

dFin, Changes in Net Financial Assets dFin is the change in net financial assets. Net financial assets are financial assets (Fna) minus financial liabilities (Fnl), with Fna = short-term investments (Compustat annual item IVST) + long-term investments (item IVAO), and Fnl = long-term debt (item DLTT) + debt in current liabilities (item DLC) + preferred stock (item PSTK).

Nop, Net payout yield Per Boudoukh, Michaely, Richardson, and Roberts (2007), total payouts are dividends on common stock (Compustat annual item DVC) plus repurchases. Repurchases are the total expenditure on the purchase of common and preferred stocks (item PRSTKC) plus any reduction (negative change over the prior year) in the value of the net number of preferred stocks outstanding (item PSTKRV). Net payouts equal total payouts minus equity issuances, which are the sale of common and preferred stock (item SSTK) minus any increase (positive change over the prior year) in the value of the net number of preferred stocks outstanding (item PSTKRV). At the end of June of each year t, stocks are sorted into deciles based on total payouts (net payouts) for the fiscal year ending in calendar year t - 1 divided by the market equity (from CRSP) at the end of December of t - 1 (Op and Nop, respectively). For firms with more than one share class, the market equity for all share classes is merged before computing Op and Nop . Firms with non-positive total payouts (zero net payouts) are excluded. Monthly decile returns are calculated from July of year t to June of t + 1, and the deciles are rebalanced in June of t + 1. Because the data on total expenditure and the sale of common and preferred stocks start in 1971, the Op and Nop portfolios start in July 1972.

Ivc, Inventory changes At the end of June of each year t, stocks are sorted into deciles based on inventory changes, Ivc, which is the annual change in inventory (Compustat annual item INVT) scaled by the average of total assets (item AT) for the fiscal years ending in t - 2 and t - 1. We exclude firms that carry no inventory for the past two fiscal years. Monthly decile returns are calculated from July of year t to June of t + 1, and the deciles are rebalanced in June of t + 1.

Rdm,R&D Expense-to-market At the end of June of each year t, stocks are sorted into deciles based on R&D-to-market, Rdm, which is R&D expenses (Compustat annual item XRD) for the fiscal year ending in calendar year $t-1$ divided by the market equity (from CRSP) at the end of December of $t-1$. For firms with more than one share class, the market equity for all share classes is merged before computing Rdm. We keep only firms with positive R&D expenses. Monthly decile returns are calculated from July of year t to June of t + 1, and the deciles are rebalanced in June of t + 1. Because the accounting treatment of R&D expenses was standardized in 1975, the Rdm portfolios start in July 1976.

Appendix B

Results for Individual Factors

B.1 Results for Constituents of Mispricing Factors

NOA	EW	VW	MV	NRP	ERC	MDP	HRP	DRP
Mean	26.29	24.44	18.03	23.18	26.58	33.33	23.06	25.45
Stdev	21.02	22.23	14.59	19.04	20.85	19.87	16.67	28.84
Skewness	-0.19	0.24	-0.10	-0.18	-0.11	-0.09	-0.31	0.14
Kurtosis	7.13	7.24	7.17	7.58	6.39	11.18	11.33	6.61
Sharpe	1.34	1.16	1.22	1.27	1.36	1.89	1.44	0.94
Sortino	1.69	1.52	1.63	1.63	1.73	2.38	1.85	1.22
UPR	2.81	2.93	2.89	2.80	2.83	2.68	2.70	2.90
Omega	4.37	3.74	3.93	4.16	4.47	6.69	4.90	3.03
VaR 5%	1.99	2.04	1.32	1.82	1.98	1.70	1.47	2.60
CVaR 5%	3.08	3.16	2.08	2.80	3.05	2.81	2.44	4.14
MDD	-54.59	-49.17	-40.11	-50.08	-46.72	-56.41	-46.84	-59.29
ENB	1.2	1.33	2.56	1.77	1.84	2.87	2.04	4.3
Gini(w)	0.00	0.6	0.68	0.59	0.47	0.66	0.71	0.6

Table B.1: Net Operating Assets (NOA)

CEI	EW	VW	MV	NRP	ERC	MDP	HRP	DRP
Mean	27.36	26.31	23.42	25.33	27.49	35.45	25.86	25.81
Stdev	20.30	21.13	14.80	18.38	20.29	19.36	16.26	31.08
Skewness	-0.36	-0.13	-0.49	-0.54	-0.39	-0.52	-0.71	-0.14
Kurtosis	5.34	5.66	5.30	5.75	5.30	7.10	8.12	6.88
Sharpe	1.28	1.08	1.64	1.17	1.39	1.92	1.55	0.81
Sortino	1.61	1.73	1.98	1.66	1.55	2.43	2.12	1.10
UPR	2.57	2.76	2.69	2.72	2.75	2.37	2.60	2.56
Omega	4.50	4.03	5.35	4.40	4.51	7.09	5.42	2.73
VaR 5%	1.72	1.73	1.35	1.61	1.69	1.65	1.36	2.56
CVaR 5%	2.84	2.94	1.90	2.55	2.74	2.61	2.38	4.46
MDD	-46.64	-46.03	-36.86	-45.20	-46.53	-44.24	-43.17	-80.33
ENB	1.2	1.41	1.76	1.77	2.86	2.05	4.11	103.25
Gini(w)	0.00	0.61	0.69	0.53	0.41	0.71	0.70	0.56

Table B.2: Composite Equity Issues (CEI)

NSI	EW	VW	MV	NRP	ERC	MDP	HRP	DRP
Mean	27.57	26.44	23.75	25.36	27.56	35.56	25.91	25.90
Stdev	20.56	21.30	14.81	18.71	20.50	19.38	16.48	31.21
Skewness	-0.31	0.06	-0.25	-0.32	-0.31	-0.35	-0.53	0.10
Kurtosis	5.46	5.93	5.51	5.76	5.48	7.39	8.31	6.94
Sharpe	1.45	1.33	1.68	1.44	1.45	2.10	1.68	0.89
Sortino	1.83	1.74	2.23	1.84	1.84	2.60	2.13	1.15
UPR	2.84	2.92	2.84	2.83	2.84	2.69	2.74	2.85
Omega	4.72	4.32	5.63	4.72	4.74	7.36	5.66	2.93
VaR 5%	1.96	1.97	1.36	1.78	1.95	1.73	1.52	2.70
CVaR 5%	2.97	2.98	2.08	2.70	2.96	2.79	2.39	4.47
MDD	-46.41	-45.79	-36.54	-44.91	-46.37	-44.03	-43.02	-80.26
ENB	1.28	1.34	2.38	1.61	1.85	2.69	1.90	3.80
Gini(w)	0.00	0.63	0.68	0.64	0.49	0.7	0.8	0.66

Table B.3: Net Stock Issues (NSI)

GP	EW	VW	MV	NRP	ERC	MDP	HRP	DRP
Mean	22.84	16.45	15.60	20.06	22.78	29.48	21.07	25.34
Stdev	18.96	18.89	13.62	17.19	18.90	18.45	15.55	27.87
Skewness	-0.27	0.12	-0.23	-0.26	-0.27	-0.32	-0.36	0.38
Kurtosis	5.90	7.48	5.64	6.32	5.91	12.15	9.72	6.79
Sharpe	1.25	0.84	1.09	1.18	1.25	1.75	1.38	0.96
Sortino	1.59	1.13	1.46	1.51	1.59	2.19	1.77	1.28
UPR	2.87	3.00	2.97	2.86	2.87	2.71	2.76	3.05
Omega	3.95	2.61	3.40	3.73	3.96	5.98	4.55	3.03
VaR 5%	1.83	1.75	1.25	1.66	1.82	1.62	1.45	2.48
CVaR 5%	2.79	2.67	1.94	2.54	2.78	2.67	2.30	3.81
MDD	-52.03	-49.13	-42.40	-50.97	-52.00	-53.60	-48.62	-60.25
ENB	1.25	1.33	2.39	1.66	1.78	2.75	1.99	3.96
Gini(w)	0.00	0.65	0.71	0.57	0.49	0.75	0.71	0.67

Table B.4: Gross Profitability (GP)

MOM	EW	VW	MV	NRP	ERC	MDP	HRP	DRP
Mean	23.87	19.81	17.62	21.28	23.87	29.95	21.50	18.56
Stdev	18.77	19.28	13.33	17.09	18.72	17.84	15.42	26.28
Skewness	-0.20	0.07	-0.12	-0.15	-0.20	-0.16	-0.06	0.07
Kurtosis	4.77	7.95	5.27	5.36	4.79	7.72	7.54	9.61
Sharpe	1.33	1.04	1.30	1.27	1.34	1.85	1.43	0.70
Sortino	1.73	1.40	1.77	1.68	1.73	2.36	1.90	0.93
UPR	2.94	3.00	2.94	2.94	2.94	2.79	2.86	2.86
Omega	4.22	3.27	4.20	4.05	4.24	6.29	4.70	2.28
VaR 5 %	1.78	1.79	1.20	1.62	1.77	1.61	1.41	2.34
CVaR 5%	2.65	2.62	1.85	2.40	2.64	2.52	2.17	3.77
MDD	-50.22	-48.42	-46.80	-41.48	-49.10	-43.37	-40.70	-79.21
ENB	1.23	1.3	2.26	1.56	1.77	2.63	1.91	3.95
Gini(w)	0.00	0.65	0.65	0.61	0.47	0.76	0.79	0.65

Table B.5: Momentum (MOM)

ACCR	EW	VW	MV	NRP	ERC	MDP	HRP	DRP
Mean	18.95	16.45	12.15	16.26	18.88	23.48	16.28	20.09
Stdev	19.18	18.89	13.45	17.28	19.09	18.24	15.53	29.34
Skewness	-0.40	0.12	-0.05	-0.45	-0.40	-0.13	-0.40	0.06
Kurtosis	8.55	7.48	5.92	9.84	8.63	14.18	13.56	6.66
Sharpe	0.98	0.84	0.81	0.91	0.98	1.34	1.01	0.69
Sortino	1.27	1.13	1.11	1.18	1.27	1.73	1.32	0.91
UPR	2.82	3.00	3.00	2.81	2.82	2.66	2.71	2.97
Omega	3.15	2.61	2.51	2.90	3.15	4.66	3.37	2.17
VaR 5%	1.78	1.75	1.25	1.60	1.78	1.55	1.39	2.65
CVaR 5%	2.84	2.67	1.91	2.56	2.82	2.63	2.28	4.19
MDD	-53.39	-49.13	-40.41	-53.23	-53.36	-52.92	-54.78	-90.17
ENB	1.22	1.44	2.39	1.6	1.72	2.6	1.89	3.82
Gini(w)	0.00	0.66	0.68	0.57	0.50	0.71	0.73	0.63

Table B.6: Accruals (ACCR)

IA	EW	VW	MV	NRP	ERC	MDP	HRP	DRP
Mean	22.84	16.45	15.60	20.06	22.78	29.86	20.74	25.81
Stdev	18.96	18.90	13.62	17.19	18.90	18.63	15.72	29.11
Skewness	-0.27	0.12	-0.23	-0.26	-0.27	-0.01	-0.71	0.36
Kurtosis	5.90	7.48	5.64	6.32	5.91	16.23	15.18	14.96
Sharpe	1.25	0.84	1.09	1.18	1.25	1.76	1.34	0.94
Sortino	1.59	1.13	1.46	1.51	1.59	2.22	1.70	1.24
UPR	2.87	3.00	2.97	2.86	2.87	2.71	2.71	2.94
Omega	3.95	2.61	3.40	3.73	3.96	6.07	4.44	3.07
VaR 5%	1.83	1.75	1.25	1.66	1.82	1.62	1.45	2.49
CVaR 5%	2.79	2.67	1.94	2.54	2.78	2.67	2.32	3.92
MDD	-52.03	-49.13	-42.40	-50.97	-52.00	-53.60	-48.62	-60.25
ENB	1.32	1.27	2.4	1.68	1.78	2.73	1.9	3.93
Gini(w)	0.00	0.63	0.73	0.63	0.50	0.77	0.75	0.66

Table B.7: Investments to Assets (IA)

AG	EW	VW	MV	NRP	ERC	MDP	HRP	DRP
Mean	20.25	16.97	9.31	17.03	20.18	22.21	15.97	12.08
Stdev	20.11	18.86	14.50	18.06	20.01	18.51	15.88	27.11
Skewness	-0.34	-0.14	-0.35	-0.41	-0.34	-0.53	-0.44	-0.15
Kurtosis	9.59	8.33	6.57	10.74	9.67	14.00	13.75	7.39
Sharpe	1.02	0.88	0.54	0.92	1.02	1.24	0.97	0.40
Sortino	1.31	1.15	0.72	1.19	1.31	1.58	1.26	0.54
UPR	2.77	2.89	2.92	2.75	2.77	2.68	2.71	2.97
Omega	3.33	2.76	1.64	3.00	3.34	4.20	3.21	1.22
VaR 5%	1.83	1.74	1.34	1.64	1.82	1.61	1.41	2.48
CVaR 5%	2.99	2.76	2.15	2.69	2.97	2.69	2.34	3.95
MDD	-61.11	-57.88	-62.50	-62.74	-61.12	-53.43	-59.52	-92.41
ENB	1.22	1.25	2.37	1.57	1.68	2.58	1.94	3.81
Gini(w)	0.00	0.62	0.68	0.57	0.45	0.69	0.7	0.64

Table B.8: Asset Growth (AG)

O-Score	EW	VW	MV	NRP	ERC	MDP	HRP	DRP
Mean	18.66	16.32	11.93	16.04	18.77	23.42	16.12	20.00
Stdev	19.06	18.57	13.30	17.05	18.98	17.91	15.31	29.27
Skewness	-0.43	-0.19	-0.36	-0.66	-0.41	-0.21	-0.67	-0.23
Kurtosis	8.40	7.36	5.83	9.80	8.59	13.86	13.29	6.61
Sharpe	0.91	0.62	0.56	0.64	0.93	1.01	0.93	0.49
Sortino	1.07	0.81	0.89	0.99	1.10	1.48	1.28	0.91
UPR	2.50	2.83	2.89	2.69	2.55	2.53	2.45	2.82
Omega	2.87	2.58	2.35	2.74	3.06	4.35	3.07	2.02
VaR 5%	1.75	1.68	1.16	1.44	1.67	1.29	1.23	2.36
CVaR 5%	2.51	2.60	1.71	2.28	2.50	2.53	2.04	4.18
MDD	-53.68	-49.15	-40.65	-53.56	-53.53	-52.99	-54.98	-90.38
ENB	1.29	1.34	2.26	1.52	1.8	2.71	2.0	3.87
Gini(w)	0.00	0.66	0.68	0.63	0.47	0.73	0.71	0.62

Table B.9: O-Score (O)

DIST	EW	VW	MV	NRP	ERC	MDP	HRP	DRP
Mean	27.50	26.38	23.63	25.07	27.39	35.46	25.59	25.63
Stdev	20.45	20.98	14.48	18.57	20.28	19.24	16.39	31.18
Skewness	-0.40	-0.03	-0.38	-0.38	-0.63	-0.38	-0.74	-0.05
Kurtosis	5.32	5.81	5.22	5.67	5.47	7.08	8.02	6.85
Sharpe	1.28	1.29	1.65	1.16	1.31	1.88	1.37	0.81
Sortino	1.83	1.73	1.99	1.53	1.68	2.47	1.95	0.86
UPR	2.62	2.84	2.58	2.79	2.78	2.49	2.66	2.62
Omega	4.60	4.23	5.30	4.71	4.52	7.35	5.41	2.85
VaR 5%	1.93	1.75	1.12	1.55	1.93	1.63	1.37	2.60
CVaR 5%	2.70	2.94	1.88	2.37	2.66	2.72	2.39	4.14
MDD	-46.70	-46.10	-36.68	-45.16	-46.44	-44.16	-43.30	-80.30
ENB	1.31	1.34	2.46	1.56	1.7	2.63	1.92	3.93
Gini(w)	0.00	0.66	0.68	0.6	0.46	0.73	0.75	0.64

Table B.10: Distress (DIST)

B.2 Results for Significant Anomalies

RDM	EW	VW	MV	NRP	ERC	MDP	HRP	DRP
Mean	19.19	17.20	14.30	16.99	19.15	29.74	17.38	21.38
Stdev	19.23	17.70	12.17	17.44	19.05	18.35	15.26	33.31
Skewness	-0.20	-0.02	-0.11	-0.19	-0.20	-0.13	-0.50	0.21
Kurtosis	6.22	7.69	7.73	6.42	6.29	12.52	10.58	7.47
Sharpe	1.00	0.95	1.10	0.95	1.00	1.78	1.11	0.66
Sortino	1.29	1.27	1.49	1.24	1.30	2.26	1.44	0.87
UPR	2.89	2.97	2.91	2.89	2.89	2.64	2.78	2.89
Omega	3.13	2.96	3.51	2.98	3.16	6.32	3.61	2.13
VaR 5%	1.86	1.66	1.11	1.69	1.84	1.59	1.44	2.91
CVaR 5%	2.83	2.49	1.73	2.56	2.80	2.64	2.25	4.79
MDD	-51.64	-44.94	-40.56	-52.03	-51.57	-48.56	-47.62	-87.97
ENB	1.32	1.35	2.4	1.72	1.74	3.01	2.11	4.09
Gini(w)	0.00	0.63	0.68	0.56	0.54	0.64	0.78	0.59

Table B.11: R&D expense-to-market (RDM)

IVC	EW	VW	MV	NRP	ERC	MDP	HRP	DRP
Mean	23.29	23.09	15.07	20.61	23.22	27.31	19.51	25.68
Stdev	18.53	17.99	12.68	16.77	18.45	16.97	14.88	27.27
Skewness	-0.24	-0.09	-0.34	-0.23	-0.24	-0.20	-0.46	0.37
Kurtosis	6.70	6.00	7.29	7.24	6.73	13.04	14.79	8.06
Sharpe	1.31	1.34	1.13	1.25	1.31	1.74	1.32	1.00
Sortino	1.68	1.74	1.51	1.62	1.68	2.21	1.71	1.33
UPR	2.82	2.94	2.88	2.81	2.81	2.67	2.69	2.98
Omega	4.27	4.25	3.62	4.08	4.27	6.06	4.46	3.24
VaR 5%	1.74	1.67	1.16	1.58	1.74	1.47	1.33	2.40
CVaR 5%	2.72	2.55	1.83	2.46	2.71	2.45	2.17	3.74
MDD	-50.37	-42.58	-54.83	-50.96	-50.31	-45.02	-52.09	-64.55
ENB	1.34	1.3	2.43	1.65	1.83	2.98	1.99	4.15
Gini(w)	0.00	0.65	0.69	0.59	0.49	0.63	0.76	0.63

Table B.12: Inventory changes (IVC)

dFIN	EW	VW	MV	NRP	ERC	MDP	HRP	DRP
Mean	27.52	26.34	23.19	25.16	27.40	35.60	25.71	26.32
Stdev	20.48	21.18	14.76	18.63	20.42	19.39	16.43	31.19
Skewness	-0.31	0.07	-0.25	-0.31	-0.31	-0.35	-0.53	0.12
Kurtosis	5.44	5.96	5.46	5.73	5.46	7.22	8.24	7.04
Sharpe	1.45	1.33	1.64	1.43	1.45	2.10	1.67	0.90
Sortino	1.82	1.72	2.15	1.81	1.81	2.57	2.10	1.16
UPR	2.88	2.95	2.89	2.86	2.88	2.73	2.77	2.89
Omega	4.61	4.23	5.33	4.58	4.61	7.18	5.49	2.93
VaR 5%	1.97	1.98	1.37	1.79	1.96	1.76	1.55	2.73
CVaR 5%	2.97	2.98	2.09	2.71	2.96	2.82	2.40	4.50
MDD	-46.41	-45.79	-36.54	-44.91	-46.37	-44.03	-43.02	-80.26
ENB	1.34	1.29	2.43	1.67	1.73	3.07	2.00	4.28
Gini(w)	0.00	0.65	0.66	0.57	0.5	0.72	0.79	0.62

Table B.13: Changes in net financial assets (dFin)

NOP	EW	VW	MV	NRP	ERC	MDP	HRP	DRP
Mean	20.61	24.47	14.92	18.11	20.58	27.26	18.05	17.64
Stdev	20.74	21.20	13.02	18.68	20.65	18.79	15.75	32.11
Skewness	-0.27	-0.05	-0.07	-0.27	-0.27	-0.36	-0.30	0.45
Kurtosis	6.17	6.14	7.79	6.88	6.21	9.25	8.71	6.35
Sharpe	1.01	1.22	1.09	0.96	1.01	1.56	1.13	0.54
Sortino	1.29	1.56	1.47	1.24	1.29	1.97	1.47	0.74
UPR	2.90	2.91	2.87	2.85	2.89	2.75	2.77	3.01
Omega	3.15	3.87	3.51	3.03	3.16	5.23	3.71	1.70
VaR 5%	2.00	1.98	1.21	1.78	1.99	1.73	1.48	2.95
CVaR 5%	3.05	3.04	1.87	2.77	3.04	2.75	2.34	4.54
MDD	-55.06	-54.71	-59.60	-56.25	-54.73	-47.25	-56.10	-67.28
ENB	1.29	1.44	2.37	1.61	1.78	3.11	2.07	4.13
Gini(w)	0.00	0.62	0.69	0.59	0.51	0.63	0.82	0.6

Table B.14: Net payout yield (NOP)

RER	EW	VW	MV	NRP	ERC	MDP	HRP	DRP
Mean	21.44	20.27	15.70	18.91	21.40	26.79	19.22	22.90
Stdev	19.11	18.96	15.09	17.54	19.09	19.73	16.37	26.99
Skewness	-0.17	-0.01	-0.18	-0.20	-0.17	-0.28	-0.39	0.11
Kurtosis	6.55	5.36	6.58	6.37	6.54	9.35	11.74	5.27
Sharpe	1.15	1.08	1.00	1.08	1.15	1.46	1.18	0.88
Sortino	1.49	1.43	1.33	1.41	1.49	1.84	1.52	1.16
UPR	2.88	3.01	2.94	2.89	2.88	2.74	2.79	3.05
Omega	3.66	3.35	3.13	3.41	3.66	4.90	3.83	2.71
VaR 5%	1.82	1.78	1.37	1.67	1.81	1.74	1.49	2.47
CVaR 5%	2.79	2.66	2.15	2.55	2.78	2.89	2.38	3.79
MDD	-53.38	-49.34	-51.18	-52.93	-53.79	-46.92	-48.97	-58.88
ENB	1.26	1.38	2.44	1.66	1.82	2.97	2.17	4.27
Gini(w)	0.00	0.67	0.73	0.53	0.52	0.67	0.8	0.58

Table B.15: Industry-adjusted Real Estate Ratio (RER)

COP	EW	VW	MV	NRP	ERC	MDP	HRP	DRP
Mean	23.61	19.65	17.47	21.06	23.60	30.00	21.09	18.70
Stdev	18.67	19.17	13.27	17.01	18.63	17.80	15.45	26.21
Skewness	-0.20	0.07	-0.11	-0.15	-0.20	-0.16	-0.22	0.10
Kurtosis	4.78	8.00	5.26	5.37	4.80	7.64	9.35	9.70
Sharpe	1.32	1.03	1.29	1.26	1.32	1.86	1.39	0.71
Sortino	1.70	1.38	1.75	1.65	1.70	2.35	1.82	0.94
UPR	2.97	3.03	2.98	2.97	2.97	2.82	2.87	2.90
Omega	4.10	3.19	4.07	3.93	4.11	6.18	4.50	2.26
VaR 5%	1.78	1.79	1.21	1.63	1.78	1.62	1.42	2.36
CVaR 5%	2.65	2.62	1.85	2.41	2.64	2.54	2.19	3.77
MDD	-50.22	-48.42	-46.80	-41.48	-49.10	-43.37	-40.70	-80.13
ENB	1.23	1.45	2.35	1.80	1.84	2.96	2.19	4.26
Gini(w)	0.00	0.62	0.73	0.54	0.5	0.64	0.75	0.67

Table B.16: Cash-based Operating Profitability (COP)

CLA	EW	VW	MV	NRP	ERC	MDP	HRP	DRP
Mean	24.43	21.19	17.77	21.81	24.43	29.68	21.20	27.79
Stdev	19.43	19.96	13.68	17.73	19.39	18.24	15.95	28.35
Skewness	-0.20	0.23	-0.09	-0.16	-0.20	-0.31	-0.45	0.12
Kurtosis	4.13	8.02	4.27	4.68	4.14	6.83	9.15	6.00
Sharpe	1.32	1.08	1.28	1.26	1.33	1.79	1.36	1.06
Sortino	1.69	1.45	1.72	1.64	1.70	2.24	1.75	1.37
UPR	2.99	3.04	3.01	2.98	2.99	2.85	2.86	3.00
Omega	4.07	3.37	3.98	3.91	4.09	5.82	4.33	3.35
VaR 5%	1.91	1.85	1.28	1.72	1.90	1.70	1.51	2.51
CVaR 5%	2.76	2.71	1.92	2.51	2.75	2.61	2.28	3.96
MDD	-49.94	-40.94	-41.82	-42.00	-49.11	-41.95	-39.99	-61.34
ENB	1.33	1.38	2.35	1.69	1.72	3.01	2.00	4.10
Gini(w)	0.00	0.61	0.71	0.56	0.46	0.69	0.76	0.63

Table B.17: Cash-based Operating Profits-to-lagged Assets (CLA)

EPRD	EW	VW	MV	NRP	ERC	MDP	HRP	DRP
Mean	14.14	14.96	11.67	12.50	13.84	17.51	12.83	14.75
Stdev	15.10	15.08	10.67	14.16	15.14	13.41	12.43	23.33
Skewness	-0.26	-0.09	-0.33	-0.83	-0.40	-0.31	-0.92	0.22
Kurtosis	8.71	8.69	17.62	17.79	10.30	19.16	22.32	8.74
Sharpe	0.87	0.94	0.97	0.80	0.85	1.28	0.94	0.60
Sortino	1.15	1.26	1.32	1.04	1.11	1.67	1.23	0.81
UPR	2.84	2.93	2.70	2.72	2.81	2.64	2.66	2.96
Omega	2.77	2.95	3.30	2.59	2.70	4.40	3.13	1.87
VaR 5%	1.46	1.41	0.89	1.32	1.46	1.16	1.13	2.10
CVaR 5%	2.24	2.14	1.52	2.10	2.25	1.97	1.84	3.32
MDD	-48.24	-40.24	-46.36	-53.36	-48.15	-43.47	-44.72	-72.00
ENB	1.26	1.30	2.32	1.61	1.80	3.06	2.15	4.28
Gini(w)	0.00	0.61	0.71	0.62	0.52	0.69	0.75	0.62

Table B.18: Earnings predictability (EPRD)