Charles University in Prague

Faculty of Social Sciences
Institute of Economic Studies

MASTER'S THESIS

Algorithmic fundamental trading

Author: **Bc. Vojtěch Pižl**

Supervisor: **doc. PhDr. Ladislav Krištoufek, Ph.D.**

Academic Year: **2015/2016**

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, May 11, 2016

Signature

Acknowledgments

I would like to express my special thanks of gratitude to my academic supervisor doc. PhDr. Ladislav Krištoufek, Ph.D. for his constructive and professional suggestions and comments during the writing of this thesis.

Abstract

This thesis aims to apply methods of value investing into developing field of algorithmic trading. Firstly, we investigate the effect of several fundamental variables on stock returns using the fixed effects model and portfolio approach. The results confirm that size and book-to-market ratio explain some variation in stock returns that market alone do not capture. Moreover, we observe a significant positive effect of book-to-market ratio and negative effect of size on future stock returns. Secondly, we try to utilize those variables in a trading algorithm. Using the common performance evaluation tools we test several fundamentally based strategies and discover that investing into small stocks with high book-to-market ratio beats the market in the tested period between 2009 and 2015. Although we have to be careful with conclusions as our dataset has some limitations, we believe that there is a market anomaly in the testing period which may be caused by preference of technical strategies over value investing by market participants.

JEL Classification	G11, G12, G14, G17
Keywords	Algorithmic trading, Factor models, Fundamentals, Value Investing

Author's e-mail	`vojtech.pizl@gmail.com`
Supervisor's e-mail	`kristoufek@ies-prague.org`

Abstrakt

Tato práce si klade za cíl aplikovat metody hodnotového investování do stále se rozvíjejícího pole algoritmického obchodování. V první části zkoumáme, jaký efekt mají vybrané fundamenty na budoucí výnosy z akcií za pomocí fixních efektů a také metody, která porovnává výnosnost portfolií sestavených pomocí velkosti firmy a hodnoty ukazatele účetní ku tržní hodnotě firmy. Výsledky ukazují, že zmíněné proměnné vysvětlují část variace výnosů z akcií, kterou nezachycuje vývoj celého trhu. V druhé části se snažíme aplikovat tyto výsledky do obchodního algoritmu. Za pomocí běžných vyhodnocovacích metod testujeme několik obchodních fundamentových strategií a zjišťujeme, že jednoduchý algoritmus, který vybírá malé firmy s vysokým ukazatelem účetní ku tržní hodnotě, překonává výnos tržního portfolia ve sledovaném období od roku 2009 do roku 2015. Ačkoliv musíme být opatrní s interpretací výsledků, jelikož naše data mají několik omezení, věříme, že je na trhu anomálie, způsobená nejspíše preferencí technických strategií oproti fundamentovým strategiím mezi účastníky trhu.

Klasifikace JEL	G11, G12, G14, G17
Klíčová slova	Algoritmické obchodování, Faktorové modely, Fundamenty, Hodnotové investování

E-mail autora	vojtech.pizl@gmail.com
E-mail vedoucího práce	kristoufek@ies-prague.org

Contents

List of Tables

List of Figures

Acronyms

BM Book-To-Market (Ratio)

PE Price/Earnings (Ratio)

CR Current Ratio

DP Dividend/Price (Ratio)

HFT High Frequency Trading

CAPM Capital Asset Pricing Model

CAGR Compounded Annual Growth Rate

Chapter 1

Introduction

Over the last decades, there has been a tremendous development in the field of algorithmic trading. Classic image of stock brokers on a trading floor of the New York Stock Exchange trying hard to find a counter-party for a trade is no longer accurate. In fact, most of this is now done using computer algorithms. But not only matching engines inside stock exchanges are automated, also other participants of the market, such as institutional investors, market makers, investment banks and others now in various degrees rely on computer algorithms. Principles of value investing, on the other hand, are well know to investors for decades. This investing paradigm, derived from ideas from the classic text *Security Analysis* of Graham & Dodd (1934), is based on buying underpriced stocks according to its fundamentals. Nonetheless, those methods are rarely associated with terms such as algorithmic trading. The reason being is that those are relatively complex and time demanding. Majority of literature on algorithmic trading seem to focus on various technical strategies or on high frequency trading. The former include methods such as mean-reversion, intra and inter-day momentum or pair trading, the latter is based on exploiting small inefficiencies in the market pricing using sophisticated algorithms that can evaluate situation and execute orders faster than any human.

While all those methods make sense in a context with developments in information technologies, fundamental analysis still play an important role to many types of investors. This thesis focuses on application of several fundamental factors that can be used to predict future stock value. Even though there is extensive research documenting effects of those variables on stocks prices (e.g. Fama & French (1993), Asness (2013), Lakonishok *et al.* (1994), and others), actual utilization of those methods for algorithmic trading purposes is not that common (to the author's best knowledge). We believe that fundamental analysis needs to catch up with recent developments to form the attractive alternative to technical strategies.

In this thesis we analyze the set of approximately 2000 companies traded on the New York Stock Exchange, NASDAQ and a few other small US based exchanges between 2009 and 2015. We proceed in three steps. Firstly, we analyze the data using a panel model. While this method is not used much for estimating individual companies in the reviewed literature due to the issues with heteroskedasticity and serial correlation, we try to apply it since it utilizes all information available in the data. In the second part, we use the method of Fama & French (1993). We form 25 portfolios based on size and book-to-market ratio in order to find, whether the market, size and value risk factors can explain common variation in average returns of those portfolios. This method has several advantages, such as elimination of large amount of noise present in individual stock returns or solving the issue of serial correlation. The last part attempts to utilize conclusions from the panel and the portfolio analysis into usable trading strategy. We select different strategies based on size, book-to-market ratio and other conditions. We then test those algorithms using common methods such as the Sharpe ratio or maximum drawdown. Furthermore, we add a bootstrap analysis which aim to minimize the impact of a *data snooping bias* discussed by Lakonishok *et al.* (1994), *i.e.* optimizing parameters of a trading strategy to outperform the market on available dataset, which will not work when applied somewhere else. This method randomly selects subset of stocks from the full sample of available data and tests the strategies using only this subset. Then this process is repeated 100 times. Although it does not eliminate the issue entirely, it presents further challenge to our algorithms. Last but not least, we add transaction costs in form of broker's commissions.

The thesis is structured as follows: second chapter summarizes the role of algorithmic trading on the financial markets and reviews the literature on fundamental factors and its application on predicting stock prices. Third chapter reviews methods used for panel estimation, portfolio formation and testing of trading strategies. In the fourth chapter we present and discuss results in the three steps summarized above. The last chapter concludes and outline limitations to solve in future research.

Chapter 2

Literature Review

In this section, we summarize the increasing role of algorithms in the financial markets today. Then, we discuss the most important fundamental factors that might be used for pricing equities, such as *book-to-market ratio* or *price-earnings ratio*, and review published research concerning its effect on stock returns. The last part focuses on technical trading strategies that are popular nowadays.

2.1 The Role Of Algorithmic Trading In The Financial Markets

In this section we try to summarize the role that recent developments in the field of algorithmic trading have on the markets, ass well as its implication for fundamental investing.

Kirilenko & Lo (2013) mentions three major developments in the financial industry that contributed to the rise of algorithmic trading. The first one is increasing complexity of the financial system that is thus benefiting more developed technologies. The second one is the development in quantitative modeling or so called financial engineering. Many academics and econometricians have focused on explaining financial markets behavior and thus contributed to the increasing importance of applying those methods. Last but not least, extensive developments in computer technology allowed us to perform tasks that were two decades ago unimaginable.

Before we move on explaining the role of algorithms for different market participants we describe basic model of the stock market developed by Koller *et al.* (2010). They assume that market participants can be separated into three basic groups:

1. Intrinsic investors

2. Traders

3. Mechanical investors

Intrinsic Investors engage in value investing - *i.e.* they analyze companies in depth with the aim to find the ones that are undervalued by the market and take long positions to profit by the time this undervaluation disappear. Intrinsic investors usually manage fewer positions and does not re-balance their portfolios that often as, for example, traders. According to Koller *et al.* (2010), these investors hold around 20 to 25 percent of the U.S. Equity market.

Traders, on the other hand, do not usually perform complex analysis of the stock and they do not intend to hold it for many years as intrinsic investors do. Instead, they aim to profit from short-term movements in the price caused by for example some technical factors (we will discuss this later) or news announcement related to company or industry. Koller *et al.* (2010) estimate the share of short-term traders to be around 35 to 40 percent of institutional U.S. equity.

Mechanical investors use various criteria or rules to make decisions. Examples include *Index funds* or *Quants*. The former build their portfolio by matching certain index (such as S&P 500), the latter rely on complex mathematical models. These investors control 35 to 40 percent of the U.S. equity according to Koller *et al.* (2010).

Now when we have established the framework we can explain the role of algorithmic trading for those groups. As we mentioned earlier, one of the foundations of algorithmic trading were developments in quantitative finance that started by the portfolio optimization theory of Markowitz (1952). This model that enables investors to compute optimal weights for their portfolios of assets using only information contained in historical prices could be considered, according to Kirilenko & Lo (2013), the first algorithmic trading strategy. Markowitz (1952) shows that optimal portfolio is determined as a maximization of the expected value of an objective function given means and variances of underlying assets. The strategy, in this case, is a number of shares to be bought or sold based on obtained weights. This was soon followed by the development of the Capital Asset Pricing Model (CAPM) by Sharpe (1964), Lintner (1965) and Mossin (1966).

Definition 2.1 (Capital Asset Pricing Model). The expected rate of returns of an assets is a linear function of a risk-free rate and the market return (under several simplifying assumptions).

$$E(R_i) = R_f + \beta_i[E(R_M) - R_f], \tag{2.1}$$

where $E(R_i)$ is an expected rate of return of an asset i, R_f is a risk-free rate (that can be proxied by yield of government bonds), $E(R_M)$ is an expected return of a market portfolio and β_i is sensitivity of an asset return to the markets fluctuations.

They introduced the *market portfolio* which is portfolio of all assets hold by all investor (under assumption that they hold they tangency portfolios). This portfolio is thus weighted by market capitalization of each asset.

These findings were according to Kirilenko & Lo (2013) critical milestone and served as a foundation for *Index fund* industry. These funds that constitutes large part of US stock market simply follows chosen index (such as the S&P 500) which is considered a market portfolio by selecting stocks contained in this index and weighting them by market capitalization. This passive investing strategy is according to Malkiel (1973) superior to any technical or fundamental analysis due to the *Efficient Market Hypothesis* (EMH).

The second example algorithmic trading in today's markets is an *arbitrage trading*. These strategies aim to profit from small discrepancies between prices of the same assets. For example stocks quoted on different exchanges. Although arbitrage is a very old concept, as markets became faster one require algorithms to find and execute arbitrage opportunities. By definition arbitrage is a riskless strategy since it requires two simultaneous transactions - buy for lower and sell for higher price. Nevertheless there is always some lag between those transactions thus it may even result in losses.

The arbitrage trading is one of the strategies employed by high frequency traders. *High frequency trading* (HFT) is a term that is nowadays associated with algorithmic trading. Generally we can define HFT as strategies that profit from speed. If we take as an example arbitrage, speed will allow a high frequency trader to exploit price discrepancies if he is faster than anyone else trying to do the same. Another example is trading on news announcements - high frequency traders can search for keywords in public news feeds using algorithms and evaluate and execute strategies in a milliseconds. By the time the other traders read and process the news, price will already reflect it.

So far we have talked about use of algorithms for decision-making: deciding how many shares of each stock in a portfolio should be bought or sold, finding pricing discrepancies and trading on them or identifying trading opportunities based on news announcements. Nevertheless, many market participants only use algorithms for *automated executions*. For example if the large institutional investor needs to re-balance a portfolio by buying or selling hundred of thousands shares. He will not do it by one single market order but by many smaller orders since the former could move the price of a stock in an unfavorable direction. Algorithms can help to time the orders appropriately or split them among several exchanges. Bertimas *et al.* (1999) derived such a strategy that minimize expected cost of trading of a block of shares over a fixed time horizon.

Related to automated executions is market making. Each exchange (stock, futures, etc.) has its designated market makers that help improve market liquidity by providing two sided quotes (buy and sell) and thus smoothing out temporary imbalances of supply and demand. Algorithms, in this case, help to achieve those objectives (Kirilenko & Lo 2013).

If we look at this short overview of algorithmic trading in a light of our separation of market participants into three groups (intrinsic investors, trader and mechanical investors), we see that the most affected group are mechanical investors, followed by traders. Intrinsic investors find use only in automated execution that helps them to submit larger orders without affecting the price. Nonetheless, most of the current algorithmic strategies do not apply to this style of investing. This thesis aim to fill this gap by investigation strategies based on fundamental information that would allow intrinsic or value investors to keep up with latest trends and developments.

2.2 Fundamental Factors

Value investing, a concept created by Graham & Dodd (1934), essentially consists of buying securities that seem under-priced based on fundamentals, *i.e.* the price of a security is low relative to, *e.g.* book value of assets, dividends, earnings and other measures of value. Strategies based on buying such securities should then outperform the market. Thorough this section we review existing research of these strategies and provide rationale behind them.

2.2.1 Book-To-Market Ratio

One of the most important fundamental factors is a book-to-market (BM) ratio defined as follows:

Definition 2.2 (Book-To-Market Ratio). Book-To-Market ratio relates the book value of the company to its market value and is calculated using the following formula:

$$\text{Book-To-Market Ratio} = \frac{\text{Book Value of a firm}}{\text{Market value of a firm}}, \tag{2.2}$$

where *Market Value* of a firm is obtained as the number of shares multiplied by the price of a share and *Book Value* could be found by adding up all tangible assets and subtracting all liabilities and stock issues ahead of common stock (Graham & Dodd 1934, pp. 548-549).

Over many years, academics (*e.g.*, Rosenberg *et al.* (1985), Fama & French (1992) or Lakonishok *et al.* (1994)) have documented returns to investment strategies based on buying high book-to-market firms. Fama & French (1992) was among the first authors to examine effects of this ratio on expected share price. They used regression approach of Fama & MacBeth (1973) (*i.e.* each month cross section of stock returns is regressed on several variables hypothesized to explain expected returns) and concluded that the BM ratio has positive effect on stock returns and together with size of a company capture the cross-sectional variation in stock returns. Furthermore they argue that *Beta* does not seem to have any explanatory power in explaining those returns and thus contradicting the asset pricing model of Sharpe (1964), Lintner (1965) and Black *et al.* (1972) (SLB model) that uses only *Beta* to estimate expected cross-section of stock returns. Their model follows up with other empirical contradictions with the SLB model such as the size effect documented by Banz (1981), the positive effect of company's leverage on average return observed by Bhandari (1988) and finally with preceding research on the BM ratio by Stattman (1980) and Rosenberg *et al.* (1985). The reason behind this effect is related to risk. Fama & French (1992) assumes two dimensions of stock risk: the first dimension is proxied by size and the second one by the Book-To-Market ratio. In other words, smaller companies with lower market value relative to its book value of assets bear higher risk and investors who choose to have those stocks in their portfolios are compensated for it.

In their other research, Fama & French (1993) extends the asset pricing model developed in Fama & French (1992) by including also U.S. government and corporate bonds (since the model should also explain the bond returns if the markets are integrated), expanding set of variables (changes in interest rate, shifts in economic conditions) and using time-series regression approach of Black *et al.* (1972) instead of cross-section regressions of Fama & MacBeth (1973) (we discuss details in the methodology section). Again, the results support the theory that size and BM ratio capture common variation in returns. In total, their model identifies five risk factors: in addition to size and BM ratio there is also an overall market factor (excess market return) and two bond market factors mentioned above. This model then explains average returns on stocks and bonds.

The result of the research by Fama & French (1992) and Fama & French (1993) is the well known *Fama-French Three Factor Model* defined as:

Definition 2.3 (Fama-French Three Factor Model).

$$R(t) - R_f(t) = \alpha + \beta[R_M(t) - R_f(t)] + sSMB(t) + hHML(t) + e(t), \quad (2.3)$$

where $R(t) - R_f(t)$ is the excess return, $R_M(t) - R_f(t)$ is the excess market return, SMB (small minus big) is the size factor (computed as the difference between small and big stock portfolios with approximately same BM ratio) and HML (high minus low) is the value factor (obtained as the difference between high and low BM portfolios of approximately same size) (Fama & French 1993).

Or alternatively *Five Factor Model* that also includes additional bond market variables introduced in (Fama & French 1993):

Definition 2.4 (Fama-French Five Factor Model).

$$\begin{aligned} R(t) - R_f(t) = \alpha + \beta[R_M(t) - R_f(t)] + sSMB(t) + hHML(t) \\ + mTERM(t) + dDEF(t) + e(t), \end{aligned} \quad (2.4)$$

where the additional variables are $TERM(t)$ that capture the effect of the interest rate changes (computed as the difference between long term government bonds and one month T-Bills) and $DEF(t)$ which serves as a proxy for shifts in economic conditions (difference between return on a market portfolio and long term government bonds) (Fama & French 1993).

Lakonishok *et al.* (1994) provides different explanation why value strategies produce superior returns. They believe that they are contrarian to "naive" strategies used by other investors, such as too optimistic extrapolation of past growth of earnings, overreaction to some good or bad news, etc. In other words, these "naive" investors tend to buy stocks that have performed well in the past and sell stocks after the significant price declines. Contrarian investors then bet against them. Contrary to Fama & French (1992), they do not believe that these strategies are fundamentally riskier, *i.e.* higher stock returns are compensation for the higher risk that value investors bear.

There are two types of stocks in the contarian model of Lakonishok *et al.* (1994): *glamour stocks* are overpriced stocks that have performed well in the past and are also expected to perform well in the future. Underpriced *value stocks*, on the other hand, have not performed well in the past and the expectations remains the same. The results shows that value strategies outperformed the strategies that involved buying glamour stocks (over the test period from 1968 to 1990). The reason being is that the actual future growth rate is lower compared to the past. In other words, those future growth rates were overestimated by market participants relative to growth rates of value stocks. Lastly, value strategies does not seem to be fundamentally riskier than then glamour strategies, which is in contrary with Fama & French (1992).

Lakonishok *et al.* (1994) also discuss following reasons for low BM ratio:

1. High proportion of intangible assets (such as R&D)
2. Growth opportunities that are not reflected in book value but rather in price
3. High temporary profits that increase stock price but not its book value
4. Overvalued glamour stock

Hence while the returns observed for high BM stocks tend to be higher, this variable cannot be understood as a synonym for overvalued glamour stocks. Nevertheless, there is a return anomaly and the best explanation for it is according to Lakonishok *et al.* (1994) the preference of glamour stocks over value stocks by both institutional and individual investors. Why is this a case? As we already discussed, investors tend to put much weight on past performance and project it into future. But this is not sole reason for it. Institutional investors might select glamour stocks simply because it appear safer and it is easily justifiable to investors in those institutions (Lakonishok *et al.* 1992). Another reason is the short time horizon for many investors (Shleifer & Vishny 1990). They seek stocks that will earn abnormal returns within few days/months but value strategies require much longer time horizons (sometimes many years). This is problem for money managers that have to show results each quarter/year. Moreover, the selection of glamour stocks by money managers could explain their poor performance compared to market index as observed by Lakonishok *et al.* (1992).

Piotroski (2000) shows that investors could shift the distribution of returns earned by their portfolios by selecting high BM firms (by 7.5% at least). Moreover, extended strategy that in addition to buying BM strong companies also short-sell those expected to lose results in 23% annual return (between 1976 and 1996). The strategy seems to be robust across time and also controls for other investment strategies. Piotroski (2000) provides the explanation for this that is again inconsistent with Fama & French (1992). Instead of assuming that those firms are fundamentally riskier, they use the intuition behind the "life cycle hypothesis" of Lee & Swaminathan (2000). They argue that firms with poor past performance tend to be subject to pessimism and investor neglect. In average, these late-stage momentum losers recover and become winners.

In the more recent research, Asness (2013) introduced the model that incorporate both the "value" and the "momentum" effects. Where the former is the relation between stock returns and its Book-To-Market ratio (discussed above), the latter represents the relation between stock returns and its recent performance. They find consistent and omnipresent return premiums across eight different markets and several asset classes.

Fama & French (2012) also propose model that captures momentum patterns observed in academic literature using the *Three Factor Model* as a building block. They again find value premiums is all examined regions (Europe, North America, Japan and Asia Pacific) and also significant momentum premiums in average returns in those regions (except Japan). Moreover, they observed higher premiums for smaller cap stocks than for larger cap stocks and thus confirmed the size effect of previous research by Fama & French (1992) and Fama & French (1993).

To estimate expected stock returns, Lyle & Wang (2014) are using book-to-market ratio and firm's return on equity (ROE). They argue that expected holding period returns are time-varying and hence investors who allocate their capital over different time horizons must take this into account. The factor models of Fama & French (1992) and Fama & French (1992) on the other hand are assuming constant expected asset returns which can lead to pricing errors and poor capital allocation. Using a cross-sectional regression they predicted expected returns from 3 months to 36 months with statistically significant coefficients (at 99% level). The trading strategy involving buying top decile (going long) and short-selling the bottom decile of 3, 12, 24 and 36 months expected returns results in abnormal future returns of 5.48%, 15.42%, 30.44% and 46.32%, respectively.

2.2.2 Price-Earnings Ratio

Another important ratio is Price-Earnings Ratio (P/E):

Definition 2.5 (Price-Earnings Ratio). P/E Ratio relates stock price of a company to its per share earnings.

$$\text{Price-Earnings Ratio} = \frac{\text{Market value per share}}{\text{Earnings per share}}, \tag{2.5}$$

where *Market value per share* is simply stock price and *Earnings per share* or EPS is portion of company's net income allocated to each share outstanding and could be obtained as follows:

$$\text{EPS} = \frac{\text{Net Income} - \text{Dividends on preferred stock}}{\text{Average number of shares outstanding}}. \tag{2.6}$$

Sometimes P/E ratio could be referred to as the *Earnings Multiple* since it shows how much investors pay for each dollar of net income. [1]

P/E ratio was used to predict stock prices by Basu (1975) and Basu (1983). In his research he found that there is a significant negative effect of P/E ratio on the returns of NYSE stocks during 1963-80 period while controlling for firm's size (in the paper, the effect is positive but Basu (1983) is using E/P instead P/E and thus the implications are inverse). On the other hand, it appears that this *earnings yield* is in inverse relation with the size of a firm since the results are not significant for larger than average NYSE firms. Therefore, since the effect of E/P is dependent on the size, its effect on expected returns is more complicated and is most likely just proxy for more fundamental factors.

Ball (1978) argues that P/E ratio is just a proxy for other factors affecting expected returns since with higher price (and thus higher P/E) risk also increases. This argument could be also applied to BM ratio which can be also

[1] Source: http://www.investopedia.com/terms/p/price-earningsratio.asp

viewed as a scaled stock price (market value is calculated using share price). Fama & French (1992) is acknowledging the argument of Ball (1978) that P/E, BM and also leverage might be regarded as scaled stock prices and are thus only extracting the information about risk. Hence some of them might be redundant when explaining expected returns and therefore Fama & French (1992) are testing the effects of those variable (together with β) jointly. Whilst the usage of P/E alone in the cross-section regression of Fama & MacBeth (1973) results in negative relationship as in studies by Basu (1983), when added to the full regression with size and BM, the size of the coefficient changes significantly (from- 4.72% to -0.87%). This sugges negative correlation of P/E and BM (note: Fama & French (1992) is using E/P instead of P/E therefore we inverted all effects) and this variable is therefore redundant in the model as BM and size seem to explain cross sectional variation in expected stock returns.

Altogether, studies on P/E ratio do not present such conclusive effect on expected stock returns. Nevertheless, we will still test its effects in this thesis together with BM ratio to test the hypothesis of Ball (1978).

2.2.3 Applications of Fundamental Factors in Trading

While the research on the fundamental factors reviewed in the previous section is extensive, the implications for actual use of those methods for stock selection is not that clear. Fama & French (1992) argues that usage of the size and BM ratio for investors who seek long-term returns depends on two assumptions:

1. Persistence, *i.e.* explanatory power of BM ratio is not deteriorating in time. Fama & French (1992) confirms this for the period cover in the research (from 1963 to 1990), but now we have additional 25 years that has not been tested yet.

2. Rationality of asset pricing. If the stock prices are irrational, as it can happen for example during crisis, the results would be jeopardized.

In this thesis, we assess performance of strategies based on findings of Fama & French (1992) and other authors discussed in preceding sections on the more recent data. Moreover, we employ additional performance measures (discussed in the methodology section).

Natural question arises when concerning abnormal returns to fundamental strategies observed in academic literature: how it could persisted for so long? Why, if the markets are efficient, this anomaly has not disappeared yet? Lakonishok *et al.* (1994) offers one explanation: in time of publication of those researches, most investors did not know about these anomalies. Even though the concepts of value investing dates back to Graham & Dodd (1934), there was simply no statistical analysis that would confirm it. While this reason could be used in 90s, today it remains questionable.

Another possible explanation also discussed by Lakonishok *et al.* (1994) is a *data snooping bias*. If we create a strategy with many parameters and then optimize those parameters, it is very likely that result would be fantastic. On the other hand, if we apply that strategy to future (not yet observed) data, or to completely different time series, there is very high probability that the performance will be nothing like it was historically. The more complex model, the higher probability of data snooping (Chan 2009, pp. 25-26). Although the models of Fama & French (1993) and Lakonishok *et al.* (1994) are quite simple, they are both built on similar datasets. Nevertheless, there are similar patterns observed on different datasets. For example Chan *et al.* (1991) observe similar anomalies Japan and Capaul *et al.* (1993) focus on France, Switzerland, Germany and the United Kingdom with similar results.

Survivorship bias is another thing to be aware of. Since most of the fundamentally based strategies have to rely on some database of historical data, there is a possibility that this database excludes stocks that disappeared due to bankruptcies, mergers & acquisitions or de-listings. Therefore only companies that survived until the time when a dataset was downloaded will be considered in a backtest. If we consider for example strategy involving buying "cheap" stocks (*e.g.* those with high BM ratio), there is a higher probability that such stocks could be about to bankrupt. However, if those stocks are omitted from a dataset, our strategy will select only those that survived and increased in value. This will result in abnormal performance but ignore the substantial part of a risk. Needless to say, such strategy would never succeed during real trading (Chan 2009, p. 24).

2.3 Technical Trading Strategies

As we discussed in our review of algorithmic trading. largest proportion of strategies is created on technical rather than fundamental basis. Whilst it is not the aim of this thesis to review and test those strategies, we will use some basic ones as a benchmark. Therefore it is worth summarizing basic types of these strategies.

Generally, we can split trading strategies into two basic categories: *mean reverting* and *momentum* strategies.

Mean reverting strategies are justified by concepts of stationarity and cointegration. In other words, stock prices tend to revert to some mean value. Such strategies then aim to buy stocks when the price is below mean value and sell it when it is above the mean. While the most of the financial time series are not stationary or mean reverting, we can construct portfolios of individual non-stationary time series that would create synthetic stationary time series. Such time series are called *cointegrated*. The easiest example is so called “pairs trading”, where we have long position in one asset and short position in another, but it could be extended to more than two assets (Chan 2013, pp. 39-51).

Momentum strategies, on the other hand, rely on some trend in stock price that is expected to persist (they are said to have “momentum”). This momentum can be caused by reaction to some news announcement (earnings, mergers & acquisitions, etc.) which is now easy to implement in code due to existence of machine-readable news feeds. Similarly these strategies can follow market sentiment and trade on it. Momentum strategies can be further split between *time series momentum* and *cross sectional momentum*, where the former is assuming positive correlation of a past return to future returns, the latter relates time series of one security to other time series (example: if one series outperformed the other, it is likely that it will keep doing so in the future) (Chan 2013, pp. 133-149).

There are other types of technical trading strategies, such as *Regime switching*, that is attempting to identify “turning point” between two regimes (such as bull and bear market) and trade on it. *Seasonal trading strategies* are based on buying and selling securities on certain days. Last but not least, there are many *High Frequency Strategies* that use speed to profit from certain anomalies and arbitrage opportunities. Whilst the importance of HFT in the financial markets is growing, it is beyond the scope of this thesis.

Chapter 3

Methodology

Most of the reviewed studies that are trying to explain and predict stock returns rely on portfolio construction. Stocks are sorted by one or more characteristics (like size or book-to-market ratio) and allocated into a small number of portfolios. Then, the returns of those portfolios are compared and if there is a significant difference, underlying characteristics have predictive power. This approach has certain advantage - it smooths out outliers and make variables less noisy. Returns and values of fundamental characteristics of individual firms tend to be very noisy. Especially ratios that can have zero (or a number close to zero) in denominator (such as PE ratio) can increase to large values that can distort the model. Nevertheless, by forming portfolios we are losing significant information in the data, which is the reason why we first try to estimate our data in full scope using panel models. Then we compare results (estimated parameters) with portfolio model based on the methodology of Fama & French (1993).

After identification of parameters that have some predictive power, we form our trading algorithm. We aim to answer the question if the result from the estimated models can be utilized in practice. We evaluate performance of each strategy using common indicators such as the Sharpe ratio and maximum drawdown. Last but not least, we incorporate transactional costs to see how it affects profitability of each strategy.

3.1 Panel Data Model Specification

We are interested in predicting stock returns R_{it} of individual companies using set (vector) of characteristics X'_{it}. General model for such prediction is the one-way error component model (Badi H. Baltagi 2005, p. 11) defined as:

$$R_{it} = \alpha + X'_{it}\beta + \mu_i + \epsilon_{it} \quad i = 1, \dots, N; \quad t = 1, \dots, T, \tag{3.1}$$

where β is the vector of estimated parameters, μ_i is the company specific effect and ϵ_{it} is the error term. We assume that errors have mean zero and are uncorrelated with the regressors, i.e. $E[\epsilon_{it} x_{it}] = 0$.

The easiest is to estimate 3.1 using pooled OLS model, which ignores the panel specification of the dataset. We can add time dimension using dummy variables but this is infeasible for many time periods. Furthermore, there is an issue with the company specific effect μ_i which has to be uncorrelated with explanatory variables otherwise the estimates are not consistent. There are most probably some specific, time invariant characteristics for each company that are correlated with its BEME ratio that we do not control for. Therefore, we have to deal with this unobserved heterogeneity. We estimate pooled OLS only to be able to perform tests for individual effects.

One possible solution for unobserved heterogeneity problem is to use fixed effect model, that is to subtract mean from each variable in 3.1. This step eliminates individual effects since these are time invariant, however, we are losing the possibility to estimate model with dummy variables such as company sector or industry.

3.1.1 Fixed Effects

One possibility is to a add vector of individual dummies and perform ordinary least squares on 3.1, but with large N the matrix to be inverted becomes large and the estimation is not feasible. Better option is demeaning variables using within transformation, that is to subtract averages over time from 3.1 (Badi H. Baltagi 2005, p. 13):

$$R_{it} - \bar{R}_i = \beta(X'_{it} - \bar{X}'_i) + (\epsilon_{it} - \bar{\epsilon}_i). \tag{3.2}$$

This step eliminates the unobserved heterogeneity and we can estimate 3.2 using the OLS.

3.1.2 First Differences

Another method that allow us to get rid of the unobserved heterogeneity is the first difference model (Greene 2012, p. 395):

$$\Delta R_{it} = \beta \Delta X'_{it} + \Delta \mu_i + \Delta \epsilon_{it}, \tag{3.3}$$

where ΔR_{it} is first difference of stock returns. We can again estimate 3.3 using OLS since $\Delta \mu_i = 0$. This method also removes any other time invariant variables (such as sector dummy variables in our case).

To decide if the fixed effect model (or first difference model) is really preferable to the pooled OLS (i.e. there are individual effects) we use the Lagrange multiplier test of T. S. Breusch and A. R. Pagan (1980), with the null hypothesis that there are no significant effects.

Since our dataset contains large number of companies (cross sectional dimension is large), we need to address the issue of heteroskedasticity and serial correlation. Therefore we use robust covariance matrix of M. Arellano (1987). This way we have heteroskedasticity and serial correlation robust standard errors.

3.2 Portfolio Formation And Stock Market Factors

Academic literature trying to explain variation in stock returns often starts by forming portfolios based on single or multiple characteristic. We employ the method of Fama & French (1993) to have another, more robust evidence of predictive power of stock's fundamentals.

Stocks are sorted each month based on size and book-to-market ratio. For size we use NYSE breakpoints (similarly as in Fama & French (1993)) because we do not want our top size quantiles to be disproportionately large compared to bottom size quantiles in terms of percentage of total market value (NYSE stocks are larger in average). For book-to-market ratio we decided to compute breakpoints individually for each sector following the evidence that BM ratio varies significantly across sectors (table 4.3). Especially for financial companies this ratio is much bigger in average. Fama & French (1993) deals with this by excluding financial companies from the dataset entirely, but we do not have a big enough dataset to do such rapid exclusions (there are 442 stocks classified as financial from 2002). Using the obtained breakpoints we form 25 portfolios based on size and BEME quantiles intersections (5×5) that serves as a dependent returns in a factor regression.

For the factors calculation we again follow Fama & French (1993) approach. Stocks are sorted into 6 portfolios based on size and book-to-market ratio. For size we have 2 groups - small and big (as a breakpoint we use NYSE median value each time period) and for BM ratio we construct 3 groups - high (top 30%), medium (middle 40%) and low (bottom 30%). Again BM sorts are sector neutral.

From the intersection of 2 size and 3 book-to-market groups we construct 6 portfolios - S/L, S/M, S/H, B/L, B/M, B/H. For example, the S/L portfolios contains small stocks with low BEME ratio, the B/M portfolios contain big stocks with medium BEME ratio and so on. Portfolio returns are value weighted, where the weight is a stock price.

Next, we define 3 risk factors related with market, size and book-to-market ratio similarly as Fama & French (1993).

RM. Market risk factor is obtained as a total return for all stocks in our dataset.

SMB. Size risk factor SMB (small minus big) is calculated, as the name suggest, by subtracting average return on 3 big stock portfolios (B/L, B/M, B/H) from average return on 3 small stock portfolios (S/L, S/M, S/H). Due to its construction, this difference isolates the size effect.

HML. Book-to-market risk factor HML (high minus low), on the other hand, isolates the BEME effect by subtracting the average of the returns on 2 low BEME portfolios (S/L, B/L) from the average of the returns on 2 high BEME portfolios (S/H, B/H).

Using the obtained risk factors and the 25 stock portfolios formed on size and book-to-market ratio we then estimate following models using time series regression.

Market model. Using only market risk factor we assess how much of the variation in returns is explained by market.

$$R(t) = \alpha + \beta R_M(t) \tag{3.4}$$

3 factor model. Based on the Fama & French (1993) three factor model (only simplification is not subtracting risk free return).

$$R(t) = \alpha + \beta R_M(t) + sSMB(t) + hHML(t) \tag{3.5}$$

Fundamental model. Not using market return as this information is not useful for stock predictions.

$$R(t) = \alpha + sSMB(t) + hHML(t) \tag{3.6}$$

We estimate above models for all 25 portfolios and compare regression's coefficients and r-squared.

3.3 Fundamental Trading Strategy

The added value of this thesis is applying gained knowledge to an actual trading strategy. As a starting point, we use 25 portfolios formed on size and book-to-market ratio defined in previous section. Our focus is on the size and the value effect, therefore we base our strategies on those two indicators. Furthermore, we add other variables to enhance the performance.

Before we describe tools that are used for evaluation, lets define portfolio returns.

Definition 3.1 (Portfolio Returns). Having a portfolio of n risky assets, we obtain portfolio return as follows:

$$R_p = \boldsymbol{w'R} = \begin{pmatrix} w_1 & w_2 & \cdots & w_n \end{pmatrix} \times \begin{pmatrix} R_1 \\ R_2 \\ \vdots \\ R_n \end{pmatrix}, \tag{3.7}$$

where $\boldsymbol{w}$ is a vector of weights (such as stock price for value weighted returns and $1/n$ for equal weighted returns) and $\boldsymbol{R}$ is a vector of individual stock returns.

To evaluate portfolio performance, we use the Sharpe ratio, that adjusts portfolio returns for risk, where the risk is standard deviation of the portfolio returns. Of course, portfolios are rebalanced on monthly or yearly basis (depending on the strategy), hence we have to recalculate the ratio after each such change. Moreover, using monthly returns is not possible since we need to calculate variance-covariance matrix to obtain portfolio variance, therefore we use daily returns.

Definition 3.2 (Portfolio Variance). Variance of a portfolio of n risky assets is calculated using the following formula:

$$\sigma_p^2 = \boldsymbol{w'\Sigma w} = \begin{pmatrix} w_1 & w_2 & \cdots & w_n \end{pmatrix} \times \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_n^2 \end{pmatrix} \times \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix}, \quad (3.8)$$

where $\boldsymbol{w}$ is a vector of weights and Σ is variance-covariance matrix of n stocks.

Definition 3.3 (Sharpe Ratio). Ex-post Sharpe Ratio is obtained as:

$$S = \frac{R_p - R_f}{\sigma_p}, \quad (3.9)$$

where R_p is realized portfolio return and σ_p is portfolio standard deviation. R_f is a risk free rate, or alternatively other benchmark such as the S&P 500 index. [1]

Other useful indicator for evaluating portfolio performance that is focused on downsides is the maximum drawdown defined below.

Definition 3.4 (Maximum Drawdown). Downside risk indicator, that is obtained as a maximum loss from a peak to a trough of a portfolio before a new peak is reached. Depth of the drawdown is measured in percentage value. Apart from depth, we will observe length of a drawdown period, length of a declining period (from a peak to a trough) and recovery period (from a trough to a new peak). [2]

Next method that we use to evaluate strategy performance is a *bootstrapping*. The reason we employ it is the *data snooping bias* mentioned earlier. We want to see how our strategies behave on different dataset. Since we unfortunately do not have any other set of companies, we reduce total number of stocks to 1000 (randomly chosen) and test each strategy's performance. Then we repeat the experiment 100 times with different, randomly chosen, subset of data and record an average return.

[1] Source: http://www.investopedia.com/terms/s/sharperatio.asp
[2] Source: http://www.investopedia.com/terms/m/maximum-drawdown-mdd.asp

Chapter 4

Data

In this chapter we describe how the data were obtained and variables constructed. Furthermore, we provide preliminary analysis and predictability tests to better understand the dataset. Last but not least, we highlight some drawbacks of our data.

4.1 Dataset

A long history of stock fundamentals is crucial for testing any value based strategy. Unfortunately, it is difficult to obtain such data without a significant financial investment. Therefore, the historical period used in this thesis is not as long as in *e.g.* Fama & French (1993). Nevertheless, we believe that this period is sufficient for testing of our hypothesis.

For fundamentals we use the "Free US Fundamental Data" from the Quandl database. This dataset provides annual fundamental data for more than 2000 companies up to 9 years of history. After deletion of a companies with no information about book value or number of shares outstanding, we are left with 2002 stocks.

Monthly stock prices were obtained from yahoo finance and merged with Quandl database. Financial rations are calculated using actual monthly stock price with the latest available fundamental information. Thus the ratios are changing each month but the underlying fundamentals only once a year.

Based on already published empirical studies we have identified several indicators (company characteristics) that should help to predict stock prices. We divide them into following groups:

Size A company size is measured as market capitalization (number of shares outstanding multiplied by share price) of company i at time t.

Valuation ratios Several ratios calculated from stock's fundamentals. Including BM ratio (book value of equity to market value of equity), PE ratio (price divided by per share earnings), DP ratio (dividend yield) and current ratio (current assets divided by current liabilities).

Market return Under the CAPM, expected return of an asset is a linear function of a risk free rate and a market return. As the market return we use value weighted portfolio of all stocks in our dataset.

Market momentum Cumulative returns of the last 1, 3, 6 and 12 months.

As a dependent variable (stock returns that we are trying to explain) we use holding period returns, where the holding period is 1, 3, 6 and 12 months (it is documented that momentum variables better explain shorter holding period returns whereas value variables best estimate longer periods).

4.1.1 Size

Size of a company is hypothesized to have negative effect on expected stock returns. This was documented by e.g. Fama & French (1992) and Fama & French (1993). Therefore, we will include this variable in our model. We measure *Size* as market capitalization of a company, which is calculated as:

$$Size = Number\ of\ shares\ outstanding \times Price\ of\ a\ share. \tag{4.1}$$

We can find descriptive statistics in the Table 4.1. Mean value is 8.5 billion USD. The maximum value of 782 billion corresponds to Apple Inc in May 2015.

4.1.2 Book-To-Market Ratio

Book-to-Market ratio (BM) is obtained as:

$$BM = \frac{Book\ value\ of\ equity}{Market\ value\ of\ equity} = \frac{Book\ value\ of\ equity}{Size} \tag{4.2}$$

Again, summary statistics for BM is in the Table 4.1. It is obvious that there are some outliers since min and max values of -176 and 204, respectively, are far from 1st and 3rd quintiles. Further analysis reveal that 61 companies have average BM ratio less than zero and for 163 companies this ratio decreased below zero at least once. On the other hand, there is only 9 companies with minimum BM ratio less than -10 and 50 companies with minimum below -1. Regarding the opposite extremes, for 47 companies, BM ratio maxed over the value of 10, nevertheless only 2 companies achieved to have mean value larger than 10.

4.1.3 Price-Earnings Ratio

PE ratio is calculated using the following formula:

$$PE = \frac{Price}{Earnings\ Per\ Share} \tag{4.3}$$

The outliers for PE ratio are even more extreme. Concerning the average PE ratio, 35 companies have it less than -100 and 59 companies over 100. 29 stocks reached maximum value over 1000 and PE of 29 companies have its minimum below -1000. Furthermore we have 380 NA values for PE variable.

4.1.4 Dividend-Price Ratio

DP ratio, or *dividend yield*, indicates how much a company pays to its investors each year relative to price of its shares. It is calculated as follows:

$$DP = \frac{Annual\ Dividend\ Per\ Share}{Price\ Per\ Share} \tag{4.4}$$

The big difference between *mean* (2.2%) and *median* (0.6%) indicates that many of companies did not pay any dividends (Table 4.1).

4.1.5 Current Ratio

Current ratio (CR) is a liquidity ratio indicating an ability of the company to pay its obligations.

$$CR = \frac{Current\ Assets}{Current\ Liabilities} \tag{4.5}$$

Unfortunately, we have 30 967 NA values for CR (not available for 427 companies).

4.1.6 Return momentum

Cumulative return variables *Ret1*, *Ret3*, *Ret6* and *Ret12* are calculated as:

$$Retx = \frac{Price_t}{Price_{t-x}} - 1 \quad \text{where} \quad x \in \{1, 3, 6, 12\} \tag{4.6}$$

From Table 4.1 we can observe that average return for 1, 3, 6 and 12 month holding period is 1.2%, 3.5%, 7.4% and 16.6%, respectively. Minimum values then reach almost -100%, which is an equivalent of loosing all the value. Maximum return (12 month) of 15000% was observed for Kingold Jewelery, whose price increased from 0.05 USD in Sep 2009 to 8.49 USD in Sep 2010. Apart from this extreme observation, only 11 companies have maximum 12 month holding period return greater than 1000% and 13 have mean return greater than 100%. For the return momentum we observe some NA values. This is caused by nonexistent stock history for new symbols and thus missing those values for 3, 6 or 12 months returns.

Table 4.1: Descriptive statistics - Explanatory variables

Variable	Min	1st Qu.	Median	Mean	3rd Qu.	Max	SD
Size	0.1	296.5	1 822	8 465	6 133	781 700	25 696
BEME	-176.5	0.264	0.502	0.700	0.896	204.5	1.658
PE	-4374	2.899	14.250	14.140	22.950	3 949	124.8
DP	0.000	0.000	0.006	0.022	0.026	11.220	0.089
CR	0.000	1.279	1.909	2.746	2.974	100.5	3.581
Ret1	-0.934	-0.048	0.009	0.012	0.062	16.600	0.146
Ret3	-0.984	-0.073	0.026	0.035	0.122	18.670	0.251
Ret6	-0.999	-0.092	0.052	0.074	0.193	30.670	0.382
Ret12	-0.999	-0.090	0.113	0.166	0.323	150.000	0.820

Source: author's computations.

4.1.7 Holding Period Returns

Our dependent variables, holding period (HP) returns for 1, 3, 6 and 12 months, are obtained similarly as return momentum, that is:

$$Returnx = \frac{Price_{t+x}}{Price_t} - 1 \quad \text{where} \quad x \in \{1, 3, 6, 12\} \tag{4.7}$$

We do not report summary statistics as those are the same as for return momentum variables in Table 4.1 (momentum returns are lagged HP returns).

4.2 Preliminary Analysis

Below is the summary by years (Table 4.2). We can see that we don't have many companies before 2009. But since 2009 number of companies is increasing from 1550 to 2002 in 2015.

Average market capitalization drops significantly in 2008 and 2009. This is the result of the financial crisis and increased number of companies in the dataset in 2009. However, this number is increasing since 2002 to 2015.

Average BM ratio decreased significantly from 2009 to 2011 (from 1.07 to 0.7), which is probably due to increasing market capitalization of all companies.

Similarly, we can observe effects of post crisis development on average PE ratio, which drops from 13.9 in 2008 to 10.3 in 2010.

Table 4.2: Summary by years

		Mean					
Year	Count	Size	BM	PE	DP	CR	Return
2007	219	13 411	0.40	10.81	0.011	2.26	-0.41
2008	577	7 919	0.74	13.89	0.019	2.36	-2.97
2009	1550	4 033	1.07	12.16	0.034	2.77	5.41
2010	1660	6 141	0.81	10.27	0.024	2.63	2.37
2011	1721	7 080	0.70	14.49	0.022	2.67	0.00
2012	1799	7 610	0.72	14.82	0.020	2.74	1.63
2013	1880	9 046	0.59	13.99	0.022	2.69	3.07
2014	1983	10 501	0.54	17.11	0.019	2.74	0.74
2015	2002	10 744	0.66	14.80	0.022	2.99	-0.45
2016	2000	10 208	0.96	13.42	0.028	3.05	-0.57

Source: author's computations.

Table 4.3 summarizes stocks by sector. We can see that largest group in our dataset is a financial sector and smallest are utilities. Largest average market size is observed for a technology sector and lowest for utilities. Companies in the financial sector have the highest BM ratio by significant margin to other sectors, which has something to do with measuring of book value of equity for those companies. For a healthcare sector, we observe the lowest PE ratio and the highest for consumer goods. One month holding period return is the largest for healthcare companies and lowest for basic materials.

It is clear from the Table 4.3 that sectors play some role in predicting returns as well as that there is some relationship between sectors and fundamental ratios.

Table 4.3: Summary by sector

Sector	Count	Size	BM	PE	DP	CR	Return
				Mean			
Basic Materials	202	9 063	0.76	14.22	0.029	2.58	0.544
Consumer Goods	147	10 963	0.50	17.86	0.018	2.25	1.501
Financial	442	6 465	1.11	14.95	0.044	2.55	1.102
Healthcare	243	8 496	0.43	5.55	0.005	4.83	1.707
Industrial Goods	196	7 014	0.64	12.48	0.017	2.48	1.208
Services	376	8 234	0.54	16.79	0.015	1.97	1.396
Technology	317	11 195	0.60	16.22	0.013	3.15	1.296
Utilities	74	6 708	0.79	8.29	0.035	1.32	1.082

Source: author's computations.

Similarly, Table 4.4 offers summary by exchange, where the stock is traded. It is not surprising that most stocks are traded on NYSE, followed by NASDAQ. Apart from those two large exchanges we have 64 companies from NYSE-MKT (former AMEX) and 3 companies from NYSE-ARCA. Finally, 8 stocks were delisted. Since the majority of stocks is either on NYSE or NASDAQ, we compare only those two (summary statistics is not very representative in a small sample).

NYSE stocks are much larger in average (11 billion compared to 4.6 billion), which is caused by many small technology companies that are traded on NASDAQ. Of 100 largest companies in our dataset (measured by average size over the whole period), 78 are from NYSE and 22 from NASDAQ. Interestingly, the biggest 3 stocks (Apple, Google and Microsoft) are traded on NASDAQ. In terms of book-to-market ratio both stock exchanges have similar average value (for NASDAQ it is 0.03 higher). In contrast, PE ratio is higher for NYSE stocks (17.5 vs 10), as well as dividend yield (many technology firms traded on NASDAQ do not pay any dividends). From the comparison of average one month holding period returns, we can see that NASDAQ stocks performed generally better.

Before we move on the next section, we should comment on drawbacks of our dataset. The major one is a limited number of delisted companies. From Table 4.4 we can see that there are only 8 and all were delisted in 2016, which probably means that this dataset started to include those only after the end 2015. This unfortunately brings some bias to our estimation models and trading simulations perform better than they would perform with all stocks available

Table 4.4: Summary by exchange

				Mean			
Exchange	Count	Size	BEME	PE	DP	CR	Return1
NASDAQ	867	4.62	0.70	10.22	0.02	4.15	1.22
NYSE	1060	11.02	0.67	17.55	0.03	1.95	0.94
NYSE-MKT	64	0.31	0.97	6.84	0.02	4.30	1.29
NYSE-ARCA	3	0.55	1.12	9.05	0.00	NaN	-0.37
Delisted	8	7.13	0.66	31.22	0.02	2.25	1.45

Source: author's computations.

at the time of selection. Another limitation is a short time period - majority of stocks do not have fundamentals available prior to 2009. This also impacts our trading models as we cannot observe long run effects in different market conditions. Lastly, even though our cross sectional dimension seems high (over 2000 stocks), we are missing some big players (*e.g.* large oil companies such as Exxon Mobil and British Petrol).

4.2.1 Predictability tests

Before we estimate our model, we test our dataset for predictability by single characteristic (that is each of our explanatory variables defined earlier in this section).

Each month we split companies in two portfolios based on deviation from median value. We have 2 portfolios for each explanatory variable: *low*, which contains all companies that have value of a given variable below median and *high*, that includes companies with value of a given variable above median.

Then we calculated mean 3, 6 and 12 month holding period return for each portfolio and performed t-test for group means. Table A.1 shows 3 panels which correspond to results for 3, 6 and 12 month holding period returns. First two columns shows average value of *low* and *high* portfolios for each variable. Next two columns displays percentage returns for those portfolios. Last two columns give us t-statistic and p-value of performed t-test.

Lets start with *Size*. The *Low* portfolio have mean market capitalization of 544 million and the *high* portfolio 16.2 billion. Returns to the *low* portfolio are in average higher than for the *high* portfolio. This difference is statistically significant for all holding period returns. This result suggest that there is negative relationship between *size* and *returns*.

For *BM* ratio, the results are opposite from what we would expect. The *low* portfolio that have average MB of 0.2 correspond to higher return than the *high* portfolio with average BM of 1.18 for all holding periods (however only for 6 and 12 month this difference is statistically significant at 10% level).

The effect of *PE* ratio is negative as we would expect. Lower PE results in higher returns for all holding periods (difference is significant at 1% level). *DP* ratio suggests that companies that pay lower or no dividend have higher returns (again significant at 1% for all holding periods) and similarly for *CR* ratio, where lower ratio (that means higher leverage) corresponds to higher return (significant at 5%, 1% and 1% for 3, 6 and 12 month holding periods).

The results for momentum variable are not significant for 3 and 6 month holding periods, but for 12 month period there is a significant positive effect.

One reason why the effect of *BM* ratio is opposite than it was observed in an empirical literature could be due to some heterogeneity in the data. In Table 4.3 we show that average BM is different in each sector and for some sectors (*e.g.* Financial) this difference is large. We try to include this effect into our portfolio creation. Instead of dividing by deviation from median of all companies, we do it separately for each sector. Results of this analysis are in Table A.2. As we can see, results remain similar for all variables except *BM* ratio, were the effect is now positive. Even though this effect is not significant for 12 month holding period, it is significant for 3 month (at 1%) as well as for 6 month (at 5%) holding period returns.

If we split the data into two portfolios based on mean rather than median, significance will improve (see Table A.3). Now the effect of *BM* is still positive and significant even for 12 month holding period returns. Moreover, return momentum variables are more significant. 3 month momentum has significant positive effect for 3 month holding period returns, 6 month momentum is significant at all returns and 12 month for 6 and 12 month return.

These results suggest that there is a positive relationship between *BM* ratio and future returns but we have to control for heterogeneity in the data.

4.3 Outliers

As we mentioned earlier, there are 163 companies with negative book-to-market ratio (in at least one period). Those observations present the issue for our panel regression model as well as for portfolio analysis. In the panel regression, it cause value coefficient to be insignificant and in portfolio formation we would need to create separate quintile for negative BM companies to separate this effect from low BM stocks. In light with those issues, we decided not to use those companies in final analysis. Same approach was followed by e.g. Fama & French (1993). Furthermore, we decided not to use stocks with average monthly return higher than 100% as this would affect the regression results. There are only 13 such stocks in the dataset. We argue that such high return is caused by some external factors rather than fundamentals.

Chapter 5

Results

As we mention in the methodology chapter, we proceed in three steps when analyzing the dataset. Firstly, we try to estimate panel data model using the fixed effect model. The goal is not to create pricing model but to examine whether our selected variables have desired effects on stock returns. Secondly, we examine size and value effects using the portfolio approach of Fama & French (1993). The purpose is to see if the variation in stock returns is captured by size and value risk factors. The advantage of this method compared to panel estimation is that we are dealing with significantly less noise as the returns and other variables are averaged. Moreover, we also avoid issues with serial correlation. The last section of this chapter is focused entirely on formation of trading strategies based on findings from the panel and the portfolio analyses. We compare several strategies using common indicators such as the Sharpe ratio and maximum drawdown. Furthermore, we add a bootstrap analysis to test our strategies on reduced dataset. Last but not least, commissions are added to trading algorithms to better reflect actual trading conditions.

The last part of this chapter is dedicated to discussion of results of the three step analysis described above. Moreover, we highlight potential issues caused by limitations of used dataset.

5.1 Panel Model Estimation

We estimated the general model 3.1 using several methods for all dependent variables. Before we comment on the results, let us explain one thing regarding the explanatory variables, which is exclusion of the market effect. Fama & French (1993) use the excess market returns in their *three factor model* to explain stock returns. But this variable is only useful *ex-post* when we know this return. The goal of our thesis is not to develop asset pricing model, but to find out which fundamental variables can predict stock returns. Therefore, the market return cannot be used even though it would improve our model. The reason is a look-ahead bias, *i.e.* trading on information that is not known at the time of selecting stocks. The same applies for time dummy variables.

The model 3.1 was estimated using 3 different methods: pooled OLS, fixed effects and first differences. We assume that there is an unobserved heterogeneity in the data (*i.e.* some factors in the error term that are correlated with our explanatory variables), thus the pooled OLS and random effect models are not consistent. Nevertheless, we want to perform tests to select models and compare estimated coefficients across models. We use cluster-robust standard errors of M. Arellano (1987) to account for serial correlation and heteroskedasticity.

5.1.1 Estimation Results

First, lets look at the pooled OLS estimation (table A.4). As we can see, there are not many significant variables. There is a negative effect of 1 month return momentum, but it is not significant for all holding periods. From the set of fundamental variables, we only see positive effect of book-to-market ratio, which is in line with our hypothesis, but only for 1 month holding period. There is no size effect at all. The last thing we want to check for pooled OLS models is R-squared which is very low (spanning from 0.5% for 1 month to 0.8% for 12 month holding period) suggesting that we barely explained any variation in returns. Nevertheless, we have to point out that the variation in returns for 2000 stocks over 7 years is very large for any model to explain. This is exactly the reason why most of the academical research use the portfolios formed on some fundamental (or momentum) characteristics. We use this approach in the next section.

In the methodology section we pointed out that there is the unobserved heterogeneity in our data - for example the ability of the management of a company that is correlated with value of fundamental ratios. To verify this, we run the Breusch-Pagan Lagrange multiplier test for individual effects. With the test statistic of 28 647, 257 840, 129 340 and 418 950 for 1, 3, 6 and 12 month holding periods, we can reject the null hypothesis that there are no significant individual effects. Therefore, the pooled OLS model is not consistent and we have to use different method.

Results of the fixed effect estimation in table A.5 show positive and significant effect of book-to-market ratio on stock returns (for 1 and 3 month holding period significant at 10% level, otherwise significant at 1% level). Moreover, there is a significant negative effect of size, confirming our hypothesis and results of e.g. Fama & French (1992) and Fama & French (1993). Momentum results are mixed. There is a negative effect of one month momentum similar as in the pooled OLS model, but only for 1 and 3 month holding periods. For 6 and 12 month periods, there is a significant and negative effect of 12 and 6 month momentum, respectively. On the other hand, there is also positive and significant effect of 3 month momentum on 12 month holding period returns. The effect of PE ratio on all holding period returns is statistically indistinguishable from zero; dividend yield has positive and significant effect (at 10%) for 1 and 12 month holding period returns and the effect of current ratio is negative and significant as in pooled OLS model (for all holding periods). R-squared of the fixed effect estimation is again very low (though at least two times higher than for pooled OLS estimation) - lowest for 1 month holding period (1.2%) and highest for 12 month holding period (8.1%).

We also run the Breusch-Godfrey/Wooldridge test for serial correlation in panel models to see if there is any serial correlation left in residuals. We were able to reject the null hypothesis of no serial correlation (of order 1) for all except 1 month holding period. This make sense since by using 3, 6 and 12 month holding periods with 1 month data frequency, we are regressing overlapping periods. Nonetheless, even with serial correlation robust standard errors, the value and size effects are still significant.

The first difference model (table A.6) results in similar effects as the fixed effect model but with higher magnitude for most of variables. One difference is PE ratio, that has now significant negative effect for all holding periods (though this effect is small compared to other variables). The second difference is current ratio, which is now insignificant. Momentum variables, on the other hand, are now significant and negative for all holding periods. Finally, R-squared for first difference estimation is much higher than for the fixed effect model with the high of 36.5% for 1 month holding period and the low of 28.0% for 12 month holding period.

5.1.2 Model Calibration

Table 5.1: Different FE models for 12 month returns

	Dependent variable:					
	Return12					
	(1)	(2)	(3)	(4)	(5)	(6)
Ret1	−0.034	−0.011				-0.055^{**}
	(0.021)	(0.017)				(0.022)
Ret3	0.029^{***}	0.022^{*}				0.014
	(0.011)	(0.012)				(0.011)
Ret6	-0.064^{***}	-0.045^{***}				-0.082^{***}
	(0.016)	(0.011)				(0.015)
Ret12	−0.017	−0.021				−0.024
	(0.012)	(0.013)				(0.015)
BM	0.248^{***}	0.213^{***}	0.124^{***}	0.255^{***}	0.124^{***}	0.095^{**}
	(0.044)	(0.036)	(0.046)	(0.062)	(0.047)	(0.045)
Size	-0.004^{***}	-0.002^{***}	-0.004^{***}	-0.004^{***}	-0.004^{***}	-0.004^{***}
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
PE	0.00001	0.00003				
	(0.00003)	(0.00003)				
DP	0.398^{**}	0.361^{**}			0.006	
	(0.182)	(0.181)			(0.147)	
CR	-0.020^{***}	-0.019^{***}		-0.022^{*}		
	(0.007)	(0.007)		(0.013)		
Market12		1.523^{***}				
		(0.046)				
Observations	71,258	71,258	109,417	84,590	109,417	91,706
R^2	0.081	0.180	0.015	0.026	0.015	0.045
Adjusted R^2	0.079	0.177	0.015	0.026	0.015	0.044
Note:				*p<0.1; **p<0.05; ***p<0.01		

Table 5.1 presents various model setups using different explanatory variables. The model (1) contains same variables as in table A.5. By adding 12

month market return to the equation, we see big improvement in explained variance (R-squared increases from 8.1% to 18.0%) and slight decrease in some coefficients (size and BM). However, the coefficient signs and significance remained unchanged. After excluding all variables except size and book-to-market ratio (model 3), R-squared decreased to 1.5% but the size and value effects are still significant. The magnitude of BM effect is smaller than in model 1, but this is due to different amount of observations (model 1 excludes stocks without complete data). We see (model 4) that by adding current ratio (CR), number of observation drops from 109 thousand to 85 thousand and BM coefficient increases to 0.255. On the other hand, by including dividend yield (DP) as in model 5, number of observations as well as BM coefficient remains the same. Current ratio is only marginally significant when added to the model with only fundamental variables, but dividend yield is not significant.

The last model includes all momentum variables and we see that there is a negative and significant effect for 1 and 6 month momentum. Again, size and BM effects are still significant and of similar magnitudes (small differences are again caused by lower number of observations due to unavailability of return momentum for new stocks).

Similar results are also achieved for different holding periods (tables A.7, A.8 and A.9 in the appendix). Size effect is always negative and significant and value effect is always positive and significant (however for shorter holding periods it is significant only at 10%). Dividend yield is not significant when used alone with size and BM ratio. Current ratio has a negative significant effect even when used alone with size and BM ratio. For market momentum, we observe negative effect for 1 and 3 month holding periods.

5.1.3 Summary

Panel data estimation that we present in this section confirms our hypothesis about negative effect of size as well as positive effect of book-to-market ratio on future stock returns. Estimated coefficients remained significant even after removing other variables from the model and changing holding period. Moreover, we used heteroskedasticity and serial correlation consistent standard errors. Other fundamental variables, such as dividend yield or current ratio, have significant effect in some models but not in all of them. Therefore we cannot draw any obvious conclusions from them. The last observation is negative and significant effect of return momentum.

5.2 Portfolio Formation

We constructed 25 portfolios based on size and book-to-market ratio. Note that we used only NYSE stocks to find quantile breakpoints for size and we also obtained BM breakpoint for given period separately for each sector. This sector neutral split is due to our earlier observation (table 4.3) that e.g. financial companies tend to have, in average, higher BM ratio than the other companies. This step is additional compared to Fama & French (1993) but they do not include financial companies at all. However, since this sector accounts for large part of our dataset (442 companies out of 2002), we decided to include it.

Table 5.2: Descriptive statistics for 25 portfolios

Size	Book-to-market ratio quintiles									
quintiles	Low	2	3	4	High	Low	2	3	4	High
	Average number of firms					Average pct value				
Small	75.90	80.00	109.17	142.76	205.79	0.28	0.29	0.37	0.39	0.39
2	43.86	55.13	57.39	51.95	34.18	0.66	0.84	0.85	0.76	0.51
3	54.75	53.17	45.26	40.60	22.39	1.66	1.57	1.33	1.20	0.66
4	61.60	57.73	40.77	31.20	21.55	4.17	3.83	2.70	2.06	1.43
Big	65.18	50.14	43.44	29.65	15.05	23.90	19.12	15.42	10.90	4.69
	Average firm size					Average BM ratio				
Small	0.49	0.47	0.44	0.36	0.25	0.19	0.45	0.64	0.88	1.88
2	1.95	1.99	1.94	1.92	1.94	0.22	0.44	0.59	0.81	1.43
3	3.95	3.85	3.83	3.85	3.84	0.23	0.41	0.57	0.78	1.39
4	8.82	8.65	8.63	8.62	8.66	0.20	0.38	0.55	0.77	1.40
Big	47.80	49.71	46.26	47.91	40.65	0.19	0.37	0.54	0.79	1.18
	Average PE ratio					Average dividend yield				
Small	3.31	13.90	6.82	10.05	9.44	1.64	1.98	1.73	2.06	3.87
2	29.35	22.53	12.30	15.63	13.86	2.31	2.35	2.25	2.51	3.58
3	24.26	19.43	11.27	13.01	13.22	1.64	1.95	2.02	1.87	2.61
4	35.27	22.02	17.75	16.15	17.64	1.20	1.97	1.87	2.13	3.37
Big	27.78	20.41	14.56	14.22	8.40	1.67	2.05	2.36	2.98	3.18

Source: author's computations.

Table 5.2 shows that the smallest size quantile contains most stocks in average. This is due to usage of NYSE stocks only, when obtaining quantile breakpoints for given time period, rather than all stocks. Even though more stocks are allocated to the lowest size quintile portfolios, the average market value (obtained as average size multiplied by average number of firms) of those

5 portfolios is only 1.72 % of the combined average value of stocks in all 25 portfolios. In contrast, the largest size quintile (all 5 portfolios) contains 74% of total value in average. The lowest BM portfolio in the largest size quintile alone accounts for almost 24% of the total combined value of all stocks.

Table 5.2 also shows that in every except smallest size quantile, both percentage value and number of stocks tend to decrease with increasing BM ratio. This has two causes (Fama & French 1993):

- Highest BM quintile is tilted towards smallest stocks.
- NASDAQ stocks (mostly small) tend to have lower book-to-market equity ratios than NYSE stocks of similar size. Therefore, small NYSE stocks are more likely to be fallen angles (big firm, low stock price) that small NASDAQ and other stocks.

Regarding average BM ratios we observe that in all BM quintiles except largest, average BM ratios are similar irrespective of Size (or slightly decreasing with size). In the top quintile, BM is increasing with decreasing size - difference between bottom and top size quintiles is 0.7.

The last two parts of table 5.2 are dedicated to price-earnings ratio and dividend yield. From the former we can see that stocks with lower BM ratios tend to have higher PE ratio (that is higher stock price compared to earnings) than stocks with higher BM ratio. This holds true for all size quintiles except the lowest one. The latter indicates that higher BM companies tend to pay higher dividends.

Table A.10 then shows average monthly returns for all 25 portfolios. There are two panels: panel A uses value-weighted average returns when constructing portfolios (where weight is a stock price), panel B assumes equal weighted portfolios (impossible to reproduce without large invested amounts). There is hardly any size effect observed for value weighted portfolios, if any, it is opposite than we assumed (returns are increasing with size). Only for highest BM quantile, there is a negative relationship between returns and size but only if we disregard low average return for small stocks with high BM which is only 0.95% per month. For equally weighted portfolios this effect is more visible, but still only for 3 out of 5 BM quantiles.

The value (book-to-market) effect is a bit stronger. For value weighted portfolios, there is a positive effect for 3 out of 5 size quantiles. For equally weighted portfolios, the relationship is positive in all size quantiles except 4th.

Furthermore, the largest 2 average monthly returns were observed for portfolios in the smallest two size quantiles and the top BM quantile (for equally weighted portfolios). For value weighted portfolios, there is low average return of 0.95% observed for small companies with high BM ratio (which is also statistically not different from zero). However in the top BM quantile and the second smallest size quantile, the average return is still the largest among all portfolios.

Table A.11 summarizes average returns for the same 25 portfolios for 3, 6 and 12 month holding periods (we are not reporting t-statistics and p-values, but everything was statistically different from zero at 99% level). We observe similar characteristics as for 1 month holding period. Again, BM effect is more present for equally weighted portfolios and size effect is now only visible for highest BM quantile. One interesting observation is that larger difference between value and equally weighted portfolios is for smaller stock portfolios suggesting that there may be some low priced outliers (concerning returns) that are shifting average return upwards.

5.2.1 Factor Estimation

We estimated our risk factors SMB and HML related to size and book-to-market ratio, respectively (summary statistics are in table 5.3). As we state in the methodology section, both factors should be unrelated to each other (SMB should mimic only size effect and HML only book-to-market effect). We verified this by obtaining correlation between SMB and HML returns which is only 0.14 (0.29 for equally weighted portfolios). Unfortunately, the average SMB return is -0.26% per month, which is against our assumed direction of size effect. However with t-statistics of -1.19, this result is statistically not different from zero. The average HML return is positive, 0.12% per month, which is consistent with our hypothesis. Nonetheless, still statistically indistinguishable from zero. When we use equally weighted portfolio returns instead of value-weighted returns, results are bit different. The average SMB return is now 0.17% which is still only 0.9 standard errors from zero, but it is now positive, suggesting that there is some size effect. For the HML factor, average value is now 0.47%, which is with t-statistic of 1.69 statistically significant at 10% level.

Regarding the market risk factor, we calculated average monthly return to all stocks in our dataset (RM) which is 1.36% for value weighted and 1.76% for equally weighted returns. This is slightly higher than 1.18% average return

Table 5.3: Summary statistics for risk factors

				Correlations			
Name	Mean	SD	t-stat	RM	SP500	SMB	HML
Panel A: Value weighted portfolios							
RM	1.36	4.69	2.64	1.00	0.95	0.37	0.39
SP500	1.18	4.12	2.62	0.95	1.00	0.18	0.38
SMB	-0.26	1.95	-1.19	0.37	0.18	1.00	0.14
HML	0.12	2.2	0.51	0.39	0.38	0.14	1.00
Panel B: Equally weighted portfolios							
RM	1.76	5.18	3.09	1.00	0.92	0.40	0.64
SP500	1.18	4.12	2.62	0.92	1.00	0.12	0.49
SMB	0.17	1.69	0.9	0.40	0.12	1.00	0.29
HML	0.47	2.52	1.69	0.64	0.49	0.29	1.00

Source: author's computations.

to SP500 index. We can see that the correlation of total market portfolio with SP500 index is very high (95% for value weighted and 92% for equally weighted returns).

5.2.2 Time Series Regression

In this section we develop time series models that aim to explain variation in stock returns (our 25 portfolios) using risk factors. Firstly, we estimate market model (3.4) using only market return, then we add fundamental factors. Finally, we drop market factor from our regression since this information is useless for prediction purposes.

Market model. Estimation results for market model are in table A.12. Clearly, market alone explain much variation in stock returns. However with r-squared ranging from 0.7 to 0.92, there is still some space for fundamental factors.

3 factor model. After inclusion of size and book-to-market risk factors into the regression (table A.13), R-squared significantly improves for some portfolios (especially small stock and with high BM ratio). Based on mostly significant coefficients (except for highest BM quintile, where only 2 out of 5 have t-statistic bigger than 3 in absolute values), we can say that SMB factor captures variation is stock returns related to size (which is not captured by the market factor). Furthermore, the estimated coefficients are clearly negatively related to size. For the HML factor, almost all coefficients are significant. This suggest that the variation in stock returns related to value effect is captured by HML factor. Moreover, the coefficients are positively related to book-to-market ratio for all size quintiles.

Fundamental model. Table A.14 shows that used alone, SMB and HML factors still capture much variation in stock returns. R-squared is higher for low size as well as for high book-to-market portfolios which suggest that those factors capture variation related to size and value. Furthermore, the coefficients (mostly significant) are still negatively and positively related to size and BM ratio, respectively.

5.2.3 Summary

In this section we examined size and value effect on portfolio level instead of company level as in the panel data estimation. Our stocks were sorted each month based on size and book-to-market ratio into 25 portfolios. Similarly as Fama & French (1993), we then examined if the variation is stock returns for those portfolios could be explained by risk factors related to size, value and the market. The results confirmed our conclusions from panel data estimation, that there is a positive effect of BM ratio on stock returns. While the size effect is not present in lower BM quintiles, it is visible in the highest book-to-market quintile. This is important information for our trading strategy - if we select high BM stocks, which are expected to perform better than low BM stocks, we can also take use of size effect to further improve returns. Similarly as in panel model, large amount of variation could be explained by the market return. Nonetheless, as we explained earlier this information is available only *ex-post* and is thus useless for trading purposes.

5.3 Fundamental trading strategy

In the preceding sections we identified some relationships between stock returns and underlying fundamentals. Now the important question arises: is there a way to utilize this knowledge to form working trading strategy? And more importantly, is the advantage sustainable? According to the efficient market hypothesis, every market inefficiency (and thus opportunity) should be immediately erased by large number of individuals participating in security trading. But if this is the case, why we and other authors observed this anomaly?

When we look back at table A.10, we see that:

- When using value-weighting, best portfolio with average monthly return of 1.93% is the intersection of highest book-to-market quintile and second smallest size quintile
- For equal-weighting, smallest stocks with largest book-to-market ratio outperformed the rest with average monthly return of 2.52%

To have a complete picture: average monthly return of the S&P 500 index in the same period is 1.18%, the average monthly return of market portfolio (i.e. portfolio formed from all stocks in our dataset) is 1.36% with value-weighting and 1.76% with equal-weighting. In other words, the top *fundamental portfolio* outperformed the market portfolio by 0.57% and 0.76% per month for value and equal weighting, respectively. Moreover, both portfolios outperformed the S&P 500 index by 0.75% and 1.34% per month.

Figure B.1 shows evolution of invested $1000 in the beginning of 2009 for the best value weighted portfolio, market portfolio and the S&P 500 index. As we can see, our size-BM portfolio outperformed the market by significant margin. The value of our investment is $4254 for fundamental strategy compared to $2810 and $2475 for market portfolio and SP500 index, respectively.

So is it possible to replicate this in actual trading strategy? We can see in the table 5.2 that there are in average 206 stocks in the lowest size/maximum BM quantile. This is to much for anyone to buy and manage. Moreover, as noted before, we would have to create equally weighted portfolio which is complicated and impossible without large capital. Value weighted portfolio was outperformed by the market. Therefore we need to define some additional parameters before we test our strategies.

5.3.1 Trading Strategies Selection

Based on the findings from previous sections, we constructed several trading strategies. For all strategies, we tested 2 versions depending on frequency of rebalancing of portfolios. The first version assumes monthly, whilst the second yearly rebalancing. Table A.15 in the appendix lists 24 strategies based on different conditions. Generally we focused on the top book-to-market quintile and the smallest size quintile since it tend to perform better than other quintiles as discovered in portfolio analysis in the previous chapter. Two exceptions are the largest size quintile and the second smallest size quintile (both within largest BM quintile) which also showed high average monthly return of 1.84% and 1.93%, respectively. Problem with the smallest size quintile is that it contains too many stocks, hence we need to apply another conditions. Selected conditions are:

- Smallest 30 stocks
- 30 stocks with smallest current ratio
- 30 stocks with highest PE ratio
- 30 stocks with highest dividend yield
- 30 stocks with lowest past returns (1, 3, 6 and 12 month)
- 30 stocks with highest past returns (1, 3, 6 and 12 month)

Additionally we tested strategies selecting 30 highest BM stocks, 30 random stocks within top BM quintile and solely momentum and mean reversion strategies without any initial fundamental selection (using 1, 3, 6 and 12 month momentum variables).

Table A.15 shows also average returns for monthly and yearly rebalanced strategies. Clearly, all solely momentum strategies were outperformed by the market (1.36% per month). So were those formed on current ratio, dividend yield and PE ratio. Neither selecting 30 stocks with the highest BM ratio did perform very well. But there are 5 strategies (bold in table A.15) that managed to outperform the market portfolio (and the S&P 500 index). We will analyze them in further detail.

Strategy A: moderately small stocks. This is strategy depicted in figure B.1. It assumes restructuring of the portfolio based on actual sorts of all stocks on size and BM ratio. Each month, we ought to buy all stocks in top book-to-market quintile that are simultaneously in second smallest size quintile. Similarly, each stock that is no longer in these quintiles, should be sold. We chose the second smallest size quantile to avoid imposing additional parameters as we need in the smallest size quintile. This portfolio contains in average only 34 stocks, hence it is manageable to maintain without large commissions. Moreover, it is the best of our 25 portfolios when using value weighting.

Strategy B: small stocks. In this strategy, we again invest only in the top BM quintile but we focus on the smallest size quintile, rather than the second smallest as in the strategy A. Since there are too many stocks, we chose only 30 smallest ones.

Strategy C: big stocks. We are again investing in the largest BM quintile, but now we chose all stocks in the largest size quintile. If the size effect is present, this strategy should perform worse than the strategies A and B.

Strategy D: past losers. Similar to the strategy B, but instead of selecting 30 smallest stocks we invest into worst companies in terms of cumulative return in the preceding month. If there is a negative effect of return momentum as we observed in our panel models, this should result in higher returns.

Strategy E: past winners. In the contrary to the strategy D, we invest in 30 best performers in terms of 6 month cumulative returns.

Table 5.4: Portfolio descriptive statistics

	Monthly re-balancing				Yearly re-balancing			
Strategy	Return	Stocks	Bought	Sold	Return	Stocks	Bought	Sold
A: moderate	1.93	34.18	6.69	6.25	1.97	33.86	19.29	12.14
B: small	3.14	30.00	4.39	4.04	1.78	30.00	16.14	11.86
C: big	1.78	15.05	1.80	1.63	1.61	15.71	7.71	5.29
D: losers	2.14	30.00	25.36	25.00	1.05	30.00	27.00	22.71
E: winners	1.79	30.00	12.94	12.58	1.82	30.00	26.29	22.00

Source: author's computations.

Table 5.4 shows average returns for all 5 strategies (without any additional costs, such as commissions), as well as average size of the portfolios and average number of stocks bought and sold each month or year. The winner strategy for monthly re-balanced portfolios in terms of average return seems to be the strategy B (see figure B.2 and B.7 that shows evolution of invested \$1000 using this strategy). Average return of 3.14% outperformed market by 1.78% (and the S&P 500 index by 1.96%) per month. Moreover, in average only 4 stocks were needed to trade each month. On the other hand, when we rebalance the portfolio only once a year, this return premium significantly decreases to 1.78% per month.

Out of the yearly rebalanced portfolios, the best strategy is the strategy A with the average monthly return of 1.97%. This method (which we examined earlier in the figure B.1) seems to be consistently good regardless of the rebalancing period.

The strategy C (figure B.3 and B.8), which invests in the largest stocks in the top BM quintile was indeed outperformed by the strategies A and B as we hypothesized, but it still shows higher average monthly return than the market portfolio (by 0.42% for monthly re-balancing and 0.25% for yearly rebalancing). The advantage of this method is small amount of stocks needed to trade.

The mean reverting strategy D, based on buying past losers, seem to outperform the market with average return of 2.14%. Unfortunately the actual simulation of invested $1000 in figure B.4 revealed that it is caused by rapid increase in the first two years. This initial successful period is followed by two large plunges and then steady decline from 2014 till 2016. Interestingly, although we observe much lower average monthly return for yearly rebalanced portfolios, the actual performance is not that bad. Still below the market portfolio though.

The momentum strategy E (*i.e.* buying past winners) outperformed the market in terms of average monthly return but ended up below both the market and the S&P 500 index by the end of 2015 as a result of large decline from its peak at the beginning of 2014 (see figure B.5). Situation is much better for yearly rebalancing for both average return (1.82% per month) and the actual performance (figure B.10).

Last but not least, in terms of trades needed to execute, the last 2 methods require significantly larger number of stocks that we need to sell and buy each month (more than the half of total portfolio). We will look at the transaction costs and how they affect results later in this chapter.

5.3.2 Sharpe Ratio

To get the Sharpe ratio, we first obtained portfolio variance using the equation 3.8. Then we estimated Sharpe ratio for all of our strategies (for both monthly and yearly rebalanced portfolios) using risk free rate as a benchmark in the formula 3.9 (labeled as *Sharpe ratio A* in the table 5.5). Simultaneously, we used the return to the S&P 500 index as a benchmark when calculating the ratio (labeled as *Sharpe ratio B* in the table 5.5).

Table 5.5: Sharpe ratio comparison

	Monthly re-balancing			Yearly re-balancing		
Strategy	Return	Sharpe A	Sharpe B	Return	Sharpe A	Sharpe B
A: moderate	1.93	2.17	0.81	1.97	2.05	0.68
B: small	3.14	1.95	0.95	1.78	1.41	0.13
C: big	1.78	1.90	0.47	1.61	1.71	0.26
D: losers	2.14	1.37	0.65	1.05	1.02	0.03
E: winners	1.79	1.77	0.61	1.82	1.92	0.66

Source: author's computations.

We observe interesting results for the strategies A and B. Even though the second strategy, which invests into 30 smallest stocks in the highest book-to-market quintile, yielded better average return over the testing period, the Sharpe ratio of 1.95 is lower than the ratio of the first strategy (2.17). This result suggest that the second strategy is riskier compared to the first one. On the other hand, when using the S&P 500 index as a benchmark in the Sharpe ratio calculation, the strategy B is preferred. For yearly rebalanced portfolios, the strategy A is preferred in terms of both return and the Sharpe ratio.

Our third strategy, that selects large instead of small companies in the top book-to-market quintile, resulted in similar Sharpe ratio as the second strategy for monthly rebalanced portfolios (when using risk free rate as a benchmark). What is interesting is that for yearly rebalanced portfolios it is even higher (1.71 compared to 1.41) even though the return is smaller.

The Sharpe ratio of the mean reverting strategy D is the smallest of all strategies for both monthly and yearly rebalancing. This is in conclusion with the previous section where we have discovered that the high monthly return is only caused by few outliers in 2009 and in the first half of 2010. On the other hand, the momentum strategy E performed very well especially for yearly rebalanced portfolios.

5.3.3 Maximum Drawdown

We calculated maximum drawdowns (that is a maximum loss from a peak to a through before new peak is reached) for all of our strategies. Table 5.6 shows depth of those drawdowns as well as other information including total drawdown length, length of the declining period and length of the recovery period.

If we compare only depth of a drawdown, the best is the strategy C (investing in large companies with large BM ratio), closely followed by the strategy A (moderately small, high BM stocks). Both strategies share the length of a drawdown as well as recovery period. Investors following the strategy C would have to suffer 22.7% loss over 5 months and then wait another 16 months before the value of their portfolio would returned back to value before the big decrease.

Even thought the drawdown of the strategy B (investing in smallest 30 stocks with high BM ratio) is relatively high (-34.5%), the length of this period is the shortest among all strategies. The same applies for recovery period, which is only 12 months long.

By this point it should be no surprise that the maximum drawdown of the mean reverting strategy is 45%. Together with the longest drawdown period, it makes it the worst strategy of all. The momentum strategy E, on the other hand, ended up with the shortest drawdown period of only 10 months.

Another interesting observation is that for the first 3 strategies, the declining period (i.e. from a peak to a through) seems to reflect overall decrease in the market during this period (the S&P 500 index declined from 1364 in the beginning of April 2011 to 1219 in the beginning of September 2011). But the value of our portfolios declined more than the market, which suggest that selected companies were largely affected.

Table 5.6: Maximum drawdowns for monthly rebalanced strategies

Strategy	Depth (%)	Length	To Through	Recovery	Start	End
A: moderate	-24.14	21	5	16	04-2011	12-2012
B: small	-34.46	14	2	12	07-2011	08-2012
C: big	-22.70	21	5	16	04-2011	12-2012
D: losers	-45.36	41	17	24	04-2010	08-2013
E: winners	-28.67	10	4	6	04-2010	01-2011

Source: author's computations.

Table 5.7 than shows results of the same analysis but for strategies that assumes yearly rebalancing of portfolios. The strategy C is again the best among all, followed by the strategy A. The depth of decline of the strategy C is slightly smaller than in table 5.6 and slightly bigger than for the strategy A, but the difference is not large. Even though the depth of maximum drawdown of the strategy C declined to -31.7%, previously attained peak still has not been recovered by the end of 2015 (when our testing period ends).

Table 5.7: Maximum drawdowns for yearly rebalanced strategies

Strategy	Depth (%)	Length	To Through	Recovery	Start	End
A: moderate	-27.04	23	5	18	04-2011	02-2013
B: small	-31.74	23	22	NA	03-2014	NA
C: big	-20.06	21	5	16	04-2011	12-2012
D: losers	-46.95	20	19	NA	06-2014	NA
E: winners	-29.93	22	5	17	04-2011	01-2013

Source: author's computations.

This analysis reveals that the best strategy in terms of maximum drawdown is investing into large companies with high book-to-market ratio (the strategy C). Similar but little worse results were observed for strategy A, that selects moderately small companies with high BM ratio.

5.3.4 Bootstrapping

To further assess the performance of selected strategies, we selected random subset of 1000 companies out of 1812 companies with positive book-to-market ratio. From this sample we obtained our 25 portfolios formed on size and BM (this time we did not use sector neutral portfolios due to the fact that our sample is smaller and we do not have enough observations for some sectors). We then simulated all investment strategies defined earlier. Whole experiment was repeated 100 times, each time different random subset of companies was selected. Table 5.8 summarizes results of this analysis for monthly formed portfolios. In first two columns we can see two different returns for each strategy. The fist one is a return obtained from the full sample (already shown in table 5.4), the second is an average return of all 100 trials obtained from bootstrap analysis. The last two columns then show how many times the strategy ended up below either market or the S&P 500 index. Note that the market portfolio is different than we defined it earlier (see table 5.3) since it is always calculated only from selected 1000 companies for each trial. The market return calculated from the full sample is 1.36% per month whereas the average return from all trials is 1.33% per month. The S&P 500 index monthly return is 1.18%.

In the appendix, there are histograms summarizing all bootstrap trials for each strategy (B.11 to B.15 for monthly rebalancing and B.16 to B.20 for yearly rebalancing).

Table 5.8: Bootstrap summary - monthly rebalanced portfolios

	Return		# of trials below	
Strategy	Full sample	Average	Market	S&P 500
A: moderate	1.93	1.63	23	16
B: small	3.14	1.86	4	1
C: big	1.78	1.21	59	49
D: losers	2.14	1.29	56	34
E: winners	1.79	1.29	56	33

Source: author's computations.

The best strategy in terms of average return and count of trials below market and the S&P 500 index is the strategy B (investing into 30 smallest stocks in the top BM quintile). On the other, there is also the largest difference between full sample and average return (3.14% vs 1.86%) suggesting that this significant premium to the market return is depending on a few stocks and when those are taken out of the sample, return drops. In contrast, return of the strategy A (moderately small, high BM stocks) only drops from 1.93% to 1.63% per month. But when we look at count of trials that ended up below the market, it is much higher than for the strategy B. In total 23 times out of 100 investors would ended below the market portfolio and 16 times below S&P 500 index.

Disappointing results are observed for the third strategy (large stocks, high book-to-market ratio) where the average return drops below market causing also large number of trials underperforming both market and S&P 500 index. The reason for such drop is probably small number of companies in the portfolios (only 15 for full sample - see table 5.4). When we further reduce the dataset, some companies that used to be in the smaller size quintile (with smaller average return of 1.26% compared to 1.84% as we can see in table A.10 in the appendix) moves to the largest size quintile causing the average return to decrease. We can see in figure B.13 that there are indeed several trials below 1% mark while the rest is well above it.

The last two strategies (mean reversion and momentum) both ended up below market portfolio in terms of average returns.

Table 5.9: Bootstrap summary - yearly rebalanced portfolios

	Return		# of trials below	
Strategy	Full sample	Average	Market	S&P 500
A: moderate	1.97	1.37	54	39
B: small	1.78	1.46	26	9
C: big	1.61	1.47	21	19
D: losers	1.02	0.96	96	83
E: winners	1.92	1.57	13	10

Source: author's computations.

Table 5.9 shows the same analysis but for yearly rebalanced portfolios. There are two surprising results. The first one is a relatively poor performance of the first strategy. With the average return of 1.37% per month, this investment method only outperformed market portfolio 46 times, which is much worse in comparison with yearly rebalanced portfolios. Also the second strategy shows worse results than for yearly rebalanced portfolios.

The second surprising result is the performance of the last strategy - that is investing in past winners. With the average return of 1.57% per month this strategy outperformed both market portfolio and S&P 500 index as well as all other strategies. Furthermore, it was outperformed by the market portfolio only 13 times out of 100 trials. On the other hand, as figure B.20 shows, there are few trials that resulted in very low average return (below 1% per month).

5.3.5 Transaction Costs

For the purposes of trading simulation with transaction costs (broker's commissions) lets assume following setup:

Invested value: $100 000

Commission: $0.0075 per share ($1.00 minimum)

Commission were obtained from the Interactive brokers for API connection (normally commission is only $0.005 per share, but we assume that trade is executed using some algorithm, thus we need API). [1]

[1]Source: https://www.interactivebrokers.com/en/index.php?f=1590&p=stocks1

Table 5.10: Trading summary with commissions

	Return (%)		Commissions ($)	
Strategy	w/o comm.	w/ comm.	Average	Total
A: moderate	1.93	1.92	17.39	1 440
B: small	3.14	3.10	222.78	18 714
C: big	1.78	1.78	9.86	829
D: losers	2.14	2.05	228.83	19 222
E: winners	1.79	1.76	102.06	8 573

Source: author's computations.

Table 5.10 compares previously reported average monthly return to same return but with commission included. As we can see, there is not a big difference in average return even though some strategies require heavy trading (see table 5.4). Reason for this is relatively low commission percentage that is not significant for monthly rebalanced strategies. Brokerage houses such as the Interactive brokers aim to day traders, i.e. investors who engage on markets on daily basis. This is clear advantage of fundamentally based algorithms - the amount of trading required is very low and so are commissions. Of course, there is a minimum value of $1 per trade therefore if we decrease initial invested amount so the algorithm will select only 1 stock per company in portfolio to purchase, commission percentage would increase. But we are assuming that investor has some funds available.

One of the largest commission paid in total was observed for the second strategy even though it does require less trading than strategies D and E. The reason being is the growth of portfolio value and thus higher amounts of stocks traded. But overall average return decreases only from 3.14% to 3.10% per month.

We only calculated transaction costs for monthly rebalanced portfolios. We do not need to perform similar analysis for yearly rebalancing since the amount of trading is much lower.

5.3.6 Summary

In this chapter we performed several evaluation methods to test all of our strategies and compared them to the market portfolio and the S&P 500 index. All strategies focused on the top book-to-market quintile, but selected different stock within this quintile. Table 5.11 shows the value of a portfolio at the end of 2015 with $1000 initial investment at the beginning of 2009 (including commissions), as well as compounded annual growth rate (CAGR) of a portfolio value.

The highest monthly return was observed for the strategy B, that selects 30 smallest stocks within the highest book-to-market quintile. With the CAGR of 36.6%, this strategy managed to beat both the market portfolio and the S&P 500 index by significant margin. However, the Sharpe ratio revealed that this investment method was also riskier than investing into moderately small stocks (strategy A). Furthermore, when rebalanced once a year, performance drops significantly (but still above market portfolio). While this strategy seem to be obvious choice for investors, there is high probability of survivorship bias caused by excluding companies that dropped out of the exchange (for example due to bankruptcy). Since we are investing in small stocks, this risk is much higher for this method than for the strategy B and especially C.

Table 5.11: Strategies final summary

	Monthly re-balancing			Yearly re-balancing		
Strategy	Return (%)	Value ($)	CAGR (%)	Return (%)	Value ($)	CAGR (%)
A: moderate	1.93	3 354	18.87	1.97	4 103	22.34
B: small	3.14	8 850	36.55	1.78	3 498	19.59
C: big	1.78	3 656	20.35	1.61	3 189	18.02
D: losers	2.14	387	-12.68	1.06	1 514	6.10
E: winners	1.79	2 072	10.97	1.82	3 323	18.71
Market portfolio	1.36	2 810	15.90	1.36	2 810	15.90
S&P 500	1.18	2 475	13.82	1.18	2 475	13.82

Source: author's computations.

Investing into moderately small (the second smallest size quintile) and big (largest size quintile) stocks with high book-to-market ratio leads to good performance in terms of both average return and CAGR. Moreover, when we look at risk adjusted return (Sharpe ratio in table 5.5) and maximum drawdown analysis (table 5.6), the results are even better than for strategy B (except Sharpe ratio of strategy C, which is slightly smaller). The only drawback of the strategy C is relatively poor performance in bootstrap analysis (table 5.8) for monthly rebalanced portfolios.

Our mean reversion strategy, on the other hand, failed for both monthly and yearly rebalanced portfolios since the average return of 2.14% was caused by several outliers in the beginning of testing period but then the value decreased to \$387 from initial \$1000. With yearly rebalancing, the performance is not that bad, but it is still below both the market portfolio and the S&P 500 index.

For yearly rebalanced portfolios, all strategies except mean reversion (D) managed to beat the market in terms of CAGR. However the best performance was observed for he strategy C (big stocks) and E (past winners). Both showed very good results in the Sharpe ratio, maximum drawdown and bootstrap analyses (see tables 5.5, 5.7 and 5.9).

5.4 Discussion of Results

The overall good performance of the trading strategies that select high BM stocks confirm our hypothesis and also results of the panel estimation in the section 5.1 as well as portfolio analysis in the section 5.2. On the other hand, not all stocks in the top book-to-market quintile tend to outperform the market and there is a need for other parameters.

When we use size factor, which should have negative effect according to our findings from sections 5.1 and 5.2, by constructing portfolios on sorting stocks in the top book-to-market quintile and selecting only smallest 30, the resulting portfolio generates outstanding return over the testing period. While this seem to be a proof of a negative size effect, we need to be careful here. Selecting small stocks is tricky since those are either undervalued companies or stock about to bankrupt. Unfortunately, our dataset excludes the latter. We discuss this and other drawbacks in the next section. Interestingly, the strategy which selects only stocks in the top size and top book-to-market quintile managed to beat the market in the testing period. Although we can argue that stocks within top size quintile tend to decrease in size with increasing BM ratio as we can see in table 5.2 and thus size effect is still visible.

The mean reversion strategy performed poorly despite high average monthly return, which is caused by outliers in the beginning of the testing period. This is in contrast with negative estimated parameters for past returns in the panel regression (section 5.1) and it raises a question whether those estimate were caused by outliers. The momentum strategy, on the other hand, managed to outperform the market when using 6 month momentum and yearly rebalanced portfolios.

Other tested variables, such as current ratio, dividend yield or price earnings ratio, show poor results when used for strategy formation. For PE ratio, it is in contrast with work of Basu (1975) and Basu (1983).

According to the efficient market hypothesis, all market inefficiencies should be immediately erased by large number of individuals trading on the same information. If this holds true, why do we observe better performance for some fundamentally based strategies? One explanation could be the limited dataset (discussed in the next section), but the predictive ability of size and book-to-market ratio was documented earlier by Fama & French (1993) on much larger data. This brings us to the second possible explanation and that is favoring of price based strategies by majority of the market thus creating an opportunity for fundamental strategies. When we look back at the stock market model of Koller *et al.* (2010), described in chapter 2, we see that intrinsic investors only hold around 20 to 25 percent of the U.S. equity market. Whereas, traders and mechanical investors control together the rest. If the majority of the market prefer to use price based strategies, it make sense that a stock value do not reflect fundamentals. Other explanation could be the preference of glamour stocks, *i.e.* companies with good past performance, over value stocks by both institutional and individual investors (Lakonishok *et al.* 1994).

The last thing we want to discuss here is the need to automation of such strategies. It seems that methods presented in this thesis could be managed without any use of algorithms since the amount of trading is low and portfolios are rebalanced only once per month (for some strategies, only once per year). We argue that this would bring the human factor into the equation and it would result in different selected companies. We are not saying that it would result in better or worse performance, but it would depend on abilities of a portfolio manager rather than quality of selected trading algorithm.

5.5 Limitations

While the results of the trading algorithms analyzed in section 5.3 seem really good, we need to stress several drawbacks of used dataset that may cause real performance to drop. Namely, those limitations are:

- Short testing period
- Incomplete stock listing
- Survivorship bias

- Data snooping bias

Unfortunately, to the author's best knowledge, obtaining dataset that would contain *e.g.* 30 years of data including bankrupted stocks is impossible without serious financial investment.

Short testing period. While the period of 7 years seem long enough to test a trading strategy, there is one big problem and that is exclusion of any major financial crisis that would surely present a challenge for any fundamentally based algorithm. On the other hand, we are comparing our strategies to the market portfolio and the S&P 500 index that both grew 16% and 14% (respectively) annually and some strategies still managed to beat them. Of course, we do not know what would happen to selected stocks during turmoil period, when fundamentals are ignored by the sentiment driven market. This would be interesting topic of future work.

Incomplete stock listing. As we saw in the bootstrap analysis, reducing the dataset can cause significant changes to any strategy. Even though 2000 companies seem to be large enough dataset, it is far from complete.

Survivorship bias. Probably the most dangerous limitation, especially for the algorithm based on investing into small stocks. The issue of companies with low market capitalization is the fact that those are either undervalued firms with good prospects or stocks just before the exit from the market. While the former stocks remained in the dataset, the latter might be excluded causing the strategy B perform better than it would in the real world application.

Data snooping bias. The issue discussed by *e.g.* Lakonishok *et al.* (1994). It is possible that we optimized the algorithms for our specific set of stocks and there is high probability that same assumption would fail when applied to different time period or different set of stocks (e.g. international stocks, small CAP stocks, etc.). This problem is closely related to the first two mentioned limitations. We tried to simulate the effect of having different set of stocks in the bootstrap analysis where we chose subset of 1000 randomly selected stocks, but we are still dealing with subset of the same dataset. On the other hand, strategies presented in this thesis are not heavily parametrized, hence we believe that the data snooping is minimized.

Any portfolio manager should test selected strategy on the dataset that minimize above mentioned issues before real time application.

Chapter 6

Conclusion

In this thesis we aim to apply methods of fundamental analysis into developing field of algorithmic trading. While the principles of value investing are known for several decades, its application to actual trading algorithms seem to be neglected in favor of technical and high frequency traders, who build their strategies mainly on price development and market sentiment.

We approach this problem in three steps. In the first step we estimate our dataset of 2000 companies using panel data model. This is rarely done in academical literature due to serial correlation and other issues. Nevertheless, our model revealed significant effects even after heteroskedasticity and serial correlation robust estimation. We observe positive effect of book-to-market ratio and negative effect of company size on future stock returns. In the second step, we apply method of Fama & French (1993) and form 25 portfolios based on the size and value effect (book-to-market effect). Then we try to explain variation in returns of those portfolios using three risk factor: market, size and value. We find that value and size effect explain significant variation in returns, thus confirming our results from panel regression. In the last part, we apply size and value effect into trading strategies. We choose 5 strategies based on size, book-to-market ratio and other conditions, such as past performance. We analyze each strategy using common methods including the Sharpe ratio and maximum drawdown. Furthermore, we add bootstrap analysis, that tests all strategies on randomly selected subsets of our dataset, and transaction costs in from of broker's commissions. The results shows that simple strategy that invests into 30 smallest stocks within top book-to-market quintile beats the market with average monthly return of 3.10% (including commissions). While this algorithm might be suspect to survivorship bias caused by excluding bankrupted companies from our dataset, we find that also selecting moderately small stocks leads to better than market performance for monthly rebalanced portfolios. On the other hand, when we change portfolios only once a year, the most successful strategy is selecting past winners within top book-to-market quintile or all companies in biggest size quintile.

According to the efficient market hypothesis, all inefficiencies in the market should be immediately erased by large amount of participants trading on the same information. Our results suggest that while this might be the case for technical based strategies, there is some space for fundamentally based algorithms. Unfortunately, we need to be careful when applying tested methods into practice since our dataset exhibits some limitations such as short time period without any serious financial turmoil and is not including failed companies. We leave those issues for a further research.

Bibliography

ASNESS, C. (2013): "Value and momentum everywhere." *The Journal of Finance* **68(3)**: pp. 929–985.

BADI H. BALTAGI (2005): *Econometric Analysis of Panel Data*. John Wiley & Sons Ltd, third edition.

BALL, R. (1978): "Anomalies in relationships between securities' yields and yield-surrogates." *Journal of Financial Economics* **6(2-3)**: pp. 103–126.

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

BASU, S. (1975): "The Information Content of Price-Earnings Ratios." *Financial Management (1972)* **4(2)**: pp. 53–64.

BASU, S. (1983): "The relationship between earnings' yield, market value and return for NYSE common stocks." *Journal of Financial Economics* **12(1)**: pp. 129–156.

BERTIMAS, D., A. LO, & P. HUMMEL (1999): "Optimal control of execution costs for portfolios." *Computing in Science & Engineering* **1(6)**.

BHANDARI, L. C. (1988): "Debt/equity ratio and expected common stock returns: empirical evidence." *The Journal of Finance* **43(2)**: pp. 507–528.

BLACK, F., M. C. JENSEN, & M. SCHOLES (1972): "The Capital Asset Pricing Model: Some Empirical Tests." *Studies in the Theory of Capital Markets* **81**: pp. 79–121.

CAPAUL, C., I. ROWLEY, & W. F. SHARPE (1993): "International Value and Growth Stock Returns."

CHAN, E. P. (2009): *Quantitative Trading: How to build your own algorithmic trading business*. Hoboken, New Jersey: John Wiley & Sons.

Chan, E. P. (2013): *Algorithmic Trading: Winning Strategies And Their Rationale*. Hoboken, New Jersey: John Wiley & Sons, Inc.

Chan, L. K. C., Y. Hamao, & J. Lakonishok (1991): "Fundamentals and Stock Returns in Japan." *The Journal of Finance* **46(5)**: p. 1739.

Fama, E. F. & K. R. French (1992): "The Cross-Section of Expected Stock Returns." *The Journal Of Finance* **47(2)**: pp. 427–466.

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

Fama, E. F. & K. R. French (2012): "Size, value, and momentum in international stock returns." *Journal of Financial Economics* **105(3)**: pp. 457–472.

Fama, E. F. & J. D. MacBeth (1973): "Risk, Return, and Equilibrium: Empirical Tests." *Journal of Political Economy* **81(3)**: pp. 607–636.

Graham, B. & D. Dodd (1934): *Security analysis*. McGraw-Hill.

Greene, W. H. (2012): *Econometric Analysis*. Pearson Education Limited, 7th edition.

Kirilenko, A. a. & A. W. Lo (2013): "Moore's Law versus Murphy's Law: Algorithmic Trading and Its Discontents." *Journal of Economic Perspectives* **27(2)**: pp. 51–72.

Koller, T., M. Goedhart, & D. Wessels (2010): *Valuation: Measuring and Managing the Value of Companies*. Wiley, 5th edition.

Lakonishok, A. Shleifer, R. Vishny, & O. Hart (1992): "The structure and performance of the money management industry." *Brookings Papers on Economic Activity. . . .* .

Lakonishok, J., A. Shleifer, & R. W. Vishny (1994): "Contrarian Investment, Extrapolation, and Risk." *The Journal of Finance* **49(5)**: pp. 1541 – 1578.

Lee, C. M. C. & B. Swaminathan (2000): "Price Momentum and Trading Volume." *The Journal of Finance* **55(5)**: pp. 2017–2069.

LINTNER, J. (1965): "The Valuation of Risky Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets." *Review of Economics and Statistics* **47(1)**: pp. 13–37.

LYLE, M. R. & C. C. Y. WANG (2014): "The Cross Section of Expected Holding Period Returns and their Dynamics: A Present Value Approach." *Journal of Financial Economics* .

M. ARELLANO (1987): "Computing robust standard errors for within group estimators." *Oxford Bulletin of Economics and Statistics* **49(4)**: pp. 431–434.

MALKIEL, B. (1973): *A Random Walk Down Wall Street.* W. W. Norton & Company, Inc.

MARKOWITZ, H. (1952): "Portfolio Selection." *The Journal of Finance* **7(1)**: pp. 77–91.

MOSSIN, J. (1966): "Equilibrium in a Capital Asset Market." *Econometrica* **34(4)**: pp. 768–783.

PIOTROSKI, J. D. (2000): "Value Investing: The Use of Historical Financial Statement Information to Separate Winners from Losers." *Journal of Accounting Research* **38(5)**: pp. 1–41.

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

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 (1990): "Equilibrium Short Horizons of Investors and Firms." *American Economic Review* **80(2)**: pp. 148–153.

STATTMAN, D. (1980): "Book Values And Stock Returns." *The Chicago MBA: A Journal of Selected Papers* **4**: pp. 25–45.

T. S. BREUSCH AND A. R. PAGAN (1980): "The Lagrange Multiplier Test and its Applications to Model Specification in Econometrics." *The Review of Economic Studies* **47(1)**: pp. 239–253.

Appendix A

Tables

Table A.1: Predictability tests

3 month return	Value		Return (%)		t-test	
Variable	Low	High	Low	High	t-stat	p-value
Size	544.00	16176.00	4.04	3.71	2.46	0.014
BM	0.20	1.18	3.93	3.82	0.80	0.424
PE	-17.16	45.60	4.16	3.59	4.22	0.000
DP	0.00	0.04	4.20	3.55	4.81	0.000
CR	1.25	4.25	4.19	3.85	2.10	0.036
Ret3	-0.09	0.17	3.60	3.58	0.20	0.842
Ret6	-0.11	0.28	3.60	3.69	-0.67	0.503
Ret12	-0.11	0.48	2.64	3.34	-5.34	0.000
6 month return	Value		Return (%)		t-test	
Variable	Low	High	Low	High	t-stat	p-value
Size	544.00	16176.00	8.29	7.65	3.10	0.002
BM	0.20	1.18	8.16	7.77	1.87	0.061
PE	-17.16	45.60	8.59	7.36	5.93	0.000
DP	0.00	0.04	8.74	7.19	7.46	0.000
CR	1.25	4.25	8.72	7.89	3.25	0.001
Ret3	-0.09	0.17	7.92	7.77	0.71	0.477
Ret6	-0.11	0.28	7.54	7.73	-0.93	0.352
Ret12	-0.11	0.48	5.75	6.82	-5.52	0.000
12 month return	Value		Return (%)		t-test	
Variable	Low	High	Low	High	t-stat	p-value
Size	544.00	16176.00	18.32	16.34	4.21	0.000
BM	0.20	1.18	18.05	16.60	3.09	0.002
PE	-17.16	45.60	18.80	15.88	6.21	0.000
DP	0.00	0.04	19.36	15.26	8.77	0.000
CR	1.25	4.25	19.36	16.88	4.22	0.000
Ret3	-0.09	0.17	16.58	16.60	-0.03	0.976
Ret6	-0.11	0.28	15.31	16.01	-1.50	0.134
Ret12	-0.11	0.48	14.25	15.27	-3.38	0.001

Source: author's computations.

Table A.2: Predictability tests - Sector Neutral

3 month return	Value		Return (%)		t-test	
Variable	Low	High	Low	High	t-stat	p-value
Size	623.00	16158.00	4.14	3.61	3.97	0.000
BM	0.25	1.13	3.69	4.06	-2.80	0.005
PE	-16.77	45.38	4.05	3.70	2.61	0.009
DP	0.00	0.05	4.09	3.58	4.14	0.000
CR	1.33	4.18	4.28	3.76	3.21	0.001
Ret3	-0.08	0.17	3.64	3.54	0.70	0.486
Ret6	-0.10	0.27	3.64	3.65	-0.08	0.933
Ret12	-0.10	0.47	2.73	3.25	-4.01	0.000
6 month return	Value		Return (%)		t-test	
Variable	Low	High	Low	High	t-stat	p-value
Size	623.00	16158.00	8.48	7.46	4.91	0.000
BM	0.25	1.13	7.72	8.22	-2.42	0.016
PE	-16.77	45.38	8.42	7.53	4.24	0.000
DP	0.00	0.05	8.50	7.25	6.44	0.000
CR	1.33	4.18	8.85	7.77	4.25	0.000
Ret3	-0.08	0.17	7.88	7.80	0.39	0.693
Ret6	-0.10	0.27	7.52	7.74	-1.06	0.287
Ret12	-0.10	0.47	5.87	6.71	-4.38	0.000
12 month return	Value		Return (%)		t-test	
Variable	Low	High	Low	High	t-stat	p-value
Size	623.00	16158.00	18.64	16.02	5.56	0.000
BM	0.25	1.13	17.17	17.50	-0.71	0.477
PE	-16.77	45.38	18.48	16.20	4.85	0.000
DP	0.00	0.05	18.77	15.37	8.08	0.000
CR	1.33	4.18	19.49	16.74	4.70	0.000
Ret3	-0.08	0.17	16.65	16.54	0.23	0.820
Ret6	-0.10	0.27	15.35	15.99	-1.36	0.174
Ret12	-0.10	0.47	14.51	15.03	-1.74	0.082

Source: author's computations.

Table A.3: Predictability tests - Sector Neutral (mean split)

3 month return	Value		Return (%)		t-test	
Variable	Low	High	Low	High	t-stat	p-value
CAP	1915.00	34219.00	3.97	3.49	4.27	0.000
BM	0.32	1.28	3.61	4.30	-4.80	0.000
PE	-14.31	43.93	5.19	3.96	6.04	0.000
DP	0.01	0.06	3.98	3.60	3.18	0.001
CR	1.55	4.22	11.43	7.67	4.87	0.000
Ret3	-0.09	0.14	1.37	1.94	-3.59	0.000
Ret6	-0.14	0.17	2.28	3.01	-3.24	0.001
Ret12	-0.08	0.23	2.33	3.18	-1.55	0.122
6 month return	Value		Return (%)		t-test	
Variable	Low	High	Low	High	t-stat	p-value
CAP	1915.00	34219.00	8.15	7.27	5.12	0.000
BM	0.32	1.28	7.61	8.53	-4.18	0.000
PE	-14.31	43.93	11.10	8.70	7.33	0.000
DP	0.01	0.06	8.27	7.20	5.75	0.000
CR	1.55	4.22	21.06	13.72	6.44	0.000
Ret3	-0.09	0.14	6.05	6.35	-1.28	0.200
Ret6	-0.14	0.17	7.88	8.48	-1.75	0.079
Ret12	-0.08	0.23	4.63	7.45	-3.59	0.000
12 month return	Value		Return (%)		t-test	
Variable	Low	High	Low	High	t-stat	p-value
CAP	1915.00	34219.00	17.80	15.45	6.94	0.000
BM	0.32	1.28	16.83	18.11	-2.86	0.004
PE	-14.31	43.93	24.23	18.87	6.57	0.000
DP	0.01	0.06	18.12	15.32	7.46	0.000
CR	1.55	4.22	34.24	22.24	6.57	0.000
Ret3	-0.09	0.14	13.18	13.42	-0.63	0.531
Ret6	-0.14	0.17	10.96	12.78	-3.42	0.001
Ret12	-0.08	0.23	12.28	15.64	-2.41	0.016

Source: author's computations.

Table A.4: Estimation Results - Pooled OLS

	Dependent variable:			
	Return1	Return3	Return6	Return12
	(1)	(2)	(3)	(4)
Ret1	-0.023^{***}	-0.038^{***}	-0.015	-0.040^{**}
	(0.009)	(0.009)	(0.010)	(0.017)
Ret3	-0.004	0.022	0.029	0.044^{**}
	(0.007)	(0.016)	(0.033)	(0.021)
Ret6	0.008^{***}	0.010	0.004	-0.018
	(0.006)	(0.020)	(0.044)	(0.019)
Ret12	-0.0004	-0.002	-0.004	-0.007
	(0.001)	(0.002)	(0.002)	(0.011)
BM	0.008^{***}	0.016	0.026	0.048
	(0.007)	(0.023)	(0.025)	(0.039)
Size	0.00000	0.00000	0.00000	-0.0001
	(0.0001)	(0.0003)	(0.001)	(0.001)
PE	-0.00000	0.00000	0.00000	0.00002
	(0.00000)	(0.00001)	(0.00002)	(0.00003)
DP	0.009^{***}	0.019	0.077	0.141
	(0.014)	(0.043)	(0.110)	(0.183)
CR	-0.001	-0.002	-0.003	-0.007
	(0.0004)	(0.001)	(0.003)	(0.007)
Constant	0.008	0.024	0.055	0.138
Observations	86,153	83,347	79,268	71,258
R^2	0.005	0.005	0.005	0.008
Adjusted R^2	0.005	0.005	0.005	0.008
Note:				*p<0.1; **p<0.05; ***p<0.01

Table A.5: Estimation Results - Fixed Effects Model

	Dependent variable:			
	Return1	Return3	Return6	Return12
	(1)	(2)	(3)	(4)
Ret1	-0.026^{***}	-0.040^{***}	-0.014	-0.034
	(0.009)	(0.010)	(0.022)	(0.021)
Ret3	-0.008	0.013	0.017	0.029^{***}
	(0.005)	(0.014)	(0.012)	(0.011)
Ret6	0.003	-0.004	-0.022	-0.064^{***}
	(0.005)	(0.018)	(0.035)	(0.016)
Ret12	-0.002	-0.005	-0.010^{**}	-0.017
	(0.001)	(0.003)	(0.004)	(0.012)
BM	0.014^{**}	0.043^{*}	0.101^{***}	0.248^{***}
	(0.007)	(0.023)	(0.026)	(0.044)
Size	-0.0004^{***}	-0.001^{***}	-0.002^{***}	-0.004^{***}
	(0.0001)	(0.0003)	(0.001)	(0.001)
PE	-0.00000	-0.00000	-0.00001	0.00001
	(0.00000)	(0.00001)	(0.00002)	(0.00003)
DP	0.026^{*}	0.070	0.196^{*}	0.398^{**}
	(0.014)	(0.044)	(0.110)	(0.182)
CR	-0.001^{**}	-0.004^{***}	-0.009^{***}	-0.020^{***}
	(0.0004)	(0.001)	(0.003)	(0.007)
Observations	86,153	83,347	79,268	71,258
R^2	0.012	0.023	0.041	0.081
Adjusted R^2	0.012	0.023	0.040	0.079
Note:		$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$		

Table A.6: Estimation Results - First Differences

	Dependent variable:			
	Return1	Return3	Return6	Return12
	(1)	(2)	(3)	(4)
Ret1	−0.307***	−0.292***	−0.256***	−0.283***
	(0.015)	(0.020)	(0.043)	(0.061)
Ret3	−0.215***	−0.198***	−0.188***	−0.190***
	(0.010)	(0.009)	(0.010)	(0.020)
Ret6	−0.106***	−0.120***	−0.131***	−0.137***
	(0.018)	(0.020)	(0.025)	(0.024)
Ret12	−0.017	−0.020	−0.020	−0.013
	(0.011)	(0.013)	(0.012)	(0.009)
BM	0.079***	0.077	0.107***	0.135***
	(0.022)	(0.048)	(0.024)	(0.047)
Size	−0.004***	−0.004***	−0.005***	−0.006***
	(0.001)	(0.001)	(0.002)	(0.002)
PE	−0.00002***	−0.00001	−0.00002**	−0.00003***
	(0.00001)	(0.00001)	(0.00001)	(0.00001)
DP	0.123***	0.136***	0.211**	0.213***
	(0.032)	(0.039)	(0.099)	(0.074)
CR	−0.001	−0.001	−0.001	−0.005**
	(0.001)	(0.001)	(0.001)	(0.002)
Constant	−0.001***	−0.004***	−0.005***	−0.006***
	(0.0003)	(0.0002)	(0.0003)	(0.0004)
Observations	84,748	81,942	77,927	69,922
R^2	0.365	0.363	0.308	0.280
Adjusted R^2	0.365	0.363	0.308	0.280
Note:			*p<0.1; **p<0.05; ***p<0.01	

Table A.7: Different FE models for 1 month returns

	Dependent variable:					
	Return1					
	(1)	(2)	(3)	(4)	(5)	(6)
Ret1	-0.028^{***}	-0.018^{**}				-0.029^{***}
	(0.009)	(0.007)				(0.007)
Ret3	-0.003	-0.008^{*}				
	(0.006)	(0.005)				
Ret6	0.004	0.007				
	(0.006)	(0.006)				
Ret12	-0.001	-0.001				
	(0.001)	(0.001)				
BM	0.018^{*}	0.018^{*}	0.015^{***}	0.020^{**}	0.015^{**}	0.015^{**}
	(0.010)	(0.010)	(0.006)	(0.009)	(0.006)	(0.006)
Size	-0.0004^{***}	-0.0002^{***}	-0.0004^{***}	-0.0004^{***}	-0.0004^{***}	-0.0004^{***}
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
PE	-0.00000	-0.00000				
	(0.00000)	(0.00000)				
DP	0.016	0.015			-0.0002	
	(0.015)	(0.014)			(0.014)	
CR	-0.001^{***}	-0.001^{***}		-0.001^{**}		
	(0.0005)	(0.0005)		(0.0004)		
Market1		1.161^{***}				
		(0.017)				
Obs.	83,364	83,364	125,882	97,320	125,882	125,882
R^2	0.012	0.109	0.010	0.011	0.010	0.010
Adj. R^2	0.012	0.107	0.009	0.011	0.009	0.010
Note:						*p<0.1; **p<0.05; ***p<0.01

Table A.8: Different FE models for 3 month returns

	Dependent variable:					
	Return3					
	(1)	(2)	(3)	(4)	(5)	(6)
Return	-0.040^{***}	-0.030^{***}				-0.042^{***}
	(0.010)	(0.009)				(0.009)
Ret3	0.013	-0.004				
	(0.014)	(0.011)				
Ret6	-0.004	0.011				
	(0.018)	(0.017)				
Ret12	-0.005	-0.004^{**}				
	(0.003)	(0.002)				
BM	0.043^{*}	0.041^{*}	0.037^{***}	0.051^{***}	0.037^{***}	0.037^{***}
	(0.023)	(0.022)	(0.013)	(0.019)	(0.013)	(0.013)
Size	-0.001^{***}	-0.001^{***}	-0.001^{***}	-0.001^{***}	-0.001^{***}	-0.001^{***}
	(0.0003)	(0.0002)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
PE	-0.00000	-0.00000				
	(0.00001)	(0.00001)				
DP	0.070	0.064			0.019	
	(0.044)	(0.041)			(0.035)	
CR	-0.004^{***}	-0.004^{***}		-0.003^{**}		
	(0.001)	(0.001)		(0.001)		
Market3		1.284^{***}				
		(0.024)				
Observations	83,347	83,347	125,860	97,303	125,860	125,860
R^2	0.023	0.129	0.019	0.023	0.019	0.020
Adjusted R^2	0.023	0.127	0.019	0.023	0.019	0.019
Note:				$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$		

Table A.9: Different FE models for 6 month returns

	Dependent variable:					
	Return6					
	(1)	(2)	(3)	(4)	(5)	(6)
Return	−0.014	−0.018				−0.041
	(0.022)	(0.022)				(0.025)
Ret3	0.017	0.002				
	(0.012)	(0.011)				
Ret6	−0.022	0.016				
	(0.035)	(0.030)				
Ret12	−0.010**	−0.013***				
	(0.004)	(0.004)				
BM	0.101***	0.091***	0.073***	0.122***	0.072***	0.072***
	(0.026)	(0.024)	(0.021)	(0.024)	(0.022)	(0.021)
Size	−0.002***	−0.001***	−0.002***	−0.002***	−0.002***	−0.002***
	(0.001)	(0.0003)	(0.001)	(0.001)	(0.001)	(0.001)
PE	−0.00001	0.00000				
	(0.00002)	(0.00002)				
DP	0.196*	0.178*			0.042	
	(0.110)	(0.106)			(0.070)	
CR	−0.009***	−0.008***		−0.007**		
	(0.003)	(0.003)		(0.004)		
Market6		1.393***				
		(0.036)				
Observations	79,268	79,268	120,365	93,054	120,365	120,365
R^2	0.041	0.156	0.027	0.039	0.027	0.027
Adjusted R^2	0.040	0.153	0.027	0.038	0.027	0.027

Note: *p<0.1; **p<0.05; ***p<0.01

Table A.10: Returns for 25 portfolios

Size quintiles	Book-to-market ratio quintiles Low	2	3	4	High	Low	2	3	4	High
Panel A: Value weighted portfolios										
	Mean returns					Standard deviations				
Small	1.03	1.09	1.23	1.14	0.95	6.58	4.80	5.39	5.07	6.43
2	1.35	1.19	1.33	1.51	1.93	7.78	5.01	5.73	6.76	6.40
3	0.93	1.03	1.33	1.68	1.35	5.41	4.50	4.92	5.70	6.69
4	1.71	1.52	1.11	1.66	1.26	4.55	4.79	5.56	5.59	6.13
Big	1.46	1.33	1.33	1.31	1.84	4.53	4.78	4.92	5.38	6.53
	t-statistics					p-values				
Small	1.44	2.07	2.09	2.04	1.35	0.15	0.04	0.04	0.04	0.18
2	1.59	2.18	2.13	2.05	2.77	0.12	0.03	0.04	0.04	0.01
3	1.58	2.10	2.47	2.70	1.84	0.12	0.04	0.02	0.01	0.07
4	3.44	2.91	1.84	2.73	1.88	0.00	0.00	0.07	0.01	0.06
Big	2.96	2.54	2.48	2.23	2.56	0.00	0.01	0.02	0.03	0.01
Panel B: Equally weighted portfolios										
	Mean returns					Standard deviations				
Small	1.43	1.23	1.41	1.65	2.52	5.25	4.72	5.61	5.34	7.24
2	1.48	1.35	1.33	1.81	2.35	6.35	5.11	5.96	6.85	7.41
3	1.59	1.15	1.65	1.45	1.80	5.07	4.73	5.45	5.21	6.76
4	1.61	1.74	1.29	1.83	1.42	4.46	4.73	5.35	5.94	6.27
Big	1.56	1.21	1.51	1.27	1.83	4.09	4.24	4.95	5.25	6.78
	t-statistics					p-values				
Small	2.50	2.38	2.31	2.82	3.19	0.01	0.02	0.02	0.01	0.00
2	2.14	2.42	2.04	2.41	2.91	0.04	0.02	0.04	0.02	0.00
3	2.88	2.22	2.77	2.56	2.45	0.01	0.03	0.01	0.01	0.02
4	3.32	3.37	2.21	2.83	2.07	0.00	0.00	0.03	0.01	0.04
Big	3.51	2.61	2.80	2.21	2.46	0.00	0.01	0.01	0.03	0.02

Source: author's computations.

Table A.11: 3, 6 and 12 Month Holding Period Returns for 25 portfolios

Size	Book-to-market ratio quintiles									
quintiles	Low	2	3	4	High	Low	2	3	4	High
Panel A: Value weighted portfolios										
	3 month holding period					6 month holding period				
Small	3.35	4.11	3.87	3.78	2.96	5.99	7.48	7.66	7.52	4.91
2	3.24	3.81	4.11	5.02	5.95	6.13	7.68	8.02	9.32	11.40
3	3.70	3.71	4.50	4.67	4.39	6.84	8.00	9.07	9.31	7.80
4	5.52	4.29	3.66	4.76	3.66	11.19	8.46	7.84	8.96	7.29
Big	4.58	4.35	4.16	4.30	5.49	9.57	8.58	8.44	8.93	10.78
	12 month holding period									
Small	8.69	13.74	14.27	14.89	10.70					
2	13.79	15.66	16.01	17.30	22.98					
3	14.63	16.37	17.93	18.42	13.72					
4	22.66	16.94	15.59	15.62	13.84					
Big	17.68	15.87	16.79	18.08	17.50					
Panel B: Equally weighted portfolios										
	3 month holding period					6 month holding period				
Small	3.96	4.70	4.71	5.04	7.58	7.43	8.96	9.41	9.91	14.13
2	4.45	4.30	4.14	5.88	7.31	9.28	8.52	8.20	10.74	14.40
3	5.10	4.04	5.03	4.58	5.32	9.18	8.47	10.04	9.29	10.32
4	5.17	4.87	4.20	5.43	4.30	10.79	9.57	8.71	10.35	8.47
Big	4.81	4.09	4.55	4.33	5.47	9.79	8.04	9.14	8.83	10.79
	12 month holding period									
Small	14.95	17.13	17.11	18.94	27.13					
2	19.76	16.89	16.64	19.08	27.57					
3	19.06	17.19	19.89	18.58	19.02					
4	21.78	18.86	16.98	18.87	16.16					
Big	18.67	15.39	17.81	18.36	18.00					

Source: author's computations.

Table A.12: Time series regression - market model

Size quintiles	Book-to-market ratio quintiles									
	Low	2	3	4	High	Low	2	3	4	High
	b					t(b)				
Small	1.16	0.91	1.07	1.02	1.22	13.28	17.21	24.40	25.72	18.86
2	1.33	0.98	1.14	1.38	1.21	12.03	24.31	24.88	28.05	17.13
3	0.83	0.87	0.95	1.09	1.27	9.27	20.59	19.19	18.72	18.31
4	0.89	0.97	1.09	1.07	1.15	22.61	30.71	21.63	19.01	17.37
Big	0.80	0.91	0.97	1.05	1.15	14.43	18.32	22.78	20.78	13.69
	R-squared									
Small	0.69	0.79	0.88	0.89	0.81					
2	0.64	0.88	0.88	0.91	0.78					
3	0.52	0.84	0.82	0.81	0.81					
4	0.86	0.92	0.85	0.82	0.79					
Big	0.72	0.81	0.86	0.84	0.70					

Source: author's computations.

Table A.13: Time series regression - 3 factor model

Size quintiles	Book-to-market ratio quintiles									
	Low	2	3	4	High	Low	2	3	4	High
	b					t(b)				
Small	1.15	0.84	0.98	0.93	0.94	13.21	14.75	22.15	22.83	23.44
2	1.02	0.98	1.01	1.24	1.04	9.61	21.17	22.36	25.38	16.04
3	0.98	0.89	0.97	1.13	1.17	11.17	18.07	17.56	18.30	16.25
4	0.95	1.00	1.02	0.97	1.05	21.59	30.43	19.31	17.83	14.99
Big	1.03	1.04	1.00	0.99	1.05	26.72	22.68	23.66	20.66	14.51
	s					t(s)				
Small	0.79	0.46	0.52	0.41	1.02	4.08	3.65	5.24	4.55	11.40
2	1.33	0.13	0.36	0.37	-0.01	5.63	1.21	3.59	3.40	-0.07
3	0.21	-0.11	-0.31	-0.52	-0.14	1.07	-1.04	-2.51	-3.80	-0.88
4	-0.09	-0.32	-0.10	-0.12	-0.06	-0.97	-4.36	-0.87	-1.01	-0.39
Big	-0.88	-0.71	-0.46	-0.29	-0.62	-10.25	-6.90	-4.93	-2.68	-3.79
	h					t(h)				
Small	-0.62	-0.04	0.07	0.16	0.70	-3.57	-0.31	0.80	1.94	8.80
2	0.59	-0.11	0.40	0.42	0.91	2.77	-1.23	4.47	4.34	7.01
3	-0.99	-0.01	0.13	0.26	0.66	-5.67	-0.08	1.22	2.11	4.59
4	-0.24	0.08	0.46	0.66	0.59	-2.70	1.17	4.34	6.07	4.23
Big	-0.50	-0.11	0.22	0.56	1.05	-6.54	-1.20	2.56	5.87	7.22
	R-squared									
Small	0.77	0.82	0.91	0.92	0.95					
2	0.76	0.88	0.92	0.93	0.87					
3	0.66	0.84	0.84	0.85	0.85					
4	0.88	0.94	0.88	0.88	0.83					
Big	0.90	0.88	0.90	0.90	0.84					

Source: author's computations.

Table A.14: Time series regression - fundamental model

Size quintiles	Book-to-market ratio quintiles									
	Low	2	3	4	High	Low	2	3	4	High
	s					t(s)				
Small	1.67	1.11	1.27	1.12	1.74	5.14	4.81	5.12	4.81	7.38
2	2.11	0.87	1.14	1.32	0.78	6.49	3.51	4.45	4.28	2.81
3	0.96	0.57	0.43	0.34	0.76	3.26	2.43	1.70	1.15	2.41
4	0.63	0.45	0.68	0.62	0.75	2.62	1.82	2.57	2.44	2.59
Big	-0.09	0.09	0.30	0.47	0.20	-0.37	0.34	1.20	1.86	0.67
	h					t(h)				
Small	0.23	0.59	0.80	0.85	1.40	0.81	2.89	3.65	4.10	6.73
2	1.34	0.61	1.16	1.35	1.68	4.67	2.79	5.12	4.94	6.81
3	-0.27	0.65	0.85	1.10	1.54	-1.03	3.18	3.79	4.22	5.53
4	0.47	0.82	1.22	1.38	1.38	2.20	3.80	5.22	6.14	5.42
Big	0.26	0.66	0.96	1.30	1.82	1.16	2.86	4.33	5.81	7.02
	R-squared									
Small	0.26	0.31	0.36	0.37	0.59					
2	0.48	0.23	0.40	0.38	0.43					
3	0.12	0.19	0.20	0.21	0.34					
4	0.14	0.20	0.32	0.38	0.34					
Big	0.02	0.10	0.22	0.34	0.40					

Source: author's computations.

Table A.15: Summary of strategies

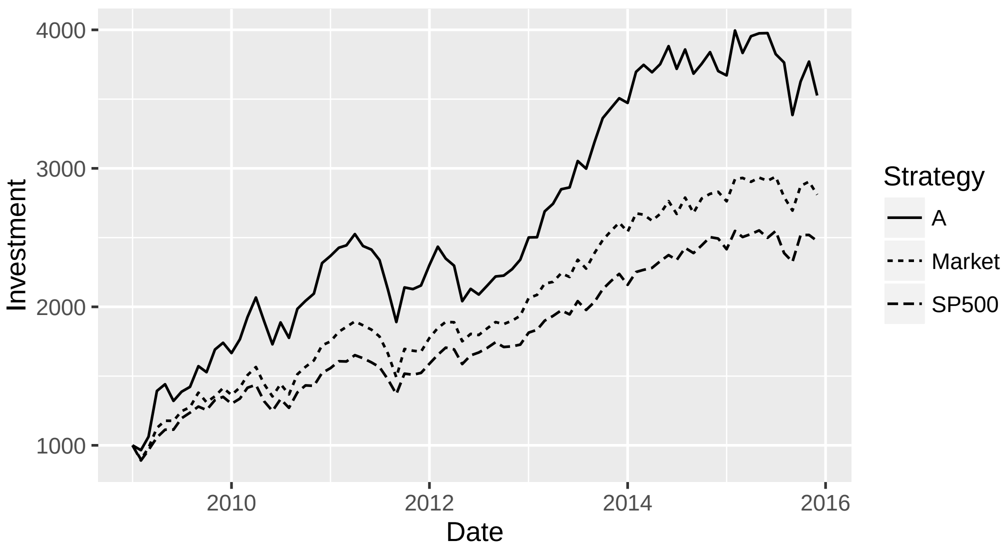
	Return		Quintile		
Strategy	Monthly	Yearly	BM	Size	Other condition
1	**1.93**	**1.97**	**5**	**2**	**-**
2	**3.14**	**1.78**	**5**	**1**	**30 smallest stocks**
3	**1.78**	**1.61**	**5**	**5**	**-**
4	0.84	0.91	5	1	30 stocks with smallest CR
5	0.79	0.49	5	1	30 stocks with highest PE
6	0.02	0.15	5	1	30 stocks with highest DP
7	0.22	0.97	5	-	30 stocks with highest BM
8	1.35	0.70	5	-	Random 30 stocks
9	**2.14**	**1.05**	**5**	**1**	**30 stocks with smallest Ret1**
10	-1.39	0.03	5	1	30 stocks with highest Ret1
11	1.42	0.89	5	1	30 stocks with smallest Ret3
12	0.84	1.39	5	1	30 stocks with highest Ret3
13	0.67	0.41	5	1	30 stocks with smallest Ret6
14	**1.79**	**1.82**	**5**	**1**	**30 stocks with highest Ret6**
15	0.10	-0.70	5	1	30 stocks with smallest Ret12
16	1.26	0.48	5	1	30 stocks with highest Ret12
17	0.08	0.03	-	-	30 stocks with smallest Ret1
18	-0.07	1.05	-	-	30 stocks with highest Ret1
19	1.06	0.83	-	-	30 stocks with smallest Ret3
20	0.35	0.63	-	-	30 stocks with highest Ret3
21	0.08	0.16	-	-	30 stocks with smallest Ret6
22	0.65	0.51	-	-	30 stocks with highest Ret6
23	-0.37	-0.53	-	-	30 stocks with smallest Ret12
24	0.64	0.65	-	-	30 stocks with highest Ret12

Source: author's computations.

Appendix B

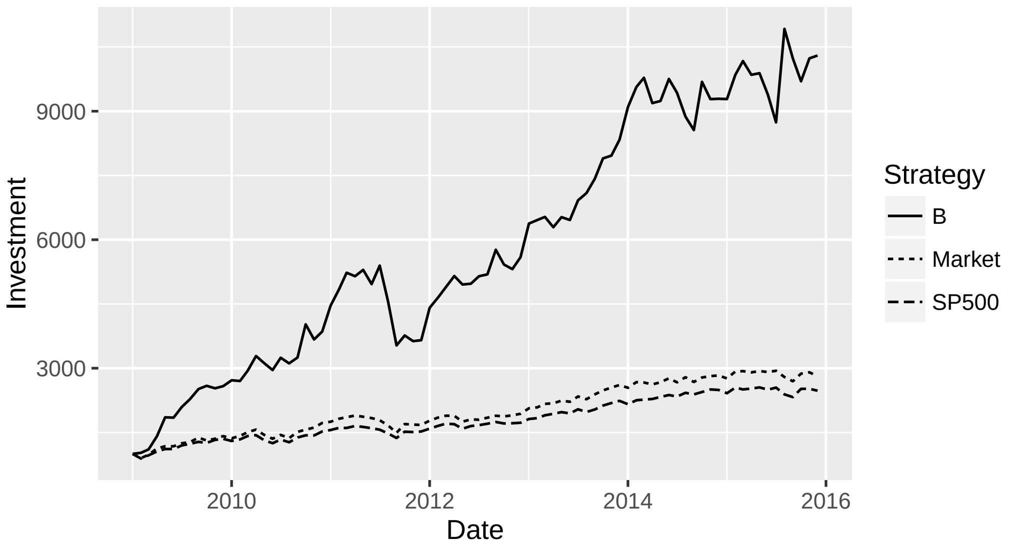
Figures

Figure B.1: Strategy A - Monthly Rebalanced



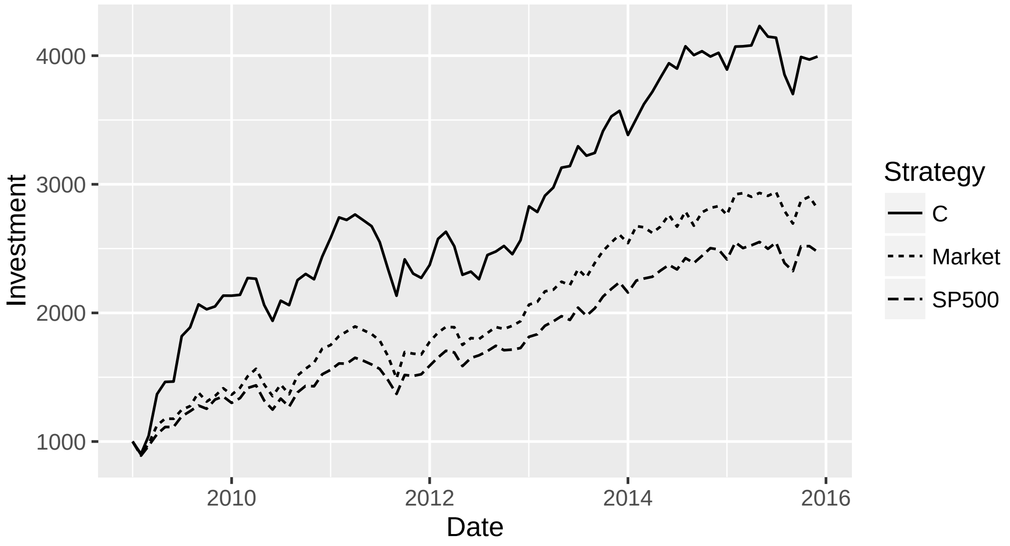
Source: author's computations.

Figure B.2: Strategy B - Monthly Rebalanced



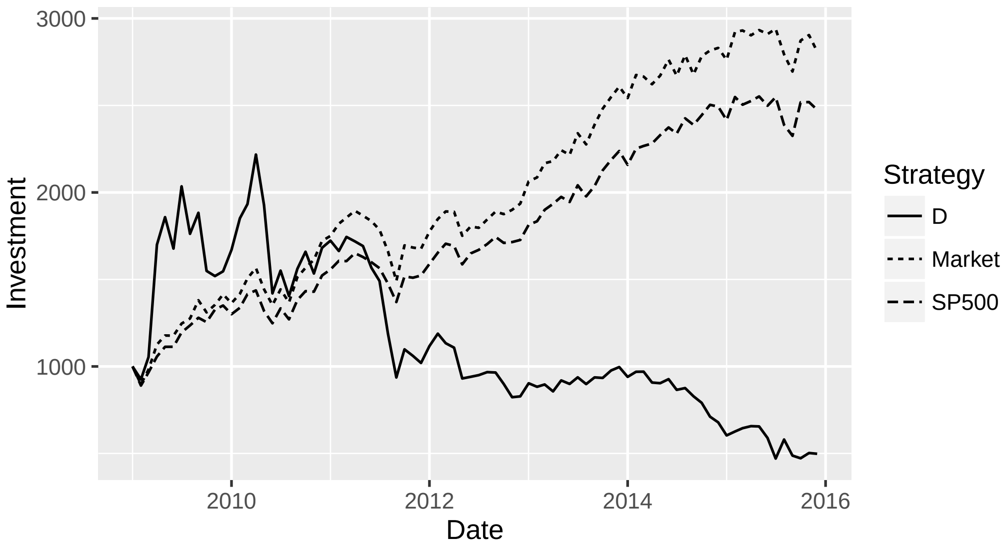
Source: author's computations.

Figure B.3: Strategy C - Monthly Rebalanced



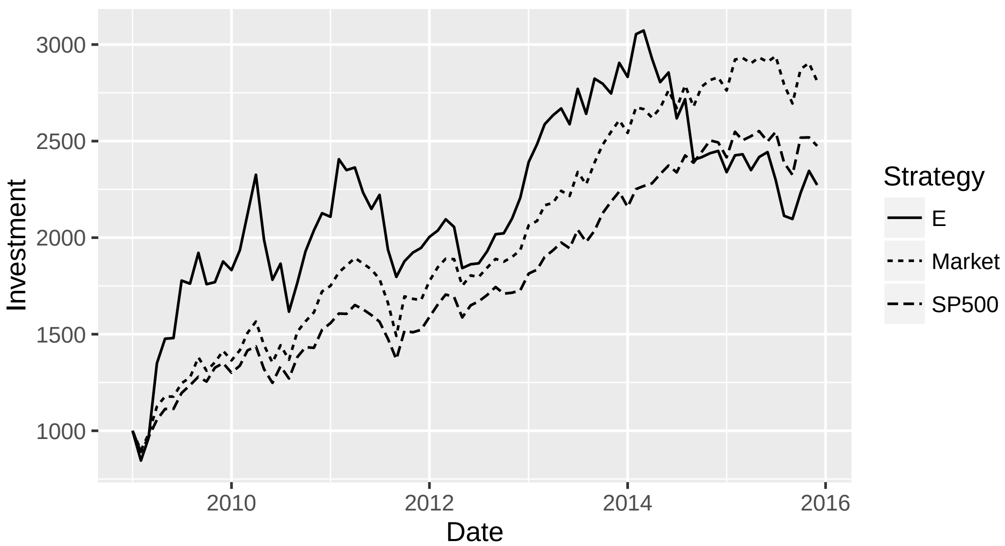
Source: author's computations.

Figure B.4: Strategy D - Monthly Rebalanced



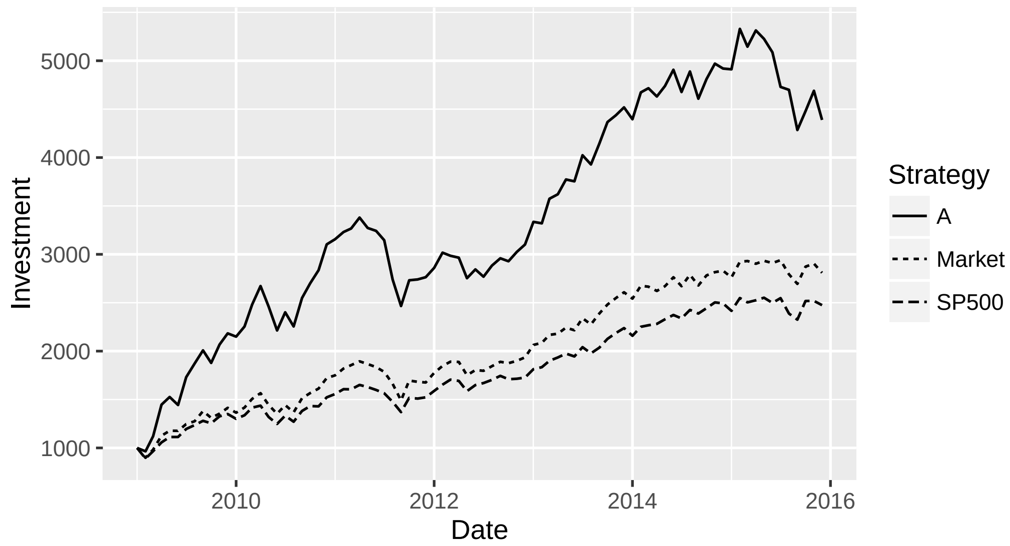
Source: author's computations.

Figure B.5: Strategy E - Monthly Rebalanced



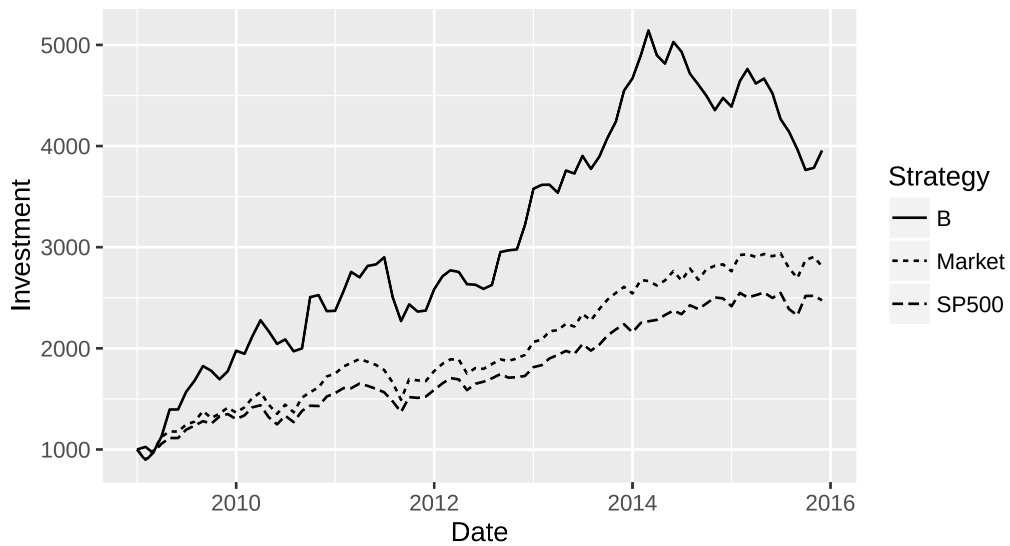
Source: author's computations.

Figure B.6: Strategy A - Yearly Rebalanced



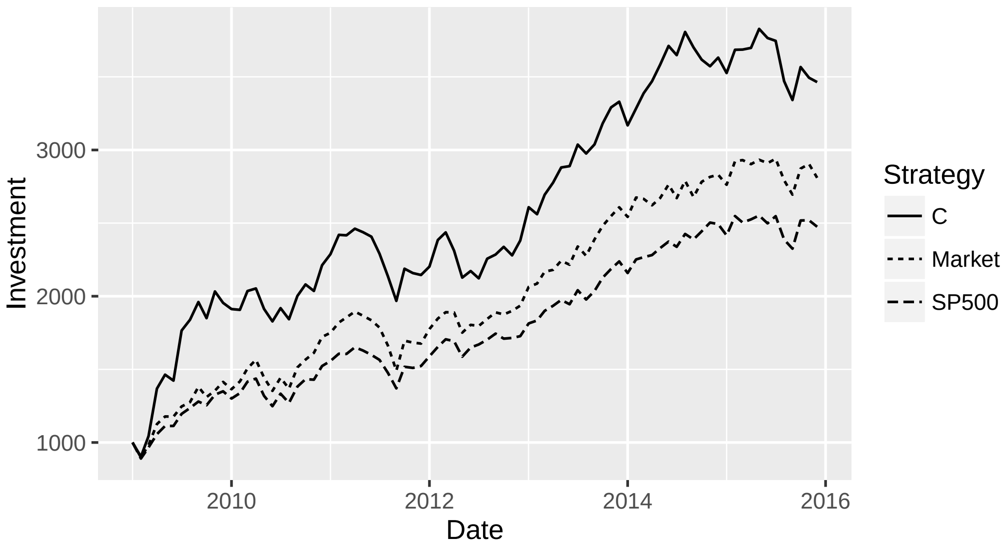
Source: author's computations.

Figure B.7: Strategy B - Yearly Rebalanced



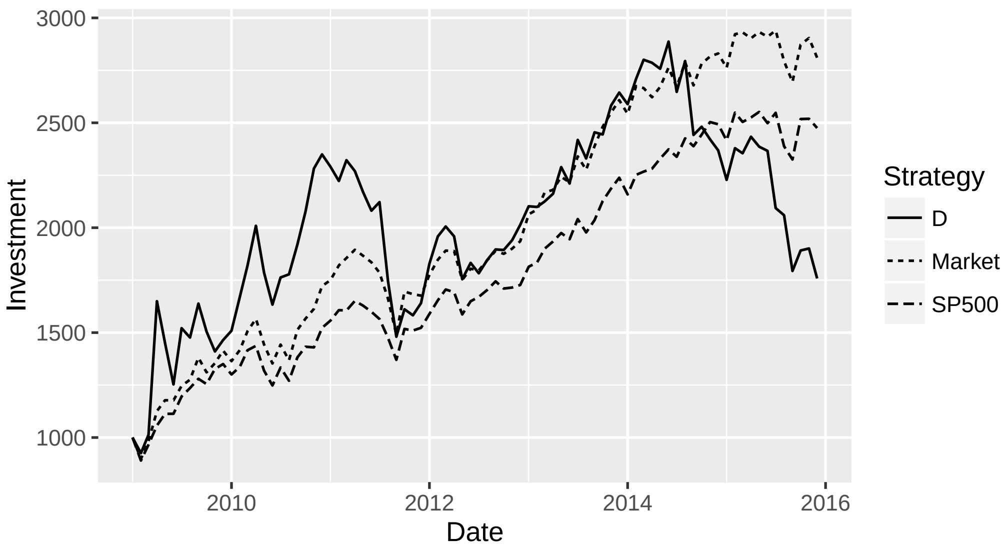
Source: author's computations.

Figure B.8: Strategy C - Yearly Rebalanced



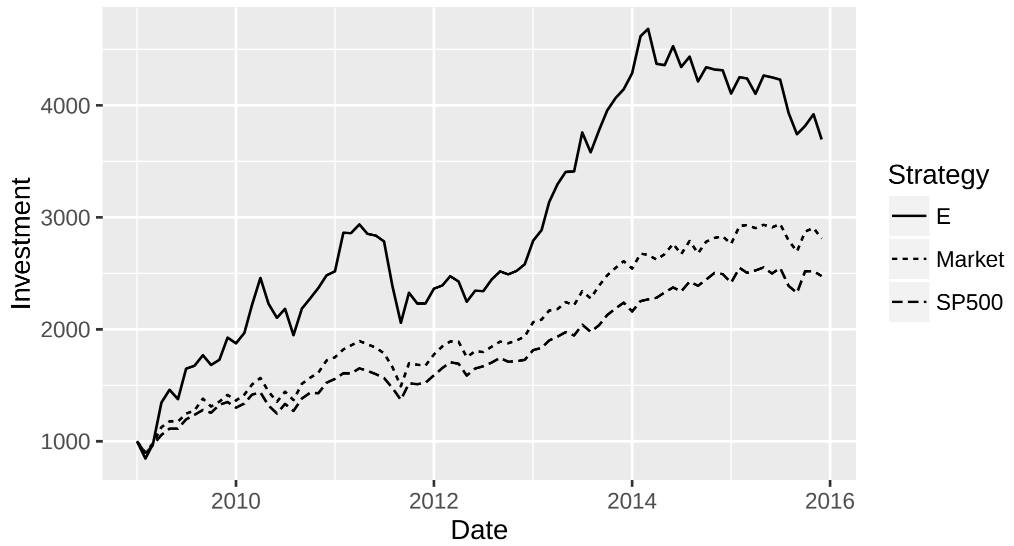
Source: author's computations.

Figure B.9: Strategy D - Yearly Rebalanced



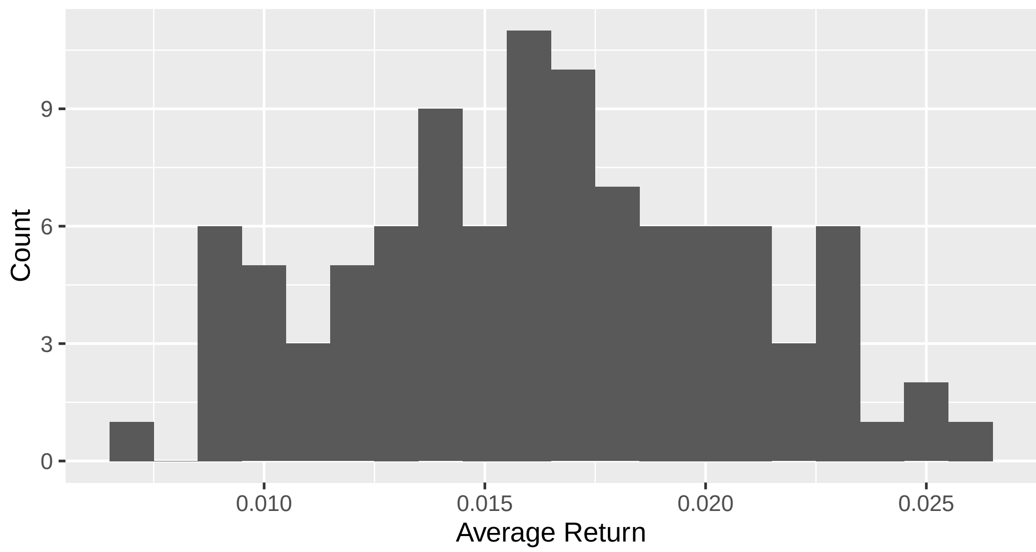
Source: author's computations.

Figure B.10: Strategy E - Yearly Rebalanced



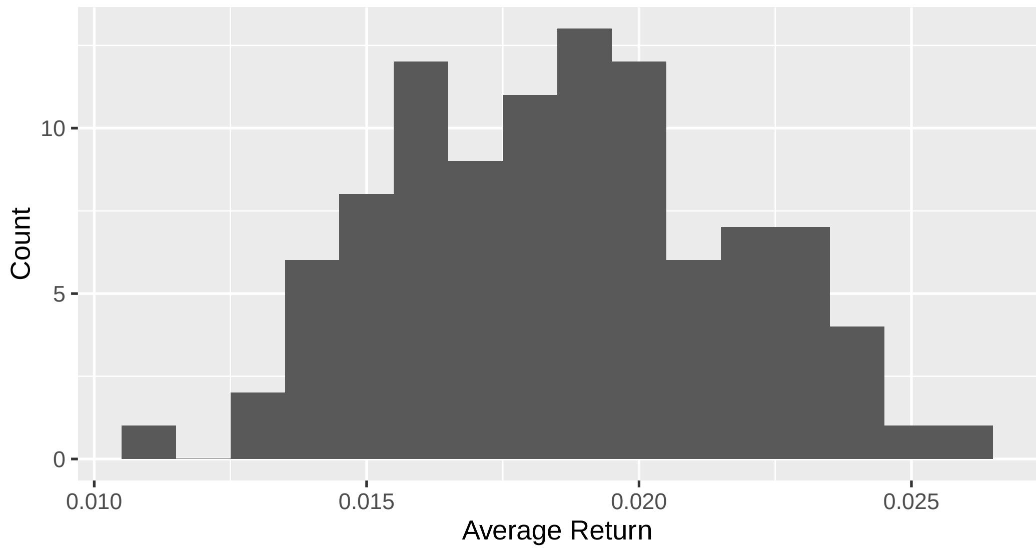
Source: author's computations.

Figure B.11: Strategy A (Monthly) - Bootstrap Histogram



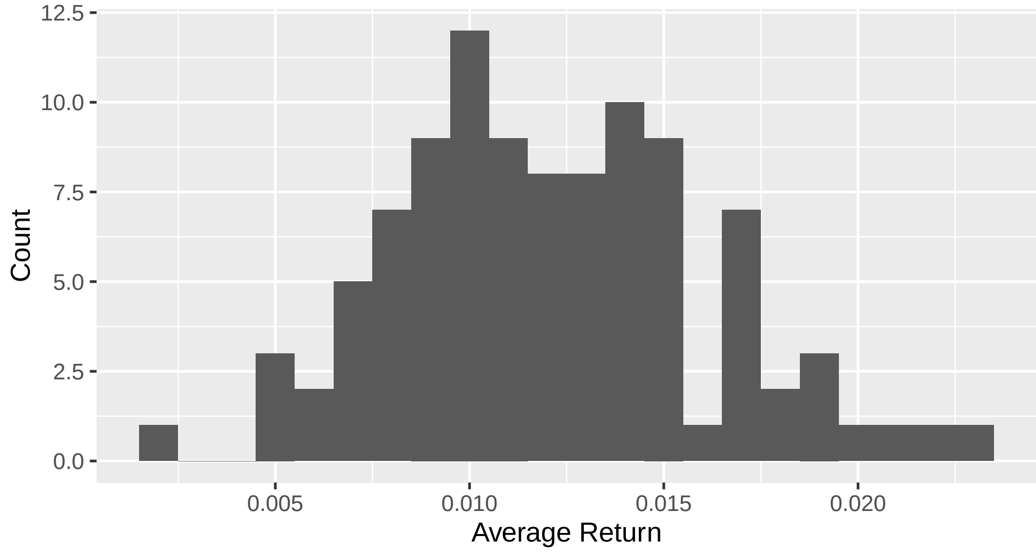
Source: author's computations.

Figure B.12: Strategy B (Monthly) - Bootstrap Histogram



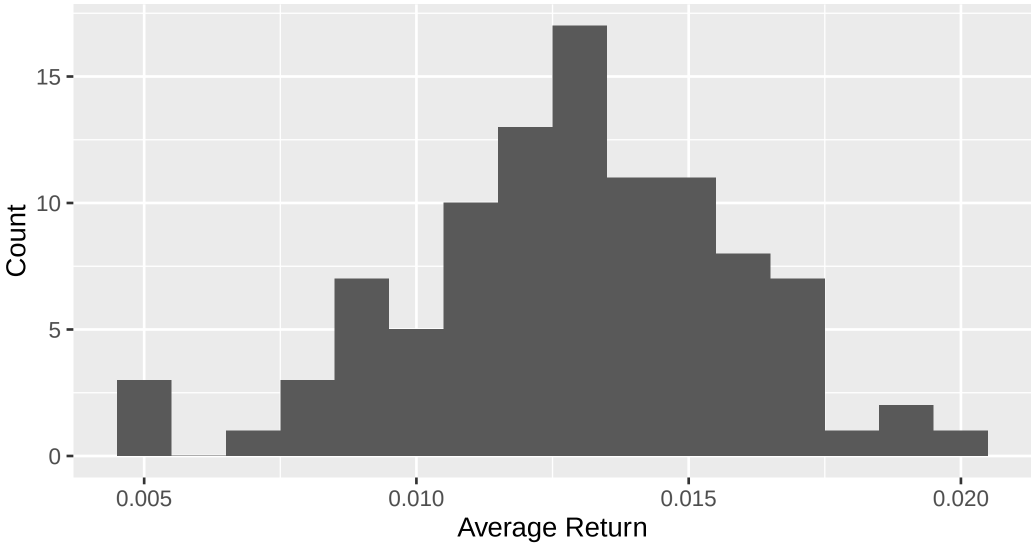
Source: author's computations.

Figure B.13: Strategy C (Monthly) - Bootstrap Histogram



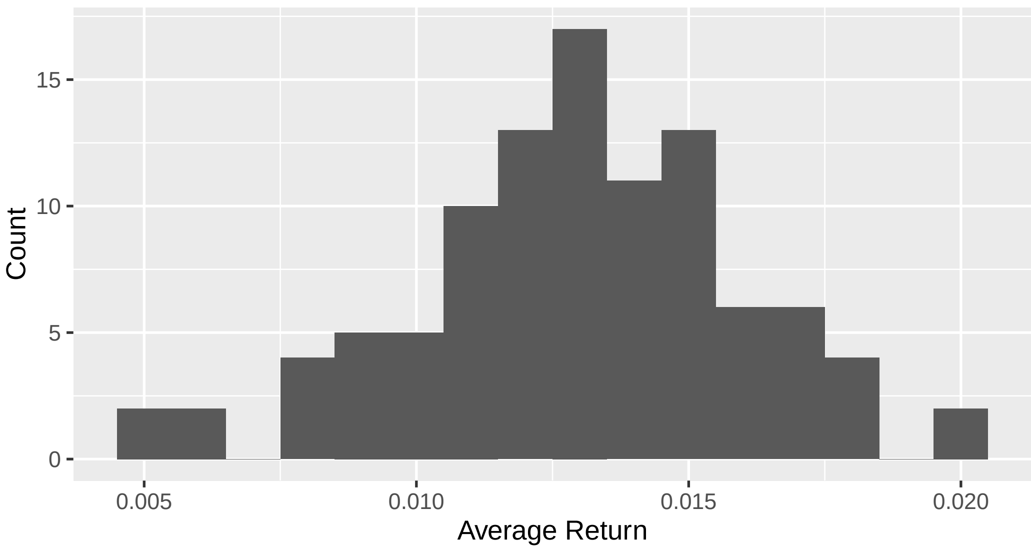
Source: author's computations.

Figure B.14: Strategy D (Monthly) - Bootstrap Histogram



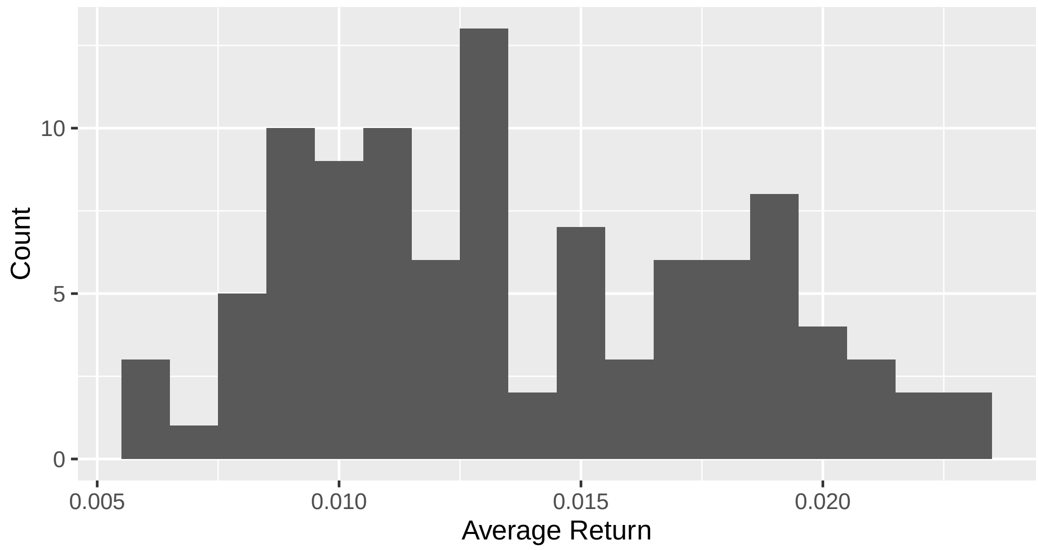
Source: author's computations.

Figure B.15: Strategy E (Monthly) - Bootstrap Histogram



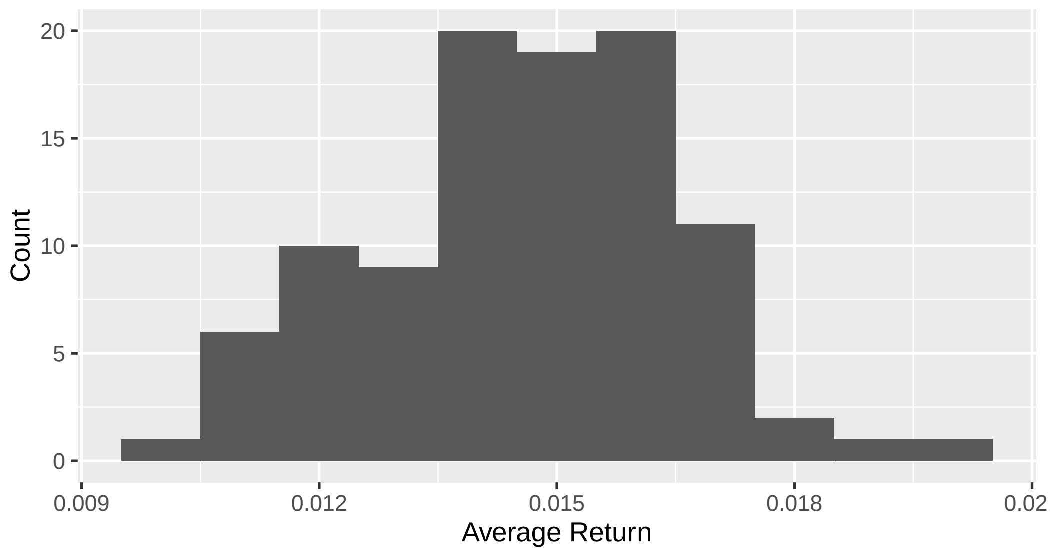
Source: author's computations.

Figure B.16: Strategy A (Yearly) - Bootstrap Histogram



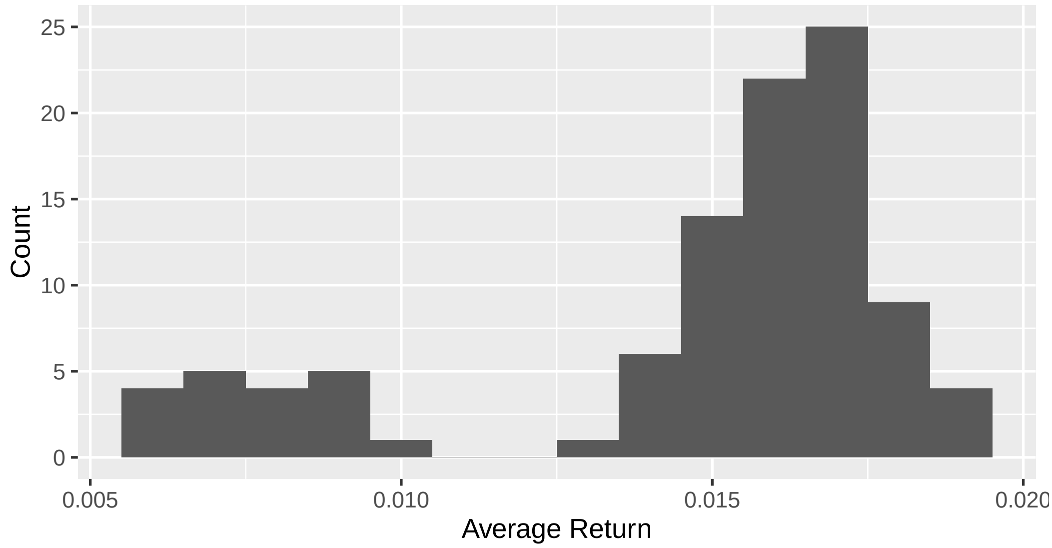
Source: author's computations.

Figure B.17: Strategy B (Yearly) - Bootstrap Histogram



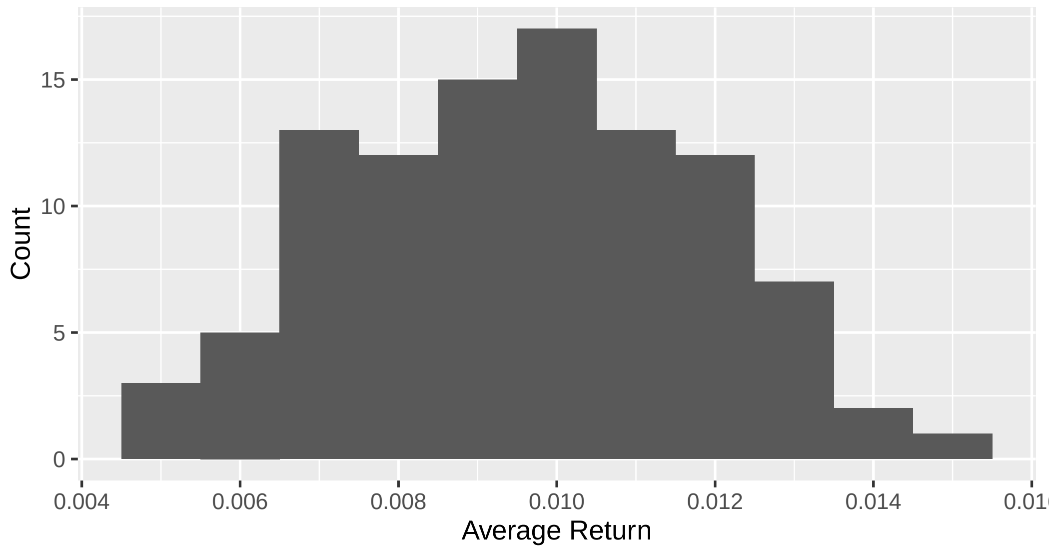
Source: author's computations.

Figure B.18: Strategy C (Yearly) - Bootstrap Histogram



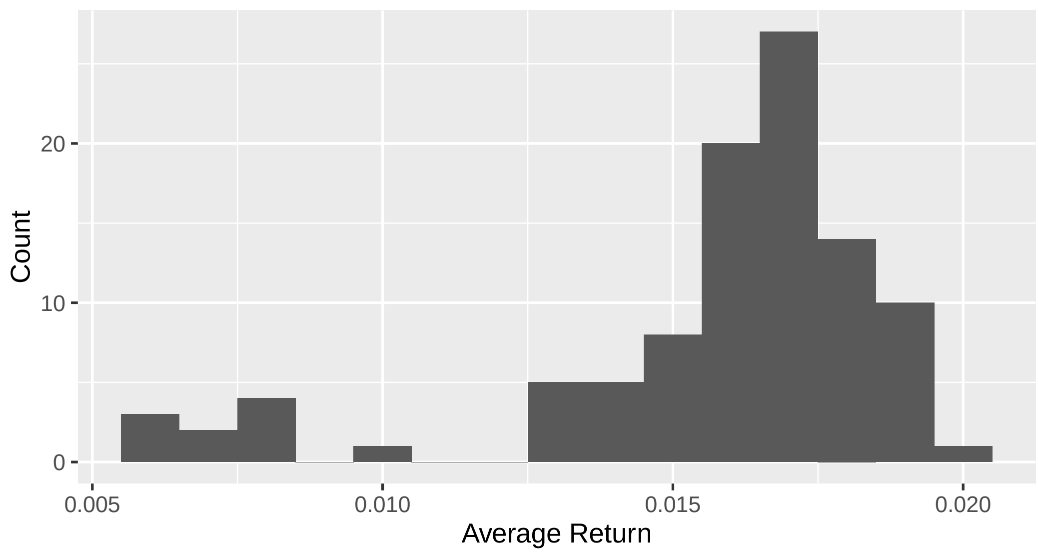
Source: author's computations.

Figure B.19: Strategy D (Yearly) - Bootstrap Histogram



Source: author's computations.

Figure B.20: Strategy E (Yearly) - Bootstrap Histogram



Source: author's computations.

Appendix C

Master Thesis Proposal

Master's Thesis Proposal

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

Author:	**Bc. Vojtěch Pižl**	Supervisor:	PhDr. Ladislav Krištoufek Ph.D.
E-mail:	vojtech.pizl@gmail.com	E-mail:	kristoufek@ies-prague.org
Phone:	603719304	Phone:	Line 312 (IES)
Specialization:	FFMaB	Defense Planned:	January 2016

Notes: The proposal should be 2-3 pages long. Save it as "yoursurname_proposal.doc" and send it to mejstrik@fsv.cuni.cz, tomas.havranek@ies-prague.org, and zuzana.irsova@ies-prague.org. Subject of the e-mail must be: "JEM001 Proposal (Yoursurname)".

Proposed Topic:

Algorithmic Fundamental Trading

Motivation:

In the last decade there has been rise in the field of algorithmic trading. Nevertheless most of the strategies seem to be focused on technical factors and short-term trading signals. Fundamental analysis that composes crucial part of value investing seems to be neglected and this has negative effect on the main function of the markets – raising capital for companies. Therefore, there is an opportunity to apply findings and conclusions from academic literature into the field of electronic trading.

One of the examples of such fundamental factor is book-to-market ratio. Fama and French [1993] have built the model (three factor model) that incorporates this ratio in the stock return prediction. Another author who examined the effect of BM ratio was J. Piotroski [2001] who presented simple accounting based fundamental strategy that generated 23% annual return. More recent research by Lyle and Wang [2014] uses also return on equity (ROE) in addition to BM ratio to model expected holding period returns. They documented highly significant predictive pooled regression slope for future quarterly returns of 0.86.

This thesis aims to apply what has been achieved in theoretical research into sensible price signals. Last but not least we will also endeavor to develop our own trading strategy based on fundamental factors.

Note: Practical part of the thesis will be externally consulted with Mr. Pravda from Pravda Capital.

Hypotheses:

1. BM ratio is a sensible determinant of future returns
2. ROE is a sensible determinant of future returns
3. There is an added value of using algorithms in value investing
4. Strategy based on fundamental factors outperforms technical based strategy

Methodology:

In the first part of the thesis, I will identify various factors that determine future stock returns. This will be done by research of existing academic contributions to this topic. Then, these factors will be used to develop various trading strategies. Theoretical models should serve as a starting point in strategy development using some adjustments. Afterwards, models will be back tested using historical data from Reuters Eikon database provided by IES (and by Pravda Capital if the IES source was not sufficient) and evaluated using performance ratios (such as Sharpe ratio). Finally, models will be calibrated by changing exogenous inputs to achieve the best result. Generally, the back testing process will be conducted as follows:

(i) Using some part of historical data to estimate the model parameters
(ii) Use model parameters in out-of-sample test (this might change according to strategy – it depends if the model parameters are fixed or re-estimated after some time period)

To better assess performance of fundamentally based strategies, I will also compare its performance with some commonly used technical strategies. Programming will be done via Matlab.

Expected Contribution:

This thesis aims to connect theoretical research in the field of fundamental analysis with the practical application of such methods in algorithmic trading. The challenge will be to employ already developed theoretical framework into models that could be used in practice by portfolio managers. The added value compared to researches by e.g. Lyle and Wang [2014] will be firstly testing whether proposed approach of using of ROE and BM ratios to model expected returns is possible (using different data, adjusted conditions, parameters, etc.) and secondly if this model is feasible for trading purposes. The second goal is to assess whether there is an added value of using algorithms for value investing purposes.

Outline:

1. Introduction and motivation:
    1.1. What is role of algorithmic trading in financial markets today
    1.2. Role of fundamental analysis
    1.3. Differences between fundamental and technical analysis and implications for trading
2. Literature review
3. Theoretical background
    3.1. Fundamental factors and strategies
    3.2. Technical factors and strategies
4. Data description
5. Methodology
6. Results
    6.1. Implementation of strategies
    6.2. Backtesting
    6.3. Performance comparison and selection of best strategies
7. Conclusion

Core Bibliography:

1. Asness, C., 2013. Value and momentum everywhere. *The Journal of Finance*, 68(3), pp.929–985.
2. Baker, N.L. & Otterloo, V., 1996. Commonality In The Determinants Of Expected Stock Returns. *Journal of Financial Economics*.
3. Boonenkamp, U., Kempf, A. & Homburg, C., 2009. Fundamental information in technical trading strategies. *CFR Working Papers*. Available at: http://ideas.repec.org/p/zbw/cfrwps/0812.html.
4. Fama, E.F. & French, K.R., 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33, pp.3–56.
5. Fama, E.F. & French, K.R., 2008. Dissecting Anomalies. *The Journal of Finance*, 63(4), pp.1653–1678.
6. Lyle, M.R. & Wang, C.C.Y., 2014. The Cross Section of Expected Holding Period Returns and their Dynamics : A Present Value Approach The Cross Section of Expected Holding Period Returns and their Dynamics : A Present Value Approach.
7. Lyle, Matthew R. and Wang, Charles C. Y., The Cross Section of Expected Holding Period Returns and Their Dynamics: A Present Value Approach (June 19, 2014). *Journal of Financial Economics*, Forthcoming.
8. Piotroski, J.D., 2001. Value Investing: The Use of Historical Financial Statement Information to Separate Winners from Losers. Available at: http://papers.ssrn.com/abstract=249455 [Accessed December 28, 2014].

Author **Supervisor**