LUND UNIVERSITY
School of Economics and Management

Department of Business Administration
Master thesis
BUSM36
June 2010

Analyst Misestimations and the Predictability of Stock Returns

Supervisor:
Måns Kjellsson

Authors:
Axel Gyllenkrok
Joel Ottosson

Abstract

Title	Analyst Misestimations And The Predictability Of Stock Returns
Authors	Axel Gyllenkrok Joel Ottosson
Supervisor	Måns Kjellsson
Keywords	Efficient markets, Behavioral Finance, Overoptimism, Financial ratios, Analyst forecasts, Stock Return Predictability, CAPM, Multi-factor models, Excess returns, Momentum effect, Book-to-market.
Purpose	The purpose of this thesis is to examine whether the predictability of stock returns based on stock market and accounting variables is related to analyst misestimations of future profits. The thesis aims to explore possible linkages between overoptimism in analyst earnings estimates and the predictability of returns, which may, if they exist, help explain excess returns from a mispricing perspective.
Methodology	The research sample is based on equity data from the NYSE and NASDAQ, with returns collected for the 1994-2009 period. The data is analyzed using portfolio sorts and Fama-Macbeth regressions.
Conclusion	Overoptimism in analyst forecasts is predictable and related to the predictability of stock returns based on stock market and accounting data. This provides support for mispricing explanations for excess returns from stock market anomalies. The momentum effect seems to be caused by lags in investor reactions to negative information, as the effect seems to be related both with analyst coverage and the degree of overoptimism of future earnings.

Table of Contents

1. Introduction

1.1 Background

Investors have always sought to profit by choosing the highest returning shares at the lowest possible risk. As statistical tools and theoretical frameworks have improved, an ever-increasing amount of research has been put into discerning what drives stock prices, in order to find attractive investment opportunities as well as to develop corporate finance models.

The development of the Capital Asset Pricing Model (CAPM) by Sharpe (1964) and Lintner (1965) pioneered a new way of looking at risk and returns by introducing the view that higher returns could only be achieved by taking higher systematic risk. This has been the dominating view in finance ever since, however during the last 25 years the model has received much criticism as empirical studies showed stock return patterns unexplained by the model. For example, Banz (1981) found that low market capitalization stocks have had abnormally high risk adjusted returns, while Rosenberg et al. (1985) showed that there has been a positive correlation between returns and firms' book value divided by their market value of equity (book-to-market). Putting these two relationships together with the beta measure used in the CAPM, Fama and French (1992)(1993), introduced a three-factor model which showed superior stock return predictability compared to the CAPM.

Due to the increasing ease of mining financial databases for relationships between firm variables and stock returns, there is an ever growing number of factors for which a correlation with stock returns have been observed; some of the most prominent include; historical stock returns (Jegadeesh & Titman, 1993), accruals (Sloan, 1996) and asset growth (Cooper, Gulen, & Schill, 2008). As with the Fama and French factors, it is debated what causes the relation between these factors and stock returns. It is also unclear whether the effects have a common underlying source, or if they are driven by fundamentally different factors. Fama and French (2006) argue that it is a problem that most research treat the effects as isolated anomalies, resulting in a large number of factors that could explain stock returns but poor understanding of the actual sources of these relationships as the knowledge of the correlation structure among the factors is limited.

Fama and French (2006)(2008) argue that the observed effects predicting stock returns are consistent with the view of stocks being rationally priced based on risk. The reason for this is that the factors also tend to be correlated with expected future cash flows, which makes it impossible to, by looking only at the relation between the factor and the return, rule out the possibility that the return differences are due to different degrees of risk. Some researchers disagree with this view, arguing that investor behavioral biases make stock prices deviate substantially from what could be considered rational based on pure risk considerations.

Lakonishok et al. (1994) explain the predictive power of market multiples on future stock returns by the tendency of investors to overestimate the future prospects of growth stocks while underestimating the future of value stocks. This is caused by investors believing that historical trends will continue in the future, resulting in prices of historical growth stocks to decline because of future negative surprises while the stock prices of value stocks increase as investors tend to become positively surprised. Further, Barberis et al. (1998) argue that investors may have trouble interpreting which historical information can be used to predict the future and which cannot.

One approach to exploring possible linkages between stock return predictability and behavioral biases is to look at deviations between stock market expectations and actual outcomes. La Porta (1996) observes a negative relation between analyst long run earnings growth estimates and future returns, suggesting that analysts are overly optimistic about firms with strong historical trends and overly pessimistic about firms with weak historical records. There also seems to be a relationship between the dispersion of analyst forecasts and future returns; higher dispersion leading to lower returns (Diether, Malloy, & Scherbina, 2002) as well as a higher predictability of stock returns based on historical information when the analyst coverage is low (Nagel, 2005). This suggests that there may be considerable information in the analyst data which possibly can provide a behavioral explanation for some of the relations between historical data and future returns.

1.2 Purpose

The purpose of this thesis is to examine whether the predictability of stock returns based on stock market and accounting variables is related to analyst misestimations of future profits. There is a disagreement among researchers whether the relation between excess returns and some variables are due to risk or mispricing. In order to contribute to this field of research, this thesis aims to explore possible linkages between overoptimism in analyst earnings estimates and the predictability of returns, which may, if they exist, help explain excess returns from a mispricing perspective.

In order to investigate whether earnings misestimations is contributing to stock return effects in a systematic manner, a comprehensive empirical review of the relations between stock return predicting variables, subsequent stock returns, and overoptimism is conducted. A secondary aim of this thesis is to increase the understanding of the relations between the examined variables and stock return patterns, which is done by including a large number of factors for which correlation with stock returns has been observed in previous research. The substantial number of variables will also make it is possible to compare the effects of overoptimism on a multitude of factors, as well as exploring the characteristics of the variable cross-sections, in terms of relations to other variables. This integrated approach, looking at a large number of variables at the same time, rather than only at a few variables, is favored by Fama and French (2006) as it makes it easier to spot inter-variable dependencies that may be crucial to consider when interpreting the underlying mechanics of the stock return structure.

1.3 Research Questions

Our main research questions are:

1. Are analyst misestimations of future earnings related to stock returns?
2. Are the included factors correlated with stock returns, and if so, is this due to analyst misestimations of future earnings?
3. Are the analyst misestimations predictable, in size and magnitude, so that returns can be predicted in a systematic manner due to analyst over or under optimism?

1.4 Delimitations

The dataset used is derived solely from the NYSE and the NASDAQ, with future returns from the 1994-2009 period. Financial, real estate and insurance companies are excluded from the sample, as are firms without Worldscope and IBES data codes (see section 3.2 for more information on this). The availability of variable values in the databases differs across companies, resulting in a substantial number of firms being excluded in some calculations.

1.5 Thesis outline

The second chapter presents the theoretical framework that the thesis is based on as well as previous relevant empirical findings. The third chapter outlines the research approach followed, including the data collection and structuring methodology. In the fourth chapter the empirical findings are presented, and the results discussed. In the final chapter the thesis is concluded by a summary of the main results.

2. Theoretical framework

2.1 The Capital Asset Pricing Model

The traditional view on the predictability of stock returns is dominated by the Capital Asset Model (CAPM) by Sharpe (1964) and Lintner (1965). The CAPM states that differences in returns across stocks over time only can be explained by different exposure to systematic risk. Thus, the only stock-specific variable needed to calculate expected returns is the beta value, so that:

$$E[R_i] = r_f + \beta_i * (E[R_M] - r_f)$$

Even though the model is appealing both in a logical and an aesthetical sense, it is not strongly supported by empirical research (Fama & French, 1992). While some researchers believe that this is mainly due to difficulties in firstly defining the concept of systematic risk, and secondly estimating this risk in an accurate fashion (Campbell & Vuolteenaho, 2004), others argue that the whole logic behind the CAPM is completely out of sync with market realities and that it is due time to focus on models that are better supported by empirical data (Haugen & Baker, 2010).

2.2 The Fama and French three-factor model

The three-factor model (FF3M) by Fama and French (1992) (1993) is one of the most prominent models in asset pricing research. The model combines the beta used in CAPM with a market value (MV) and a book-to-market (BM) factor. The reason for including these factors in the model is the strong relations between their historical values and subsequent stock returns.

The notion that stocks with low market capitalization tend to outperform stocks with high market capitalization is widely spread both in asset pricing research and in the investor community. Banz (1981) observes that this holds even after adjusting for differing risk cross firm sizes. Further he finds that while the effect is generally present across a full sample of stocks, it is especially pronounced for the smallest stocks, which have particularly high returns.

Rosenberg et al. (1985) observe a positive relation between the book-to-market ratio (BM) (the book value of a firm divided by its market capitalization) and expected returns. They show that high BM stocks outperform stocks with low BM. This relation is confirmed by Fama and French

(1992), observing that the BM ratio, together with the market value have explanatory power on the cross-section of average stock returns for the 1963-1990 period. Just as in the case of the market value effect, the relationship remains even after adjusting for the return expected by the CAPM.

In the FF3M model, the exposure to the market beta, the market value factor (SMB) and the book-to-market factor (HML) determines the expected return:

$$E[R_i] = r_f + \beta_i^M * \left(E[R_M] - r_f\right) + \beta_i^{SMB} * E[R_{SMB}] + \beta_i^{HML} * E[R_{HML}]$$

Fama and French (1993) argue that the book-to-market and the market value effect on stock returns are due to these factors being proxies for risk, captured by the FF3M. The empirical support for this view is based firstly on the fact that the factor premiums have survived over a long time period. This is puzzling for proponents of the efficient market view since rational investors should arguably have taken advantage of the effects, which would have led to price adjustments resulting in the effects disappearing, but could be justified if the factors are actually related to risk for which investors require a premium. The second argument for the risk explanation of the FF3M is that there is substantial covariance between stocks in the same portfolio sorted on the FF3M factors, suggesting that those stocks are exposed to a common risk factor. Fama and French suggest that this risk factor, at least in the case of BM, is distress risk (Fama & French, 1993). If high BM firms are subject to high underlying distress risk, that could explain both the higher returns of these firms, the return covariance among them, and the fact that the premium has sustained over time, without relaxing market efficiency assumptions.

It is debated whether the premiums observed through the FF3M really are due to risk associated directly with the factor loadings, or if the model variables predict stock returns because they tend to be characteristics of stocks with certain expected returns because of other reasons. For example, Daniel and Titman (1997) argue that the FF factors are not risk factors per se, and put forward evidence that stocks in FF3M portfolios tend to comove not because of exposure to the model factors, but due to similar characteristics of the companies in the portfolio in terms of for example industry, geographic location and capital structure. However, despite its theoretical drawbacks, the FF3Ms strong ability to predict stock returns has made it a standard risk-adjustment model used in finance literature when studying the correlations between firm specific factors and the cross-section of stock returns.

2.3 Stock return predicting variables

After the initial discovery of the FF3M factors, subsequent research has found a large number of factors that predict the cross-section of future stock returns in a fashion similar to the FF3M model. In this section these factors are divided into a number of categories based on their source, however, the categories are not set in stone but created purely for didactic reasons, and the factors may well be categorized differently.

2.3.1 Stock trading and trends

De Bondt and Thaler (1987) study the relation between past and future stock returns, comparing previous winners to previous losers. Stocks that have performed poorly over a three to five year period outperform those that have performed well during the same period when looking at returns over the subsequent three to five year period. Jegadeesh & Titman (1993) on the other hand, find that when using a six to twelve month historical period instead, the old winners outperform the old losers during the coming six to twelve months. They argue that this short term momentum effect, and the longer term De Bondt & Thaler (1987) reversal effect, are not due to systematic risk but due to the market overreacting to long-term prospects of firms while under reacting to the short term prospects.

Building on the Jegadeesh & Titman (1993) momentum effect, Carhart (1997) expands the FF3M model by combining it with a momentum factor, thus forming a four factor model. He argues that even though the FF3M is superior to the CAPM in predicting future returns, the model is not doing a good job of predicting returns for portfolios of stocks with very high or low historical returns. Carhart (1997) shows that his model improves accuracy compared to both the FF3M and the CAPM.

Datar et al. (1998) measure the effect of liquidity on stock returns by looking at the stock turnover rate, i.e. the number of shares traded divided by the total number of shares outstanding, and find that liquidity is negatively related to stocks returns, even after controlling for market value, beta and B/M. This might be justified from a risk perspective since illiquid stocks might be more costly for investors to trade.

Ang et al. (2006) find that high idiosyncratic volatility in cross sectional stock returns result in low future returns, even after adjusting for liquidity, book to market, size and momentum effects. Although Nagel (2005), unlike Ang et al. (2006), uses aggregate volatility rather than only the idiosyncratic component, his findings supports those of Ang et al. (2006) by showing a correlation between high volatility and low returns, especially for stocks with constraints on short selling.

2.3.2 Market Multiples

Market multiples are ratios between stock market data, most commonly the share price, and accounting based numbers, e.g. earnings or sales. Market ratios have been used for a long time by investors in assessing the value of stocks, and are still commonly used to determine whether a certain stock is worth buying or not (Subrahmanyam, 2010).

Basu (1983) observes a relation between the price-to-earnings ratio (PE) of stocks and expected returns. NYSE listed stocks with low PEs earns higher risk-adjusted returns than high PE firms in the 1963-80 period. Further to this Lakonishok (1994) find that the cash flow-to-price ratio (CP) is related to average stock returns in the 1968-1989 period in a similar way as the above factors and Barbee (2008) conclude that the price-to-sales (PS) ratio also can predict future returns.

Research shows that there are varying degrees of correlation between the effects of market multiples. This is explained by the tendency of firms with high market value of equity compared to earnings to also on average have high of market value compared to sales, book value of equity or cash flow. It is possible that one or a few of the market multiples alone captures substantially all of the value/growth effect, while the other multiples exhibits a relation with returns solely because of a correlation with the main market multiple. Because of conflicting results in different studies, it is however not clear which one, if any, of the multiples that dominates the others. Fama & French (1992) show that the effect of PE on expected stock returns is already captured in the BM measure. When BM is controlled for, the PE ratio no longer exhibit any significant correlation with stock returns. However Lakonishok etl al. (1994) find that the BM effect is weaker than both the PE and CP effect and that CP is clearly the strongest factor among the three. In a more recent study, Barbee (2008) find that the PS ratio dominates the other three market multiples and that neither BM nor CF or PE exhibit any significant relation to expected returns.

2.3.3 Firm growth, investments and profitability

Future returns are also related to pure accounting data independent of the current market price. Sloan (1996) divides firm earnings into one cash flows and one accruals part and tests whether the market correctly prices both components. Accruals are defined as the part of earnings related to accrued income and expenses, or more specifically the change in current assets minus the change in current liabilities minus depreciation, all divided by average total assets. Sloan’s conclusion is firstly that the cash flow component of earnings is more persistent than the accrual component and secondly that firms with low accruals significantly outperform high accrual firms, providing support to the view that investors focus too much on earnings and fail to incorporate the different characteristics of the underlying cash flow and accrual component.

According to Fairfield et al. (2003), the growth in net operating assets (NOA) can be divided into two components; one consisting of the growth in accruals and the other of the growth in long-term NOA. Since both components have a similar effect on future growth, they argue that the accruals effect is part of a larger NOA effect. Firms with high NOA growth tend to have lower stock returns than firms with low NOA growth. Hirshleifer et al. (2004) find that the effect of NOA to total assets is stronger than that of growth in NOA, and that the effect extends three years into the future from the time of portfolio formation.

Capital investment (CI) is another accounting based factor that is correlated with returns according to a number of research papers. Abnormally high CI leads to significantly lower returns for five subsequent years (Titman, Wei, & Xie, 2004). Similarly, the growth in CI is also negatively related to returns, with higher returns for low CI growth firms than for high CI firms (Anderson & Garcia-Feijóo, 2006). While Titman et al. (2004) interprets the CI effect as investor under-reaction to manager empire building related over-investments, Anderson and Garcia-Feijóo (2006) believe that the effect is consistent with risk explanations.

While the capital investment effect is due to expansions of the asset side of the balance sheet, there is also a documented relation between expansions of the liability side and future stock returns. Loughran and Ritter (1995) document a negative relation between stock issuances and future returns. The effect is present for a period of five years after the issuance and is valid both for initial public offerings and seasoned equity offerings. There is also evidence of a reversed effect when reducing the number of shares outstanding. Ikenberry et al. (1995) show that open market repurchases of shares led to significantly higher returns during the 1980-1990 time period. They interpret this effect as a result of management exploiting private information in order to time repurchases when the stock is undervalued and that the market under estimate the significance of these actions. Evidence of this, they argue, is that the stock repurchase effect is strong for stocks with high BM values (more likely to be undervalued), while being non-existent for stocks with low BM values. Putting both the issuance and the repurchase effect together, Pontiff and Woodgate (2008) documents a general net stock issuance effect, where the net effect of stock issuances/repurchases predicts the cross-section of stock returns. The effect is present even after taking variables like BM, market value and momentum into account, but puzzlingly only in the post-1970 period.

Cooper et al. (2008) look at the growth in total assets and find that asset growth is a strong predictor of future stock returns. They also show that asset growth captures part of the accruals, NOA, capital investment and net stock issuances effect. They argue that this is because growth in assets is correlated with both NOA and CI, and also with stock and debt issuances since an asset increase leads to a corresponding increase of the equity and liability side of the balance sheet. Interestingly, the asset growth effect on returns exceeds that of the FF3M.

Haugen & Baker (1996) investigate the correlation between firm profitability and abnormal returns and find that both return on equity and return on assets are significant in predicting returns. Fama & French (2008) also find that return on equity is positively correlated with stock returns, but when excluding small firms the effect of profitability on returns is solely due to correlation with the FF3M model.

2.4 Behavioral explanations

There is a large field of research describing possible behavioral biases that may explain why the factors above are related to future stock returns. Lakonishok et al. (1994) put forward the idea that the predictive power of the market multiples are due to a tendency for investors to over extrapolate past growth when predicting the future. By doing this, investors tend to overestimate the growth of growth stocks and underestimate the growth of value stocks. The same concept is mentioned by Cooper et al. (2008) as the most likely explanation for the asset growth effect. Titman et al. (2004) suggest a different, but somewhat related, explanation for the capital investment effect. They believe that the negative relation between capital investments and future stock returns is due to investors underestimating manager's inclination to grow the size of their firms for their own rather than the shareholders gains.

Barberis et al. (1998) argue that investors tend to focus too much on previous earnings and returns and thus overreact to strong historical development. They explain this by outlining two psychological factors that cause under and overreactions; "representativeness" means that analysts interpret recent events as typical even though these events could not be used to predict the future, "conservatism" on the other hand implies that analysts are set in their ways and therefore slow at incorporating new information. Barberis et al. (1998) argue that analysts tend to focus more on the strength of new evidence than on the weight of the evidence; where strength refers to clearly noticeable evidence, with high impact on historical events, such as continuous high earnings; and weight is the statistical significance of the information in predicting the future, such as the degree of randomness of historical events. Tying this together with the empirical stock trends evidence, "conservatism" may create the Jegadeesh and Titman (1993) momentum effect through slow diffusion of new information, while "representativeness" may lead to overreactions causing the longer term De Bondt & Thaler (1987) reversals.

2.4.1 Analyst forecasts and returns

La Porta (1996) examine the relations between analyst forecasts and future stock returns and finds that there is a strong negative correlation between expected earnings growth and future returns. The stocks with low earnings growth expectations realize the highest future returns, without showing any sign of being more risky; in fact, the low expectation stocks tend to be more resistant than average to downward stock market movements. Another interesting finding is that the expected earnings growth is related to the BM factor, as firms with high growth expectations tend to have low BM values. This is quite obvious as the low BM value implies that substantial future value creation is necessary to justify the valuation, but taken together with the evidence that the high analyst expectations firms also tend to underperform. This finding is in line with the Lakonishok et al. (1994) hypothesis that the value/growth effect is caused by investor over-optimism about the future of historical growth stocks.

By observing the market reaction to earnings announcements, La Porta et al. (1997) find that 25-30% of the BM effect on stock returns can be explained by expectation errors as investors extend past growth rates too far into the future and therefore tend to underestimate the growth of high BM firms and overestimate the growth of low BM firms. La Porta et al. (1997) argue that the remaining effect can be attributed to investors gravitating towards low BM companies as those companies tend to be better known.

Easterwood & Nutt (1999) look at analyst forecasting errors. They find that in general, analysts tend to be overoptimistic, but react differently to different types of information. Instead of constantly over or underreacting, analysts tend to systematically overreact to positive information and underreact to negative. Easterwood & Nutt (1999) also examine analyst forecast revisions in order to spot their reaction to previous forecasting errors. In line with the aforementioned relation between the information content and over and under reactions, analysts tend to revise their future earnings forecast too little downward when the historical forecast error is negative and too much upward when the forecast error is positive. This evidence is consistent with the view of Daniel et al. (1998) who define two attributes that they believe contribute to investor forecasting errors. They argue that analysts are "overconfident", relying too much on their own private information, which results in systematic overreaction to private information and under reaction to public information. Further, analysts exhibit a considerable propensity to overreact to information that supports their already established views while underreacting to information that challenges their beliefs, a concept that Daniel et al. (1998) call "biased self-attribution".

Diether et al. (2002) show that companies with high dispersion of analyst forecasts earn abnormally low future returns when compared to similar companies with lower forecast dispersions. These findings are especially strong for smaller companies and for companies that have previously performed poorly. Nagel (2005) shows that the effect of dispersion is dependent on short selling constraints. If investors who believe that the stock is overpriced are inhibited from short selling, the stock will be overvalued due to the lack of any downward pressure on prices.

Hong et al. (2000) argue that stocks with low analyst coverage show a more distinct momentum effect due to slower information flow, making the stock price reaction to new information sluggish. Thus, the hedge return in taking a long position in a portfolio of strong momentum stocks and a short position in a portfolio of weak momentum stocks is stronger for stocks with low analyst coverage than for those with high coverage. Hong et al. (2000) argue that this is due to analyst coverage acting as a proxy for information flow, as less covered stocks publish less information thus aggravating analyst errors outlined in behavioural models such as Barberis et al. (1998).

3. Methodology

3.1 Hypotheses

The three main hypotheses of this study are outlined below.

Hypothesis 1

Analyst earnings misestimates are related to future returns, so that stocks with negative earnings surprises exhibit low returns and stocks with positive earnings surprises exhibit high returns.

Hypothesis 2

The studied variables are related to future stock returns, so that they can be used to predict returns.

Hypothesis 3

The studied variables are related to analyst misestimates of future earnings in a systematic way, so that erroneous estimates can explain part of the stock return predictability.

3.2 Overall data sample

The data used in this thesis is collected from ThomsonReuters financial databases. Datastream is used for stock market related variables, Worldscope for accounting data and IBES for analyst forecasts. Because the goal is to relate the results of this thesis to previous findings, data from the United States is used in order to avoid having to consider geographical discrepancies when relating the results to previous influential studies.

The sample consists of stocks listed on the New York Stock Exchange and NASDAQ during the 1994-06-30 to 2008-06-30 period. We include only securities that are listed as equities. The stocks must also be recorded in Datastream as primary securities, major securities, traded in US Dollars, and have either IBES or Worldscope data codes.

Special consideration is taken to decrease the risk of survivor biases in the data. The term survivor bias refers to the possibly systematic distortion that might be the result of not including stocks that have been delisted during the testing period. To reduce this risk, all stocks, both those that are still traded and those that were delisted during the survey period are included.

Financial, insurance and real estate firms are excluded from the sample. This is, in line with the reasoning of Fama & French (1992), due to the special characteristics of the financial statements of these firms that may make concepts like book value or operating assets different in terms of economical meaning compared to other industries. Thus, as in (Fama & French, 2008) all firms with Standard Industrial Classification Codes in the 6000-6999 range are excluded. Stocks were the sector is stated as “Unquoted Equities” are also excluded.

Only companies with fiscal year ends between 1 December and 31 March are included in the sample due to the risk of otherwise including not yet public data (see section 3.3.1 for more information on this). The fiscal year ends are obtained from the Worldscope database, and when this information is missing in Worldscope, from IBES. The negligible number of stocks for which year dates are missing in both the Worldscope and the IBES databases are removed from the sample in order to avoid distortions.

As the database coverage across the researched variables varies greatly, we chose to include stocks that have values on at least one of the researched variables in the basic sample and then simply exclude the stocks with missing values in each individual regression.

After the adjustments mentioned above, the total sample consists of 4,305 stocks. The number of stocks included in each year is outlined in chapter 4, table II.

3.3 Variables

3.3.1 Basic principles

Monthly data is collected over the 20 year period 1989-06-30 – 2009-06-30. This equates to 240 observations per stock for every variable. The future returns are calculated with the last of June as the base date over a 15 year period ranging from 1994-06-30 to 2009-06-30. Thus, the data from the 1989-06-30 – 1994-06-30 period is used solely for calculating some of the explanatory variables. In the end of June all the returns are reset, the portfolios rebalanced and the regressions recalculated.

The minimum requirement for a stock to be included in the calculations in June each year is that a return has been recorded for that year. As Datastream records a constant stock price for delisted stocks after the delisting, and thus zero returns, special consideration is needed in order to avoid including delisted stocks in the annual calculations. Hence, included in the sample each year are only stocks for which a return has been recorded, or, when the registered return is zero; where trading volume has been recorded in the month at the end of the return measurement period.

An issue to take seriously when using data from financial databases is the risk of attributing data to a point in time when it was not yet publicly available. In that case, correlations with future returns that did not exist in reality might appear if using information that was not available to the investor community at the time, and thus not priced by the market. In order to counter this risk of getting look-ahead biased results, it is important to pay particular attention to the methodology used by ThomsonReuters for the different data types. Some data type values are recorded in the database at the date they were announced; while others are recorded at the time period they refer to. Most stock market related variables (market value, liquidity, returns etc.) are in this thesis assumed to be publicly available to investors without delay at the date they refer to. Those variables are therefore generally obtained from the end of June each year. Most accounting related data is recorded in the database at a date before announcement. Thus, the end of December in the year t-1 is used for most of these variables in order to make sure that the information is available to the market. The principle of using the end of June for market data and the end of December for company data may however differ depending on the variable, because of certain characteristics of the data and reporting intervals (see section 3.3.2 for more information on the methodology used for calculating the individual variables).

The fact that the fiscal periods of some companies do not end at the end of the calendar year but at another date may result in a look-ahead bias if that leads to including variables containing not-yet public information. Since most of the variables are obtained from the end of December the year before, it may at a first glance seem like the risk of including not yet public information because of differing fiscal years is only present for companies with a fiscal year-end later than the last of December. However, due to the Datastream/Worldscope methodology in reporting accounting variables, the problem is probably more severe for firms with year-ends prior to December. If December values are used for those companies, the values for most accounting related variables actually refer to the numbers made public after the end of the following fiscal year, thus resulting in a severe look-ahead bias.

3.3.2 Included variables

Below follows a description of the methodology used for calculating the including variables. Datastream/Worldscope/IBES codes for each of the variables are displayed in brackets. In order to make it meaningful to compare the results from this thesis to those of earlier work, similar calculation principles are used when possible.

Stock returns

Stock returns are defined as the returns between the last trading day in June each year (time t) and a historical or future date. The returns used are the one, three and six month returns, as well as the one, two and three year returns, all of which originally are calculated from monthly returns.

The return index (RI) provided by Datastream is used for calculating stock returns. The index starts at 100 when the stock is listed after which further development corresponds to the theoretical return that would have been obtained if dividends were to be reinvested in the stock. (Ince & Porter, 2006) discuss the risk of getting incorrect results when using Datastream to calculate returns. They point out that some returns are wrong due to Datastream input errors, but revert to the correct values in the next month. In line with the recommendation of (Ince & Porter, 2006), the returns are adjusted for this in the following way; the return is set to missing if R_t or R_{t-1m} is greater than 300 % when at the same time $(1 + R_{t-1m}) * (1 + R_t) - 1$ is lower than 50 %.

Returns are, as previously mentioned, set to missing when both the return and the number of traded shares are zero. By doing this delisted stocks are excluded after the delisting. A manual review of the returns has also been conducted in order to remove unrealistic extreme returns from the sample. Although this exercise is somewhat arbitrary, there is a pattern in that unrealistically high returns almost always are concentrated to the first month of stock listing or to the delisting month.

$$R_{t+x} = \frac{Return\ Index\ (RI)_{t+x}}{Return\ Index\ (RI)_{t}} - 1$$

Book-to-market

The inverted book-to-market (BM) value provided by Datastream (PTBV) is used for the BM calculations. The PTBV variable is based on the book value per share for the same year, even though this value was not yet known at the time. This means that in order to avoid look-ahead biased results, a historical value must be used. Thus, for the returns starting in the end of June, the reported PTBV value for the end of December the year before is used. This is in line with the Fama & French (2008) methodology.

$$BM = \frac{1}{Price\ to\ book\ value\ per\ share\ (PTBV)_{t-6m}}$$

Cash flow-to-price

The same methodology as used for BM is used for the Cash flow-to-price (CP) value. The calculation is based on the value reported by Datastream in the end of December the preceding year in order to avoid taking not yet public information into account.

$$CP = \frac{1}{Price\ to\ cash\ flow\ per\ share\ (PC)_{t-6m}}$$

Earnings-to-price

The earnings-to-price ratio could be obtained directly from Datastream without any additional calculations (PE), but it is not appropriate to use this datatype when comparing the earnings-to-price effect to the other market multiples since the Datastream methodology for calculating the ratio is different to the one used for the other market multiples. The EPS number in the database is the last trailing 12 months EPS made public. Thus, in order to make the calculation consistent with those of the other market multiples, the ratio is calculated using the earnings per share recorded in the database in the end of March in year t (a time when most companies have made their year-end results public), divided by the share price in the end of December in year t-1.

$$EP = \frac{Earnings\ per\ share\ (EPS)_{t-3m}}{Adjusted\ share\ price(P)_{t-6m}}$$

Sales to price

In order to ensure consistency with the other market multiples, the share price at the end of the year before the return calculation point is divided by the sales per share for that year.

$$SP = \frac{Adjusted\ share\ price\ (P)_{t-6m}}{Sales\ per\ share\ (WC05508)_{t-6m}}$$

Market Value

Market value (MV) is defined as the market capitalization of the stock in the end of June each year. A potential problem faced when using Datastream to obtain market values is that the values recorded in the database refer to the particular security rather than to the company as a whole. This makes the market values of companies that have more than one class of equity listed appear smaller than they really are. However, as most U.S. companies have either only one class of common stock, or one dominant class, it is not likely that the results would be much different if the firm market value was to be used instead.

$$Market\ Value\ (MV) = MV_t$$

Asset Growth

Asset growth is, as in Cooper et al. (2008), defined as the percentage growth in total assets between year t-1 and t-2. This differs from the Fama and French (2008) methodology who calculate asset growth on a per share basis in order to make the asset growth variable independent from the net stock issuance variable.

$$Total\ Asset\ Growth\ (TASG) = \frac{Total\ Assets\ (DWTA)_{t-6m}}{Total\ Assets\ (DWTA)_{t-18m}} - 1$$

Capital Expenditures Growth

The growth in capital expenditures is defined as the percentage growth in annual capital expenditures from year t-1 to year t-2. This is consistent with the variable used by Anderson and Garcia-Feijóo (2006) but differs from Titman et al. (2004) who instead calculate an abnormal capital expenditures measure by dividing the capital expenditures for year t-1 by the average for the last three years.

$$Capital\ Expenditures\ Growth\ (CEG) = \frac{Capital\ Expenditures\ (WC04601)_{t-6m}}{Capital\ Expenditures\ (WC04601)_{t-18m}} - 1$$

Return on Assets

The Return on Assets figure reported in Worldscope is used for this variable. The measure refers to the ROA for the year prior to the return calculation point.

$$Return\ on\ Assets\ (ROA) = ROA(WC08326)_{t-6m}$$

Return on Equity

The Return on Equity variable is based on the ROE number reported in Datastream for year t-1.

$$Return\ on\ Equity\ (ROE) = ROE(DWRE)_{t-6m}$$

Liquidity

The liquidity measure used is the number of shares traded during last six months in proportion to total number of shares. In line with Datar et al. (1998), stocks that have had splits during the last three months (changes in the split adjustment factor) are excluded in order to avoid stock split related distortions of the measure.

$$Liquidity\ (LIQ) = \frac{Split\ adjustment\ factor\ (AF) * \sum_{i=-3}^{-1} Traded\ volume\ (VO)_i}{\sum_{i=-3}^{-1} Number\ of\ shares\ outstanding\ (NOSH)_i}$$

Beta

Beta is calculated over the last 60 months prior to the forward return period. Stocks need to have recorded returns in every month during this period to be included. The index used is S&P500 (S&PCOMP(RI)).

$$\beta_i = \frac{\sigma_{i,S\&P}}{\sigma_{S\&P}}$$

Net stock issuance

The method applied by Pontiff & Woodgate (2008) is used when calculating the net stock issuance. The logged number of split-adjusted shares outstanding at time t minus eleven months is deducted from the same variable at time t. The NSI observation is set to missing if the number of shares or adjustment factor is missing for any of the two observations required to calculate the factor. This is done in order to avoid obtaining distorted results following IPOs or de-listings.

$$\begin{aligned} &Net\ Stock\ Issuance\ (NSI) \\ &\quad = \ln\left[\frac{Number\ of\ shares\ (NOSH)_t}{Split\ adjustment\ factor\ (AF)_t}\right] \\ &\quad - \ln\left[\frac{Number\ of\ shares\ (NOSH)_{t-11m}}{Split\ adjustment\ factor\ (AF)_{t-11m}}\right] \end{aligned}$$

Net Operating Assets

Net Operating Assets (NOA) is defined as in (Hirshleifer, Hou, Teoh, & Zhang, 2004) and divided by one year lagged assets. In line with the methodology used by these researchers, missing values for the variables total debt and minority interest are set to zero while missing values for the other datatypes lead to the NOA number being excluded from the sample.

$$NOA = \frac{Total\ Assets_{t-6m} - Cash_{t-6m}}{Total\ Assets_{t-18m}} - \frac{Total\ Assets_{t-6m} - Total\ Debt_{t-6m} - Minority\ Interest_{t-6m} - Shareholders\ Equity_{t-6m}}{Total\ Assets_{t-18m}}$$

Accruals

In line with the methodology used by Hirshleifer et al. (2004), non-missing values for current assets, cash, current liabilities and depreciation are required for the observation to be included in the sample, while missing values for short term debt and taxes payable results in that the sub-part of the accruals number being set to zero. The reason for those variables not being required is that the resulting sample reduction of that requirement is not justified by the minimal impact on overall accruals those variables normally have.

$$Accruals\ (ACC) = \frac{\Delta Current\ Assets\ (WC02201)_{t-6m} - \Delta Cash\ (WC02005)_{t-6m}}{0.5 * (Total\ Assets(DWTA)_{t-6m} + Total\ Assets(DWTA)_{t-18m})} - \frac{\Delta Current\ Liab.(WC03101)_{t-6m} - \Delta Short\ t\ debt\ (WC03051 - 18232)_{t-6m} - \Delta Taxes\ payab.(WC03063)_{t-6m}}{0.5 * (Total\ Assets(DWTA)_{t-6m} + Total\ Assets(DWTA)_{t-18m})} - \frac{Depreciation\ (WC01148)_{t-6m}}{0.5 * (Total\ Assets(DWTA)_{t-6m} + Total\ Assets(DWTA)_{t-18m})}$$

Volatility

Volatility (VOL) is defined as the standard deviation of the previous twelve months individual monthly stock returns.

Analyst expected earnings growth

The analyst expected EPS growth (EEPSG) is calculated using IBES median estimates for two years future EPS numbers (F2MD) compared to the last reported annual figure recorded in the same database (F0EPS). One year forecasts, as well as mean values, are also obtained, but since the median two year forecast seems to have the greatest predictive power on stock returns, only this number is used in the calculations

$$EEPSG = \frac{Median\ \ EPS\ forecast_{t+24} - Actual\ EPS_t}{Actual\ EPS_t} - 1$$

Diether et al. (2002), in their article on dispersion of analyst estimates, point out that the way IBES treats stock splits will erroneously smooth out small to medium changes in the reported EPS number. Historical numbers are adjusted using a split-adjustment factor and rounded to the nearest cent, so that the EPS figures are not distorted by stock splits, however this will make small historical EPS changes disappear in the database if the number of shares outstanding has increased substantially because of stock splits. For example, a stock with reported EPS of $1.00 for year t-15 and expected EPS of $1.09 for year t-14 with subsequent cumulative 20-fold stock splits. This would result in a reported EPS after adjustment of $0.05 for both year t and year t+1, i.e. 0 % even though the actual growth was 9 %. The difference in expected EPS growth is however not considerably different in the beginning of the survey period, as would have been expected in the presence of the split problem.

Scaling the expected earnings change with the earnings estimate has two drawbacks; firstly, the formula above cannot be used for stocks with negative EPS at time t; secondly, stocks with very low profitability may have very high expected earnings growth percentagewise, without for that sake being considered growth stocks. To adjust for this, stocks with zero or negative EPS at time t, are excluded, as are stocks with zero or negative return on asset ratios. Stocks with EPS forecasts of zero are also excluded because these database inputs are likely the result of mistakenly including missing forecasts rather than actual zero consensus estimates. However, a bias towards very high expected growth rates among low profitability firms still persist, and it may be important to consider the impact of this when interpreting the results.

Analyst earnings overoptimism

The misestimate variable is defined as the error made by analysts in their EPS forecasts (median two year forecasts) compared to the actual EPS number, scaled by the EPS forecast. The value is positive for overoptimistic forecasts and negative for underoptimistic forecast. As forecasts tend to be overoptimistic on average, this variable is throughout the thesis often referred to as analyst overoptimism.

$$\mathrm{EERRD}_{\mathrm{t,t+2\ yr}} = \frac{2\ year\ median\ EPS\ estimate_t - Actual\ annual\ EPS_{t+24}}{2\ year\ EPS\ estimate_t}$$

This methodology differs from that used by Easterwood and Nutt (1999) in that they use the share price rather than the EPS estimate in the denominator. In this thesis, given the purpose of relating misestimates to variables like book-to-market and asset growth, the methodology of Easterwood and Nutt (1999) would not be suitable because the share price relative to EPS tend to be higher for growth stocks. Thus, if the error was to be divided by the share price, the part of the value/growth premium related to analyst misestimations would be underestimated.

Just as for expected earnings growth, the EERRD number may be biased for low profitability firms, as even small deviations from expected EPS numbers may be very large percentagewise. It is therefore important to consider the possibility of obtaining very large deviations from the actual EPS number when the initial EPS number is low. As observations containing negative EPS forecasts would distort the results severely when using this formula, those observations are excluded from the sample.

As an alternative measure, an analyst misestimate number without direction is used, i.e. the variable is measuring the error in positive terms regardless if it is positive or negative. This factor is abbreviated EERR in the portfolio tables.

Analyst earnings dispersion

The analyst dispersion is defined as the spread of two year EPS forecasts in terms of standard deviation (F2CV), as reported in the IBES database. Also for this variable, it is important to consider the possible distortion due to stock splits found by Diether et al. (2002) (see above). Even though the database description of the F2CV variable does not clearly tell whether the variable is calculated based on split-adjusted numbers or not, the fact that the average dispersion is quite stable over the whole time-period indicates that the dispersion number is likely based on the original EPS data rather than on adjusted figures. However, in order to further reduce the risk of having zero dispersion for a large number of stocks, two year forecasts rather than one year forecasts are used. Only positive dispersion numbers are used in the calculations, the number of observations with a zero dispersion is however negligible, and likely a result of a very low number of forecasts rather than caused by the split-problem. The analyst forecast dispersion variable is abbreviated SDEST in the regression tables.

Analyst coverage

The analyst coverage is defined as the number of one year future EPS estimates (F1NE) recorded in the IBES database. This variable is abbreviated NEST in the regression tables.

3.3.3 Variable List

Below follows a list of abbreviations for the variables used in the portfolio and regression analysis.

Table I
Variable abbreviations for variables included in the Portfolio and Regression Analysis

Code	Description	Included years
F6M	Logged future 6 months return	15
F12M	Logged future 12 months return	15
F24M	Logged future 24 months return	14
F36M	Logged future 36 months return	13
H12M	Logged historical 12 months return	15
H36M	Logged historical 36 months return	15
MV	Market value (million dollars)	15
LIQ	Liquidity	15
BETA	Beta value	15
VOL	Volatility	15
ACC	Accruals	15
TASG	Logged asset growth	15
NSI	Net stock issues	15
CEG	Logged capital expenditures growth	15
NOA	Net operating assets	15
ROA	Return on assets	15
BM	Book to market	15
EP	Earnings to price	15
CP	Cash flow to price	15
SP	Sales to price	15
NEST	Number of analyst 1 year earnings estimates (annual median values)	15
SDEST	Standard deviation of analyst 2 year earnings forecasts (annual median values)	15
EEPSG	Logged median 2 year expected earnings growth (annual median values)	15
EERRD	Logged median 2 year expected earnings error (annual median values)	14

EERR	Logged median 2 year expected earnings error, without direction (annual median values)	14
LEERD	Logged median 2 year expected earnings error at time t+6 months (annual median values)	14

3.4 Evaluation tools

This section contains a description of the calculation techniques used when analyzing the data, as well as explanations of the tables displaying the results.

3.4.1 Portfolio Characteristics

In the portfolio analysis, the data is divided into portfolios based on the value of each of the investigated variables. For each year in the time period 1994-2008, 10 portfolios are formed based on the sorted value of the variable displayed in the heading of each table. The portfolios are sorted from 1 to 10, where portfolio 1 contains the 10 % of stocks with the lowest numbers on the sorting variable for each year, and portfolio 10 contains those 10 % of stocks with the highest numbers.

Panel A reports the characteristics, in terms of average annual portfolio values for each of the variables in this study. Thus, an average 12 months return of 10 % in one of the years for portfolio 10 is given a 6.66 % weight $\left[0.066 = \frac{1\,year}{15\,years}\right]$, when calculating the average 12 months return for portfolio 10, resulting in a total return contribution of 0.066*0.1 for the final portfolio. It is important to note that not all companies have observations for all of the variables. Thus, the portfolio characteristics might have been slightly different if they would have had complete data.

In the right section of each table, the differences between the highest and the lowest portfolio is reported (H-L) followed by t-statistics and statistic significance for this number.

The T-statistic is calculated based on the difference between the highest and the lowest portfolio:

$$t(H-L) = \frac{\overline{H-L}}{\sigma(\widehat{H-L})/\sqrt{n}}$$

Where $\overline{H-L}$ is the average of the annual difference between the highest and the lowest portfolio, $\sigma\left(\widehat{H-L}\right)$ the standard deviation of this number, and n the number of portfolio formations.

The statistical significance is calculated based on a two tailed t-distribution, with degrees of freedom of n-(k+1), where n = the number of years in the sample, and k = the number of parameters (1).

For obvious reasons, some variables are not possible to obtain for all of the years, e.g. 36 months future return from June 2008. For those variables, there will be less than 15 portfolio averages. All numbers that refer to changes from one time period to another, i.e. returns, asset growth, capital expenditures growth, expected earnings growth, as well as the earnings error numbers, are logged. Further, the numbers reported for the variables that contain analyst estimates refer to medians instead of averages. This is because of a tendency of extreme values for these variables to distort the results.

Due to the similarity of some variables, and limited space in this paper, not all variables are included in the portfolios. Some of the 1 and 3 months future stock returns are excluded, together with 1, 3, 6 and 24 months historical returns and return on equity, return on assets, total earnings estimate error, and volatility. Some of these are however displayed as characteristics of portfolios sorted on other variables. See Table I in section 3.3.3 for a list of abbreviations used for the portfolio variables.

3.4.2 Cross-Portfolio Analysis

For a few variables, cross-portfolios are conducted in line with the methodology used by Nagel (2005). This is done in order to capture return differences within portfolios caused by different characteristics on another variable. In this sense is it similar to a multivariate regression with two independent variables, but does, like the single portfolio analysis, better capture non-linear relations in the cross-section than does a regression.

For the cross-portfolios, the exact same technique is used as for the standard portfolio analysis with the only difference that each portfolio is divided into 5 sub portfolios sorted on a secondary variable. That is; the 12 months stock returns are first sorted on the main variable and divided into 5 equally large portfolios. Then, each portfolio is divided into 5 portfolios sorted on a secondary variable, resulting in a matrix containing 25 cells. Hedge portfolio returns, accompanied with statistic significance measures, are reported to the right of the table for hedge portfolio returns between portfolio 1 and 5 on the primary variable but within, for example, portfolio one on the secondary variable, and in the bottom of the table for hedge returns between portfolio 1 and 5 on the secondary variable but in, for example, portfolio 2 on the primary variable.

When performed, the results from the cross-portfolio analysis are displayed in panel C in tables III-XXIII. Appendix A holds a summarized table with all the 12 months returns of variable sorted portfolios.

Although not mentioned in (Nagel, 2005), it is important to bear in mind that the portfolio boundaries are not fixed to certain numbers, but are always distributed evenly with the same number of observations for every portfolio in all the cross-portfolios. This leads to important implications when interpreting the hedge portfolio values and statistical significance when the two variables are correlated. Even though a stock is recorded in the high portfolio on the secondary variable and in, say, portfolio 5 on the primary one, it is far from certain that the same stock would have been in the high portfolio on the secondary variable if placed in portfolio 1 on the primary one. Thus, if the limit for being included in the high portfolio is, for example, getting lower for every higher portfolio on the primary scale, the hedge portfolio return between portfolio 1 and 5 on the primary scale would tend to represent diagonal hedge returns from the upper left corner to the lower right, rather than horizontal hedge returns from the upper left corner to the upper right. Thus, even though the cross-portfolio could still be used and present interesting results when the two variables are correlated, it is important to have this in mind when interpreting the results.

3.4.3 Cross-Sectional Regressions

Panel B of each Table IV to XXIII contains cross-sectional regressions, and Appendix B contains a summary of univariate regressions based on each of the variables. Before running the regressions, two adjustments are made to the variables. Firstly, all explanatory variables and returns used in the regressions are logged. Secondly, the explanatory variables are winsorized at the 0.5 and the 99.5 percentiles in order to reduce the distorting impact of sample outliers. This means that the bottom 0.5 % values in the sample are set to equal the 0.5 percentile value and the top 0.5 % values to equal the value of the 99.5 % percentile.

In accordance with standard asset pricing research practice, (Fama & MacBeth, 1973) regressions are constructed based on a combination of annual cross-sectional regressions according to the methodology outlined below;

The observations for each variable (s_j) are regressed against the corresponding stock return of the following period ($\mathrm{R}_{\mathrm{t+1}}$) , resulting in an equation of the form:

$$\mathrm{R}_{\mathrm{t+1}} = \gamma_0 + \gamma_j s_j + \varepsilon$$

To estimate the regression described above, the Fama and MacBeth (1973) two step cross-sectional regression approach is used. In the first step, one regression is run for each year in the sample (15 regressions in total for most variables). In the second step, $\gamma_j$is estimated as the sample mean of the annual cross-section regression coefficient, $\widehat{\gamma_{Jt}}$. A t-test is then conducted based on the standard error ($\sigma(\widehat{\gamma_{Jt}})/\sqrt{n}$)of the γ_j estimate.

$$t(\gamma_j) = \frac{\gamma_j}{\sigma(\widehat{\gamma_{Jt}})/\sqrt{n}}$$

Where $\sigma(\widehat{\gamma_{Jt}})$ is the standard deviation of the $\widehat{\gamma_{Jt}}$ time-series and n the number of years in the sample.

Then, the statistical significance is calculated based on a two tailed t-distribution, with degrees of freedom of n-(k+1), where n = the number of years in the sample, and k = the number of parameters including the intercept.

After regressing the variables separately, multiple independent variables are run in the same regression using a method equivalent to the one above, resulting in the following formula:

$$\mathrm{R}_{t+1} = \gamma_0 + \gamma_1 s_1 + \gamma_2 s_2 + \ldots + \gamma_j s_j + \varepsilon$$

Panel B in tables III to XXIII displays univariate and multivariate cross-sectional regressions with 12 months future returns as the dependent variable and, as dependent variables, some of the most interesting variables from the portfolio analysis. Every second line in the regression tables makes up one regression, with the dependent variables in the table heading and the regression intercept and coefficients displayed in the row. The line below displays t-statistics in brackets for the intercept and each coefficient. Statistic significance is displayed in numbers following the t-statistics figure where 3 corresponds to a statistic significance below 0.01, 2 to a significance below 0.05, and to 1 a significance below 0.10.

Each regression line contains empty spaces for the variables that are not included in that regression.

Since one regression is run for every time period in the sample, the Fama-MacBeth type of regressions do a great job adjusting for differing sample sizes across time periods. The problem with this regression technique is that they tend to focus on the number of time periods rather than the size of the samples in each time period when estimating the statistical significance. When the number of time periods is small, the standard errors tend to become large and the degrees of freedom small, resulting in very high requirements on the relations between the independent and dependent variable to obtain significant results. The sample in this thesis is large but includes only 15 measurement points, resulting in the significance between the variables from the annual regressions being harshly punished when combined using the Fama-MacBeth technique. This makes the Fama-MacBeth regressions a less than perfect technique in this particular case, however, if the variables obtains significant results in the regressions, that significance is probably very robust.

4. Empirical Results

4.1 Research Sample

Table II
Sample Statistics - Firms

Table II displays annual data for the stocks included in the sample from which the calculations are made. Panel A shows the number of included stocks for each year in during the survey period. Results are reported by exchange and in total for the whole sample.
Panel B displays the total market capitalization for all stocks in the sample, reported per year and exchange.

Table II: Panel A - Number of firms by exchange and year, 1994-2008

Exchange	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
NYSE	1,052	1,036	1,007	1,028	1,029	986	941	900	895	885	890	908	911	901	908
NASDAQ	829	961	1,188	1,325	1,379	1,393	1,493	1,460	1,377	1,330	1,342	1,371	1,384	1,400	1,359
Total	1,881	1,997	2,195	2,353	2,408	2,379	2,434	2,360	2,272	2,215	2,232	2,279	2,295	2,301	2,267

Table II: Panel B - Total market capitalization by exchange and year ($B), 1994-2008

Exchange	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
NYSE	2,429	2,927	3,618	4,519	5,445	5,711	5,777	6,536	6,637	6,690	6,985	7,116	7,525	8,045	8,898
NASDAQ	159	272	415	522	693	767	972	1,013	1,140	1,345	1,446	1,528	1,538	1,839	2,053
Total	2,588	3,199	4,033	5,041	6,238	6,479	6,749	7,550	7,777	8,035	8,432	8,644	9,064	9,885	10,951

4.2 Analyst Misestimates

In order to test the hypothesis that the degree of realized overoptimism in earnings forecasts is related to returns, portfolios are constructed which are sorted on this variable. Table III below displays the characteristics of the cross-section of portfolios sorted on the difference between the earnings estimate for two year ahead earnings (for the fiscal year ending 18 months ahead as the base date is the end of June) and the actual earnings number reported for that year. This difference is then divided by the earnings estimate, as demonstrated in the methodology chapter (section 3.3.1) resulting in positive numbers for overestimated earnings and negative numbers for underestimated earnings. The reason for using two years earnings estimates rather than one year estimates is that stock market pricing by definition is forward looking and that part of next period earnings will already be priced in at the end of any given year. Using a one year estimate instead would mean that earnings expectations would refer to the fiscal year ending only 6 months into the future, thus likely having limited explanatory ability for 1-3 years stock returns. This conception is confirmed in regressions showing that the 2 year overoptimism measure has stronger explaining abilities than the one year measure on 6 months or longer return horizons.

Table III
Portfolios Sorted on Earnings Estimates Overoptimism

Panel A shows characteristics of portfolios sorted on Earnings Estimates Overoptimism. The portfolios are created annually according to the methodology in section 3.4.1 and the variable abbreviations are listed in Table I, section 3.3.3. The numbers in the table below are the portfolio averages for the 1994-2008 period. The upper variable is the sorting variable, on which the portfolios are sorted from low to high. The right side of the table reports the difference between the highest and lowest portfolio for each variable and the statistical significance of the resulting number.

Table III: Panel A - Characteristics of portfolios sorted on Earnings Estimates Overoptimism, 1994-2008

	Low		Variable sorted portfolios					High			Significance		
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**H-L**	**T-stat**	**P**
EERRD	-1.156	-0.251	-0.100	-0.016	0.054	0.125	0.210	0.314	0.465	0.756	1.911	9.02	0.000
F6M	0.137	0.076	0.031	0.003	-0.046	-0.083	-0.110	-0.175	-0.210	-0.285	-0.422	-8.67	0.000
F12M	0.341	0.193	0.129	0.075	0.012	-0.048	-0.099	-0.177	-0.267	-0.447	-0.788	-10.55	0.000
F24M	0.509	0.378	0.290	0.219	0.129	0.047	-0.022	-0.129	-0.269	-0.567	-1.076	-21.55	0.000
F36M	0.566	0.443	0.363	0.284	0.197	0.098	0.022	-0.070	-0.200	-0.521	-1.087	-18.33	0.000
H12M	0.222	0.190	0.139	0.110	0.085	0.059	0.024	0.023	-0.020	-0.121	-0.344	-7.05	0.000
H36M	0.500	0.503	0.395	0.410	0.352	0.386	0.346	0.381	0.352	0.278	-0.222	-2.99	0.011
MV	4,592	5,623	7,295	10526	8,359	7,336	5,842	3,621	3,336	1,379	-3,213	-4.22	0.001
LIQ	0.193	0.185	0.156	0.151	0.141	0.160	0.163	0.177	0.194	0.231	0.037	2.53	0.026
BETA	1.037	0.946	0.878	0.830	0.839	0.873	0.934	0.958	1.061	1.160	0.123	2.03	0.065
VOL	0.124	0.111	0.098	0.093	0.092	0.097	0.107	0.119	0.127	0.151	0.027	4.02	0.002
ACC	-0.027	-0.021	-0.031	-0.028	-0.025	-0.028	-0.025	-0.016	-0.013	-0.003	0.024	4.54	0.001
TASG	0.169	0.164	0.157	0.150	0.137	0.164	0.163	0.189	0.203	0.265	0.096	4.76	0.000
NSI	0.036	0.032	0.024	0.028	0.027	0.026	0.030	0.037	0.045	0.060	0.025	3.13	0.009
CEG	0.219	0.240	0.255	0.213	0.220	0.264	0.246	0.282	0.303	0.392	0.174	3.06	0.010
NOA	0.728	0.723	0.727	0.746	0.724	0.769	0.749	0.774	0.775	0.754	0.026	0.76	0.463
ROA	0.098	0.090	0.089	0.092	0.091	0.090	0.109	0.090	0.089	0.081	-0.017	-0.98	0.344
BM	0.519	0.448	0.433	0.425	0.427	0.430	0.451	0.434	0.471	0.487	-0.032	-2.24	0.045
EP	0.056	0.055	0.053	0.054	0.055	0.054	0.055	0.053	0.052	0.048	-0.008	-2.38	0.034
CP	0.132	0.114	0.112	0.112	0.112	0.112	0.115	0.110	0.113	0.109	-0.023	-3.08	0.010
SP	1.433	1.237	1.115	1.130	1.081	1.169	1.228	1.180	1.242	1.303	-0.130	-1.06	0.312
NEST	9.553	9.831	10.65	11.16	11.15	10.44	9.750	9.045	8.925	7.592	-1.962	-4.93	0.000
EEPSG	0.303	0.263	0.220	0.191	0.189	0.195	0.201	0.238	0.276	0.287	-0.016	-0.33	0.747
SDEST	0.239	0.117	0.102	0.093	0.097	0.103	0.111	0.123	0.147	0.232	-0.006	-0.14	0.890
EERR	0.465	0.194	0.093	0.055	0.070	0.126	0.210	0.314	0.465	0.756	0.290	5.04	0.000
LEERD	-0.516	-0.161	-0.071	-0.015	0.036	0.095	0.165	0.253	0.415	0.765	1.281	27.20	0.000

As expected, the degree of overoptimism is a very strong driver of stock returns. The returns of the portfolios sorted on overoptimism seem to spread across the cross-section in a linear fashion, with a huge return spread of almost 80 % between the highest and the lowest overoptimism portfolio. As the actual overoptimism is only known ex post, it is of course not possible to construct these portfolios in reality. The reason for analyzing stock returns of portfolios sorted on the overoptimism measure is rather to judge whether over or underperforming compared to expectations has a strong impact on stock returns, and that is obviously the case. Still, the likelihood of earnings estimates being overoptimistic, as table III indicates and which will be expanded on further, is partly predictable and related to a large number of the variables researched in this thesis.

On average, analysts tend to be overoptimistic. However this overoptimism does not seem to be spread evenly in the cross-section, as there is a tendency for stocks in the overoptimistic portfolios to have higher historical growth rates in assets and capital expenditures, and lower values on the market multiples; i.e. characteristics that previous research have shown predict low returns.

4.3 Historical Returns

Below follows a detailed review of the characteristics of portfolios sorted on 12 months historical stock returns and 36 months historical stock returns, as well as Fama-MacBeth regressions for these variables.

Table IV shows a negative relation between 12 months historical returns and 6 months future returns. After 6 months, the returns are reversing, resulting in lower long-term returns for the high portfolios. From Table V, a negative relation between 36 months historical returns and 12-36 months future returns can be observed. These results are consistent with the Jegadeesh and Titman (1993) short term momentum effect and the longer-term return reversal effect documented by De Bondt and Thaler (1987).

Table IV
Portfolios and regressions for 12 months historical returns

Panel A shows characteristics of portfolios sorted on 12 months historical return. The portfolios are created annually according to the methodology in section 3.4.1 and the variable abbreviations are listed in Table I, section 3.3.3. The numbers in the table below are the portfolio averages for the 1994-2008 period. The upper variable is the sorting variable, on which the portfolios are sorted from low to high. The right side of the table reports the difference between the highest and lowest portfolio for each variable and the statistical significance of the resulting number. Panel B displays Fama-Macbeth regressions created in line with the methodology in section 3.4.3. Every second line represents a regression with the regression variables in the column headings and the statistical significance reported in brackets below the coefficients. Panel C presents cross-portfolios created using the methodology described in section 3.4.2. The portfolios are first sorted on the horizontal variable and then every portfolio is sorted by an intersecting variable reported vertically. The right and lower parts of the table displays the statistical significance of each hedge portfolio

Table IV: Panel A - Characteristics of portfolios sorted on 12 months historical return. 1994-2008

	Low			Variable sorted portfolios				High			Significance		
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**H-L**	**T-stat**	**P**
H12M	-0.961	-0.434	-0.236	-0.099	0.005	0.097	0.193	0.307	0.478	0.941	1.902	16.24	0.000
F6M	-0.143	-0.089	-0.076	-0.044	-0.031	-0.020	-0.012	-0.013	-0.037	-0.057	0.086	2.62	0.022
F12M	-0.028	0.005	0.007	0.024	0.034	0.046	0.047	0.039	-0.014	-0.036	-0.008	-0.10	0.918
F24M	0.058	0.070	0.059	0.096	0.121	0.109	0.098	0.074	-0.003	-0.085	-0.144	-1.13	0.278
F36M	0.160	0.134	0.118	0.176	0.205	0.198	0.184	0.136	0.040	-0.070	-0.230	-1.81	0.094
H36M	-0.696	-0.205	-0.059	0.114	0.169	0.239	0.327	0.410	0.495	0.807	1.503	15.77	0.000
MV	491	1,542	2,482	3,898	4,708	4,940	4,789	4,898	4,101	2,043	1.552	3.30	0.006
LIQ	0.254	0.147	0.130	0.126	0.269	0.123	0.129	0.146	0.170	0.364	0.110	0.95	0.360
BETA	1.308	1.097	1.014	0.923	0.881	0.848	0.868	0.884	1.006	1.095	-0.213	-1.61	0.131
VOL	0.203	0.158	0.137	0.120	0.113	0.112	0.110	0.122	0.147	0.235	0.032	1.54	0.148
ACC	-0.019	-0.030	-0.037	-0.030	-0.030	-0.035	-0.033	-0.034	-0.032	-0.019	0.001	0.05	0.961
TASG	0.155	0.133	0.128	0.116	0.119	0.119	0.121	0.139	0.158	0.200	0.045	1.44	0.174
NSI	0.093	0.058	0.047	0.030	0.036	0.035	0.037	0.051	0.072	0.111	0.018	1.59	0.136
CEG	0.179	0.147	0.126	0.107	0.101	0.097	0.105	0.124	0.126	0.155	-0.024	-0.36	0.726
NOA	0.668	0.686	0.683	0.689	0.673	0.688	0.696	0.698	0.726	0.702	0.034	0.29	0.778
ROA	-0.130	-0.036	0.013	0.026	0.028	0.053	0.057	-0.014	0.037	-0.014	0.115	2.78	0.015
BM	0.688	0.630	0.606	0.610	0.552	0.544	0.511	0.489	0.478	0.481	-0.207	-3.60	0.003
EP	0.121	0.091	0.067	0.075	0.063	0.077	0.059	0.061	0.059	0.065	-0.055	-1.67	0.118
CP	0.147	0.137	0.135	0.140	0.129	0.128	0.130	0.121	0.117	0.139	-0.008	-0.50	0.628
SP	2.227	1.801	1.752	1.645	1.911	3.328	1.448	1.468	1.417	1.692	-0.535	-1.38	0.191
NEST	5.542	6.879	7.666	8.155	8.995	9.151	8.831	8.381	7.926	6.093	0.551	0.97	0.351
SDEST	0.700	0.398	0.327	0.227	0.143	0.144	0.135	0.152	0.150	0.240	-0.460	-4.02	0.001
EEPSG	0.099	0.223	0.239	0.239	0.256	0.270	0.295	0.344	0.410	0.590	0.491	8.67	0.000
EERRD	0.422	0.270	0.195	0.124	0.094	0.050	0.038	0.027	0.019	-0.036	-0.458	-8.05	0.000
EERR	0.481	0.348	0.264	0.203	0.179	0.160	0.168	0.166	0.196	0.288	-0.193	-5.26	0.000
LEERD	0.239	0.150	0.103	0.058	0.050	0.023	0.025	0.027	0.035	0.038	-0.201	-4.83	0.000

Table IV: Panel B – Fama-MacBeth Annual Return Regressions based on 12 months historical return

Intercept	H12M	BM	MV	LEERD	EERRD	SDEST
0.051	0.016	0.049				
(0.86)	(0.54)	(2.24)[2]				
0.066	0.017	0.044	-0.003			
(0.68)	(0.55)	(1.87)[1]	(-0.34)			
0.094	-0.056			-0.411		
(2.87)	(-1.57)			(-13.81)[3]		
0.099	-0.122				-0.442	
(3.26)	(-3.54)[3]				(-15.47)[3]	

Table V: Panel C – 6 months returns for portfolios sorted on 12 months historical returns and Analyst coverage

12 months historical return	Analyst coverage						Significance		
	0	**L**	**2**	**3**	**4**	**H**	**H-L**	**T-stat**	**P**
Lowest	-0.123	-0.129	-0.160	-0.118	-0.088	-0.072	0.057	1.42	0.179
1	-0.061	-0.069	-0.069	-0.065	-0.032	-0.017	0.052	1.13	0.279
2	-0.005	-0.033	-0.045	-0.036	-0.022	-0.001	0.032	0.96	0.355
3	-0.001	-0.023	-0.023	-0.014	-0.019	-0.010	0.013	0.50	0.625
Highest	-0.020	-0.045	-0.064	-0.051	-0.056	-0.018	0.027	0.74	0.472
	Significance								
H-L	0.103	0.084	0.096	0.067	0.032	0.054			
t-stat	3.19	3.28	3.32	2.12	0.73	1.06			
P	0.007	0.006	0.006	0.054	0.478	0.308			

Table V
Portfolios and regressions for 36 months historical returns

Panel A shows characteristics of portfolios sorted on 36 months historical return. The portfolios are created annually according to the methodology in section 3.4.1 and the variable abbreviations are listed in Table I, section 3.3.3. The numbers in the table below are the portfolio averages for the 1994-2008 period. The upper variable is the sorting variable, on which the portfolios are sorted from low to high. The right side of the table reports the difference between the highest and lowest portfolio for each variable and the statistical significance of the resulting number. Panel B displays Fama-Macbeth regressions created in line with the methodology in section 3.4.3. Every second line represents a regression with the regression variables in the column headings and the statistical significance reported in brackets below the coefficients. Panel C presents cross-portfolios created using the methodology described in section 3.4.2. The portfolios are first sorted on the horizontal variable and then every portfolio is sorted by an intersecting variable reported vertically. The right and lower parts of the table displays the statistical significance of each hedge portfolio.

Table V: Panel A - Characteristics of portfolios sorted on 36 months historical return. 1994-2008

	Low			Variable sorted portfolios				High			Significance		
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**H-L**	**T-stat**	**P**
H36M	-1.502	-0.657	-0.307	-0.065	0.126	0.294	0,463	0,668	0,958	1,642	3,144	26,55	0.000
F6M	-0.067	-0.082	-0.031	-0.039	-0.028	-0.030	-0.037	-0.053	-0.048	-0.088	-0.021	-0.62	0.544
F12M	0.092	0.014	0.051	0.027	0.045	0.018	0.028	-0.003	-0.009	-0.085	-0.177	-3.08	0.009
F24M	0.177	0.081	0.074	0.078	0.099	0.089	0.096	0.047	0.014	-0.135	-0.312	-2.63	0.021
F36M	0.259	0.144	0.144	0.169	0.176	0.158	0.171	0.109	0.058	-0.116	-0.375	-3.23	0.007
H12M	-0.454	-0.182	-0.065	-0.008	0.033	0.086	0,113	0,183	0,246	0,469	0,923	11,47	0.000
MV	274	1,008	2,583	2,323	3,181	3,569	4,384	4,243	3,557	3,339	3,065	4.31	0.001
LIQ	0.140	0.141	0.133	0.137	0.138	0.142	0.150	0.163	0.200	0.269	0.129	7.71	0.000
BETA	1.362	1.179	1.096	0.990	0.917	0.879	0.898	0.937	0.968	1.130	-0.232	-1.82	0.091
VOL	0.224	0.167	0.148	0.134	0.124	0.123	0.122	0.132	0.147	0.183	-0.041	-2.16	0.050
ACC	-0.056	-0.044	-0.044	-0.032	-0.028	-0.032	-0.034	-0.031	-0.031	-0.003	0.053	5.49	0.000
TASG	-0.096	0.038	0.064	0.089	0.111	0.109	0.143	0.165	0.223	0.339	0.435	13.22	0.000
NSI	0.110	0.049	0.038	0.028	0.034	0.038	0.043	0.050	0.072	0.083	-0.027	-1.51	0.156
CEG	-0.344	-0.113	-0.020	0.043	0.090	0.130	0.155	0.199	0.284	0.436	0.779	14.78	0.000
NOA	0.458	0.576	0.621	0.609	0.687	0.660	0.693	0.683	0.707	0.763	0.306	6.89	0.000
ROA	-0.200	-0.069	-0.029	-0.011	0.032	0.052	0.063	0.061	-0.033	0.061	0.261	7.64	0.000
BM	0.997	0.753	0.667	0.605	0.584	0.538	0.501	0.452	0.409	0.322	-0.675	-5.36	0.000
EP	0.372	0.074	0.066	0.068	0.063	0.062	0.060	0.058	0.060	0.051	-0.321	-1.86	0.086
CP	0.251	0.148	0.144	0.137	0.135	0.129	0.121	0.115	0.112	0.090	-0.161	-3.20	0.007
SP	3.283	2.259	2.054	1.809	1.586	1.420	1.413	1.367	1.257	1.064	-2.219	-2.55	0.024
NEST	4.394	5.684	6.962	7.692	8.399	8.360	8.516	8.037	7.655	7.696	3.302	6.18	0.000
SDEST	0.742	0.370	0.285	0.265	0.213	0.183	0.142	0.153	0.181	0.183	-0.559	-4.49	0.001
EEPSG	0.280	0.322	0.276	0.263	0.263	0.278	0.292	0.315	0.351	0.483	0.203	3.37	0.005
EERRD	0.190	0.254	0.126	0.124	0.081	0.095	0.082	0.072	0.061	0.069	-0.121	-3.33	0.006
EERR	0.302	0.331	0.242	0.222	0.196	0.187	0.188	0.204	0.214	0.267	-0.035	-1.10	0.294
LEERD	0.047	0.135	0.061	0.051	0.040	0.051	0.050	0.046	0.036	0.074	0.027	0.61	0.551

Table V: Panel B – Fama-MacBeth Annual Return Regressions based on 36 months historical return

Intercept	H36M	BM	MV	LIQ	TASG	LEERD
0.049	-0.031	0.035				
(0.88)	(-1.89)[1]	(1.57)				
0.047	-0.032	0.033	0.000			
(0.55)	(-2.12)[1]	(1.36)	(0.02)			
0.065	-0.032			-0.406		
(1.18)	(-2.00)[1]			(-2.28)[1]		
0.039	-0.027				-0.133	
(0.78)	(-1.60)				(-4.26)[3]	
-0.117	--0.015			-0.320	-0.117	
(1.35)	(-1.03)			(-1.80)	(-4.02)[2]	
0.127	-0.033			-0.106	-0.024	-0.407
(3.50)	(-1.89)[1]			(-0.70)	(-1.19)	(-13.12)[3]
0.125	-0.035			-0.124		-0.408
(3.48)	(-2.02)[1]			(-0.80)		(-13.37)[3]
0.107	-0.046					-0.042
(3.46)	(-2.05)[1]					(13.06)[3]

Table V: Panel C – 6 months returns for portfolios sorted on 36 months historical returns and Analyst coverage

36 months historical return	Analyst coverage						Significance		
	0	L	2	3	4	H	H-L	T-stat	P
Lowest	-0.095	0.064	-0.002	-0.032	0.048	0.030	0.125	-0.035	-0.59
1	-0.034	0.004	-0.033	0.022	0.034	0.004	0.038	-0.031	-0.63
2	-0.039	0.033	0.019	0.047	0.034	0.033	0.072	0.039	1.26
3	-0.015	-0.005	-0.048	-0.030	0.004	-0.005	0.010	-0.014	-0.38
Highest	-0.058	-0.080	-0.057	-0.064	0.021	-0.080	-0.022	0.078	1.22
	Significance								
H-L	0.037	-0.121	-0.078	-0.026	-0.113	-0.008			
t-stat	0.872	-2.48	-1.73	-0.59	-1.53	-0.13			
P	0.399	0.028	0.107	0.567	0.151	0.895			

The standard earnings estimate error measure cannot be used for historical returns because there is a clear risk that the historical returns reported has already priced in future changes in earnings estimates. This is due to the fact that information leading to earnings estimates revisions is available to the market prior to the date the revisions are recorded in the IBES database. Thus, there will appear to be a relationship between historical returns and earnings estimate errors even though there is no such relationship. To counter the distorting effect caused by this lag between changed investor earnings perceptions and changed earnings estimates in the database, an alternative measure based on 6 months future estimates is used (LEERD). Thus, the future estimate error reported in June is based on the earnings estimate 6 months into the future, at a date when all estimate changes that occurred prior to the end of June should have been recorded in the database. Although using this measure dramatically reduces the relation between 12 months past returns and estimate overoptimism, there is still a highly significant correlation between the two variables.

The analyst overoptimism variable contains interesting information that may explain the momentum effect as a result of mispricing. For 36 months historical returns there is no considerable difference in the overoptimism measure across portfolios, thus failing to explain the low future returns for portfolios with high 36 months historical returns. For 12 months historical returns however, the story is completely different. While there still is no big difference between the earnings estimate overoptimism of medium and high historical return portfolios, the low portfolios display significant overoptimism. Although it is difficult to say exactly when the overoptimism will be translated into lower returns, the pattern in the overoptimism measure clearly corresponds to 6 months forward returns, which are lower for the low portfolios but not very different among the rest. Further evidence indicating that the momentum effect is related to overly optimistic earnings forecasts is that, regressed together with the overoptimism measure, the coefficient for historical 12 months returns turns negative.

According to the empirical results above, the puzzling momentum effect followed by subsequent reversals observed in previous studies, as well as in this one, is caused by overoptimism about the future of stocks that have performed poorly in the 12 months period prior to portfolio formation. Removing the overoptimism effect; the momentum effect disappears, resulting in a reversal effect. This is in line with Barberis (1998) hypothesis about “analyst conservatism”. If analysts and investors are slow at incorporating new information, there will be a lag in earnings forecast revisions for stocks where the conditions have changed a lot recently (which has probably been the case for those stocks with very high or low 12 months historical returns). Because of this lag; the forecasts for stocks that performed poorly during the previous 12 months will temporarily be too high; which will subsequently be eliminated resulting in lower returns (the momentum effect).

The evidence provided by Easterwood and Nutt (1999); that analysts tend to underreact to negative information and overreact to positive, might explain why the momentum effect is not linear in the cross-section but driven by the historically bad portfolios. As these portfolios exhibit considerably larger earnings estimation errors and 6 months future returns than all the other portfolios, the opposite is not true for the historically good portfolios that have lower returns than the middle portfolios and about the same overoptimism numbers.

Hong et al. (2000) argue that the momentum effect is driven by stocks with slow information flow, using analyst coverage as a proxy for this. If the previous argument is true that the momentum effect is caused by lags in investor's reactions to negative information leading to overoptimism about the future of previous losers; the stronger momentum effect for stocks with slow information flow observed in Hong et al. (2000) should be present in this data as well. As a test for this, a cross-portfolio between 12 months historical returns and analyst coverage is formed (Table IV, Panel C). The results from this clearly show that the momentum effect, just as expected, is significant only for stocks with low analyst coverage. There are similar indications for portfolios sorted on 36 months return, but here, both stocks with high analyst coverage and no coverage at all have no reversal effect, while it seems to be more common in the middle portfolios.

The 36 months past return measure, showing very strong long term reversals, is clearly correlated with a number of other researched variables. For example, the high 36 months historical returns portfolios have asset growth, liquidity, and book-to-market characteristics that all predict low future returns, and the opposite goes for the low historical portfolios. This makes it difficult to distinguish the reversal effect from the effect of those variables. However; while portfolios with high asset growth and liquidity tend to be related to overoptimistic earnings expectations, this is not manifested in terms of differing degrees of overoptimism (LEERD) in the 36 months historical return sorted portfolios.

4.4 Beta, Market Value and Liquidity

Contrary to what the CAPM states, beta is negatively related to future returns, with an increasing effect over time up to at least 36 months. Looking at the characteristics of the beta sorted portfolios, it becomes obvious that this might be due to correlation between beta and other variables. High beta stocks tend to be volatile, liquid and have low book-to-market ratios and low profitability. Since the high (low) beta portfolios tend to have variable characteristics that predict low (high) returns, the return spread is likely caused by exposure to these factors rather than to beta per se. Further data suggesting that this is the case is that beta has a positive relation with overoptimistic EPS estimates and that the negative coefficient for beta in the Fama-MacBeth regressions drops to almost zero when the earnings overoptimism measure is included in the regression. This suggests that it is possible that the "pure beta effect" is actually influencing returns positively, but the influence of the aforementioned factors distorts the picture, making beta insignificant or even negative.

Table VI
Portfolios and regressions for Beta

Panel A shows characteristics of portfolios sorted on Beta. The portfolios are created annually according to the methodology in section 3.4.1 and the variable abbreviations are listed in Table I, section 3.3.3. The numbers in the table below are the portfolio averages for the 1994-2008 period. The upper variable is the sorting variable, on which the portfolios are sorted from low to high. The right side of the table reports the difference between the highest and lowest portfolio for each variable and the statistical significance of the resulting number. Panel B displays Fama-Macbeth regressions created in line with the methodology in section 3.4.3. Every second line represents a regression with the regression variables in the column headings and the statistical significance reported in brackets below the coefficients.

Table VI: Panel A - Characteristics of portfolios sorted on Beta. 1994-2008

	Low			Variable sorted portfolios				High			Significance		
	1	2	3	4	5	6	7	8	9	10	H-L	T-stat	P
BETA	-0.084	0.308	0.491	0.642	0.790	0.940	1.115	1.341	1.672	2.486	2.569	16.95	0.000
F6M	-0.046	-0.011	-0.012	-0.027	-0.023	-0.020	-0.030	-0.048	-0.055	-0.074	-0.027	-0.87	0.399
F12M	0.030	0.065	0.060	0.040	0.051	0.051	0.039	0.036	0.013	-0.044	-0.074	-1.44	0.174
F24M	0.124	0.137	0.129	0.126	0.122	0.110	0.087	0.085	0.062	-0.021	-0.145	-2.42	0.031
F36M	0.222	0.232	0.223	0.229	0.195	0.199	0.166	0.168	0.126	0.020	-0.202	-2.68	0.019
H12M	0.119	0.077	0.076	0.081	0.080	0.076	0.077	0.023	0.002	-0.030	-0.148	-1.66	0.121
H36M	0.297	0.273	0.250	0.266	0.260	0.251	0.258	0.161	0.124	0.059	-0.238	-1.63	0.127
MV	2,498	3,268	5,319	5,587	5,154	5,292	5,175	5,570	4,410	1,943	-555	-0.79	0.445
LIQ	0.113	0.108	0.113	0.119	0.123	0.137	0.150	0.160	0.184	0.234	0.121	12.75	0.000
VOL	0.125	0.099	0.102	0.105	0.110	0.118	0.124	0.138	0.159	0.206	0.080	4.66	0.000
ACC	-0.038	-0.035	-0.039	-0.035	-0.033	-0.034	-0.026	-0.030	-0.029	-0.037	0.001	0.15	0.883
TASG	0.106	0.086	0.098	0.096	0.099	0.102	0.111	0.106	0.123	0.125	0.019	0.63	0.540
NSI	0.050	0.029	0.026	0.033	0.029	0.026	0.028	0.045	0.045	0.080	0.030	3.43	0.004
CEG	0.116	0.121	0.093	0.051	0.073	0.079	0.115	0.070	0.104	0.083	-0.032	-0.53	0.607
NOA	0.679	0.683	0.686	0.705	0.692	0.656	0.682	0.638	0.631	0.530	-0.148	-2.76	0.016
ROA	0.126	0.074	0.065	0.051	0.049	0.024	0.041	0.013	0.00	-0.170	-0.297	-2.96	0.011
BM	0.637	0.628	0.612	0.606	0.561	0.544	0.530	0.554	0.517	0.499	-0.138	-3.91	0.002
EP	0.071	0.084	0.066	0.062	0.061	0.063	0.059	0.059	0.057	0.096	0.024	1.07	0.306
CP	0.156	0.145	0.149	0.148	0.135	0.132	0.125	0.127	0.119	0.108	-0.049	-3.65	0.003
SP	1.840	1.629	1.977	1.724	1.609	1.737	1.769	1.754	1.485	1.513	-0.327	-2.02	0.064
NEST	6.569	7.391	8.364	8.989	9.482	9.817	9.696	9.541	9.166	8.138	1.570	3.41	0.005
SDEST	0.233	0.897	0.152	0.155	0.135	0.153	0.185	0.227	0.295	0.455	0.222	1.50	0.158
EEPSG	0.240	0.221	0.228	0.252	0.275	0.273	0.304	0.308	0.351	0.387	0.147	3.55	0.004
EERRD	0.054	0.048	0.056	0.059	0.059	0.072	0.085	0.114	0.128	0.194	0.141	2.44	0.031
EERR	0.176	0.142	0.155	0.179	0.171	0.196	0.220	0.221	0.257	0.346	0.170	3.51	0.004
LEERD	0.019	0.022	0.027	0.039	0.036	0.048	0.047	0.061	0.065	0.102	0.084	2.37	0.036

Table VI: Panel B – Fama-MacBeth Annual Return Regressions based on Beta

Intercept	Beta	BM	MV	LIQ	EERRD
0.083	-0.022	0.36			
(1.43)	(-0.92)	(2.03)[1]			
0.085	-0.024	0.034	-0.001		
(1.09)	(-1.01)	(1.65)	(-0.08)		
0.089	-0.012			-0.387	
(1.59)	(-0.65)			(-2.43)[2]	
0.093	-0.003				-0.394
(2.74)	(-0.15)				(-12.93)[3]
0.107	0.012			-0.157	-0.406
(2.64)	(0.90)			(-0.95)	(-14.86)[3]
0.124	0.013	0.024		-0.102	-0.408
(2.49)	(0.94)	(1.30)		(-0.74)	(-14.77)[3]

Table VII
Portfolios and regressions for market capitalization

Panel A shows characteristics of portfolios sorted on market capitalization. The portfolios are created annually according to the methodology in section 3.4.1 and the variable abbreviations are listed in Table I, section 3.3.3. The numbers in the table below are the portfolio averages for the 1994-2008 period. The upper variable is the sorting variable, on which the portfolios are sorted from low to high. The right side of the table reports the difference between the highest and lowest portfolio for each variable and the statistical significance of the resulting number. Panel B displays Fama-Macbeth regressions created in line with the methodology in section 3.4.3. Every second line represents a regression with the regression variables in the column headings and the statistical significance reported in brackets below the coefficients.

Table VII: Panel A - Characteristics of portfolios sorted on market capitalization. 1994-2008

	Low			Variable sorted portfolios				High			Significance		
	1	2	3	4	5	6	7	8	9	10	H-L	T-stat	P
MV	20	59	109	185	301	478	786	1,432	3,191	25164	25143	12.16	0.000
F6M	-0.075	-0.088	-0.055	-0.058	-0.080	-0.068	-0.061	-0.053	-0.033	-0.021	0.053	1.10	0.290
F12M	0.092	0.024	0.020	-0.022	-0.055	-0.041	-0.028	-0.019	0.002	0.015	-0.078	-1.34	0.204
F24M	0.242	0.088	0.050	-0.005	-0.057	-0.020	-0.032	0.001	0.047	0.052	-0.190	-2.00	0.067
F36M	0.384	0.148	0.115	0.053	-0.019	0.044	0.008	0.060	0.104	0.126	-0.257	-2.13	0.053
H12M	-0.255	-0.117	-0.065	0.009	0.062	0.091	0.123	0.145	0.135	0.151	0.406	6.84	0.000
H36M	-0.433	-0.132	-0.023	0.142	0.269	0.382	0.420	0.431	0.439	0.477	0.910	8.51	0.000
BETA	0.892	0.942	0.975	1.006	1.013	0.959	1.034	0.985	0.972	0.905	0.013	0.16	0.879
LIQ	0.968	0.111	0.125	0.149	0.176	0.194	0.199	0.192	0.187	0.172	-0.795	-1.26	0.231
VOL	0.222	0.181	0.166	0.164	0.153	0.140	0.135	0.118	0.102	0.085	-0.137	-16.46	0.000
ACC	-0.047	-0.032	-0.030	-0.025	-0.024	-0.023	-0.028	-0.029	-0.031	-0.037	0.010	0.91	0.381
TASG	-0.001	0.081	0.120	0.165	0.189	0.198	0.211	0.193	0.165	0.148	0.149	6.51	0.000
NSI	0.081	0.079	0.079	0.065	0.069	0.057	0.058	0.045	0.033	0.031	-0.050	-3.55	0.004
CEG	-0.064	0.056	0.078	0.160	0.195	0.228	0.219	0.200	0.167	0.171	0.235	5.08	0.000
NOA	0.596	0.583	0.592	0.633	0.724	0.714	0.721	0.816	0.796	0.734	0.138	3.34	0.005
ROA	-0.134	-0.087	-0.079	-0.041	0.023	0.048	-0.020	0.076	0.064	0.077	0.211	7.35	0.000
BM	1.120	0.741	0.640	0.572	0.511	0.478	0.444	0.429	0.416	0.353	-0.767	-7.02	0.000
EP	0.216	0.080	0.080	0.092	0.062	0.056	0.056	0.056	0.053	0.052	-0.165	-2.59	0.023
CP	0.234	0.170	0.204	0.142	0.124	0.116	0.115	0.112	0.114	0.107	-0.127	-4.08	0.001
SP	11.48	2.327	2.396	1.660	1.495	1.297	1.128	1.140	1.140	0.895	-10.58	-2.14	0.052
NEST	1.492	2.167	2.894	3.482	4.501	5.314	6.764	8.484	12.15	18.89	17.39	26.43	0.000
SDEST	0.839	0.578	0.620	0.367	0.299	0.264	0.195	0.176	0.149	0.109	-0.730	-3.84	0.002
EEPSG	0,279	0.426	0.364	0.354	0.366	0.351	0.325	0.306	0.274	0.255	-0.023	-0.12	0.909
EERRD	0,053	0.400	0.249	0.229	0.168	0.142	0.096	0.069	0.061	0.026	-0.026	-0.07	0.943
EERR	0,462	0.501	0.360	0.336	0.272	0.276	0.235	0.192	0.162	0.130	-0.332	-4.22	0.001
LEERD	0,179	0.251	0.124	0.140	0.101	0.083	0.055	0.044	0.035	0.015	-0.164	-1.26	0.231

Table VII: Panel B – Fama-MacBeth Annual Return Regressions based market capitalization

Intercept	MV	BM	TASG	EERRD	NEST
0.062	-0.001	0.053			
(0.73)	(-0.12)	(1.83)[1]			
0.090	-0.009		-0.159		
(1.06)	(-1.11)		(-4.24)[3]		
0.168	-0.012			-0.435	
(2.39)	(-1.42)			(-15.37)[3]	
0.005	-0.009				0.004
(0.06)	(-0.65)				(1.57)

The portfolio sorts show that, there is a negative, although not strong significant, market value effect over the longer-term, in line with previous research (Banz, 1981). This fact is also consistent with the logic of the FF3M model, predicting that market capitalization is a factor with explanatory power on stock returns. There are however, a number of aspects that puts the size effects inclusion the FF3M into question. The first problem is that market value is correlated with book-to-market. This can be seen clearly in the portfolio characteristics table VII, and the effect of it is that market value has basically a zero relation to returns when combined with BM in Fama-MacBeth regressions (see table VII, panel B). The second problem is that the returns are not decreasing linearly with higher market capitalization. Instead, the middle portfolios exhibit the lowest returns. Thus, the size effect is solely due to the micro cap stocks having higher returns than the rest. If these stocks are removed from the sample, the size effect will reverse, leading to higher returns for bigger firms. The FF3M, predicting returns using factor loadings, requires that the effect on returns do not change direction with higher exposure to the factor. There might be distress risk related reasons for the high returns in the lowest market capitalization portfolios as these stocks seem to be financially constrained. They have low profitability, poor historical returns, and a considerable amount of share issuances; possibly in order to provide improved liquidity for the firms. The uncertainty about the future is also manifested in the very high dispersion of analyst earnings forecasts.

The characteristics of the analyst variables for market capitalization sorted portfolios also need to be explained. The smaller the firm is, the higher the earnings growth expected by analysts. These forecast do however tend to be more and more overly optimistic the smaller the company is, resulting in large estimate errors. It may seem puzzling that, although analysts tend to overestimate future earnings of the smallest companies considerably, the stock returns of the smallest companies are higher than for the rest of the sample. This could be explained by a general tendency of analysts to issue buy recommendations for small companies in order to generate business. A sell recommendation for a small company will likely not lead to nearly as much trading volume as a buy recommendation because the pool of possible buyers is immensely larger than the pool of possible sellers.

Table VIII
Portfolios and regressions for Liquidity

Panel A shows characteristics of portfolios sorted on Liquidity. The portfolios are created annually according to the methodology in section 3.4.1 and the variable abbreviations are listed in Table I, section 3.3.3. The numbers in the table below are the portfolio averages for the 1994-2008 period. The upper variable is the sorting variable, on which the portfolios are sorted from low to high. The right side of the table reports the difference between the highest and lowest portfolio for each variable and the statistical significance of the resulting number. Panel B displays Fama-Macbeth regressions created in line with the methodology in section 3.4.3. Every second line represents a regression with the regression variables in the column headings and the statistical significance reported in brackets below the coefficients.

Table VIII: Panel A - Characteristics of portfolios sorted on Liquidity. 1994-2008

	Low			Variable sorted portfolios				High			Significance		
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**H-L**	**T-stat**	**P**
LIQ	0.015	0.033	0.051	0.070	0.092	0.115	0.147	0.192	0.270	0.531	0.516	9.51	0.000
F6M	-0.018	-0.035	-0.038	-0.048	-0.055	-0.052	-0.082	-0.092	-0.123	-0.130	-0.112	-1.83	0.091
F12M	0.096	0.056	0.043	0.031	0.010	0.009	-0.033	-0.069	-0.107	-0.149	-0.244	-2.92	0.012
F24M	0.219	0.139	0.107	0.114	0.031	0.057	-0.014	-0.057	-0.145	-0.222	-0.441	-3.22	0.007
F36M	0.341	0.252	0.201	0.196	0.100	0.117	0.036	-0.032	-0.119	-0.242	-0.584	-4.19	0.001
H12M	-0.041	-0.032	-0.061	-0.025	0.001	0.018	0.023	0.045	0.072	0.102	0.143	1.88	0.083
H36M	-0.055	-0.038	-0.050	0.015	-0.002	0.092	0.118	0.210	0.329	0.484	0.539	5.95	0.000
BETA	0.689	0.819	0.947	1.009	1.050	1.103	1.178	1.230	1.341	1.517	0.828	16.99	0.000
MV	405	2,098	3,102	2,784	2,609	2,203	2,010	1,595	1,610	1,721	1.315	6.51	0.000
VOL	0.144	0.139	0.137	0.137	0.138	0.144	0.156	0.165	0.180	0.229	0.086	4.63	0.000
ACC	-0.028	-0.035	-0.036	-0.039	-0.036	-0.034	-0.028	-0.023	-0.017	-0.024	0.004	0.39	0.706
TASG	0.091	0.090	0.089	0.095	0.116	0.138	0.144	0.202	0.257	0.354	0.263	8.13	0.000
NSI	0.106	0.051	0.046	0.055	0.050	0.047	0.060	0.070	0.096	0.115	0.008	0.25	0.809
CEG	0.076	0.043	0.054	0.078	0.098	0.126	0.177	0.197	0.259	0.323	0.247	3.82	0.002
NOA	0.671	0.651	0.620	0.625	0.679	0.725	0.664	0.680	0.775	0.827	0.156	1.41	0.182
ROA	-0.048	-0.018	-0.023	-0.011	-0.004	0.004	-0.013	-0.004	-0.023	-0.021	0.027	0.52	0.612
BM	0.751	0.799	0.647	0.587	0.557	0.519	0.493	0.458	0.407	0.338	-0.412	-6.09	0.000
EP	0.165	0.096	0.085	0.077	0.062	0.086	0.064	0.060	0.048	0.044	-0.121	-1.47	0.166
CP	0.180	0.170	0.147	0.142	0.137	0.123	0.122	0.108	0.095	0.082	-0.098	-3.82	0.002
SP	3.366	11.37	1.964	1.694	4.499	1.441	1.492	1.251	1.010	0.935	-2.431	-2.87	0.013
NEST	2.921	4.616	6.001	6.985	7.481	7.603	7.712	7.712	8.117	8.906	5.986	11.54	0.000
SDEST	0.547	0.253	0.269	0.268	0.291	0.222	0.283	0.306	0.345	0.293	-0.254	-1.05	0.314
EEPSG	0.247	0.291	0.260	0.266	0.288	0.311	0.351	0.373	0.433	0.449	0.203	5.76	0.000
EERRD	0.018	0.089	0.067	0.079	0.081	0.082	0.130	0.132	0.152	0.212	0.194	2.71	0.018
EERR	0.263	0.191	0.170	0.177	0.209	0.212	0.246	0.257	0.272	0.347	0.084	1.38	0.192
LEERD	-0.014	0.045	0.041	0.037	0.049	0.056	0.074	0.082	0.085	0.122	0.136	2.81	0.016

Table VIII: Panel B – Fama-MacBeth Annual Return Regressions based on Liquidity

Intercept	**LIQ**	**BM**	**MV**	**EERRD**	**TASG**	**NEST**	**H36M**
0.093	-0.402	0.043					
(1.51)	(-2.76)[1]	(2.07)[1]					
0.091	-0.376	0.041	-0.001				
(1.02)	(-2.54)[1]	(1.73)	(-0.09)				
0.126	-0.283			-0.438			
(3.30)	(-1.66)			(-16.25)[3]			
0.082	-0.380				-0.146		
(1.51)	(-2.16)[1]				(-4.36)[3]		
0,025	-0,371					0,004	-0,023
(0,44)	(-2,03)[1]					(2,34)[2]	(-1,44)

There is a strong relation between stock liquidity and future returns, getting stronger the longer the time period of future returns. With an annual return spread of almost 25 % between the low and high liquidity portfolio, the difference seems too large for a liquidity risk premium. It is more likely that the liquidity measure predicts returns so well because it captures part of the effects on return of other, correlated, variables. Most of the variable characteristics for the highest liquidity portfolio predict negative returns; the historical 36 months returns of stocks in the portfolio are high, their market values are low, their asset growth and capital expenditure high, and book-to-market low, suggesting that these variables might have a significant impact on the returns of liquidity sorted portfolios. There is also a strong overoptimism pattern for liquidity sorted portfolios, where earnings estimates tending to be more overoptimistic the more liquid the stock is.

Even though the liquidity effect on stock returns may come from exposure to the variables characteristics of the portfolios, the liquidity measure may have predictive abilities in itself by identifying overhyped or forgotten companies. This seems to be supported by the Fama-MacBeth regressions, where the liquidity effect tends to remain relatively robust even after adjusted for other factors. The stocks in the liquid portfolios seem to have most characteristics of what makes up interesting investments. Firstly, they have grown quickly resulting in high historical returns. Secondly; analysts believe the growth will continue in the future, resulting in high valuations in terms of market multiples. Thirdly; even though they are not the biggest companies on the stock market, they have attracted a considerable number of analysts, securing a steady news flow to the investors. Finally, the stocks are both heavily traded and volatile, securing a thrilling ride for the investor. It is not difficult to imagine that these companies, if the stock market is influenced by behavioral biases driving up prices for glamorous companies too much, are excellent candidates for the sell list. The exact opposite is true for the illiquid stocks, which are boring investments in every respect, except for that investors actually make money by investing in those stocks.

4.5 Market Multiples

Table IX
Portfolios and regressions for Book to Market

Panel A shows characteristics of portfolios sorted on Book to Market. The portfolios are created annually according to the methodology in section 3.4.1 and the variable abbreviations are listed in Table I, section 3.3.3. The numbers in the table below are the portfolio averages for the 1994-2008 period. The upper variable is the sorting variable, on which the portfolios are sorted from low to high. The right side of the table reports the difference between the highest and lowest portfolio for each variable and the statistical significance of the resulting number. Panel B displays Fama-Macbeth regressions created in line with the methodology in section 3.4.3. Every second line represents a regression with the regression variables in the column headings and the statistical significance reported in brackets below the coefficients.

Table IX: Panel A - Characteristics of portfolios sorted on Book to Market. 1994-2008

	Low			Variable sorted portfolios				High			Significance		
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**H-L**	**T-stat**	**P**
BM	0.089	0.181	0.253	0.321	0.394	0.476	0.569	0.691	0.882	1.677	8	8.89	0.000
F6M	-0.089	-0.069	-0.072	-0.063	-0.056	-0.042	-0.035	-0.027	-0.028	-0.030	0.059	0.96	0.352
F12M	-0.087	-0.059	-0.033	-0.011	-0.002	0.026	0.039	0.043	0.065	0.106	0.194	2.07	0.058
F24M	-0.134	-0.070	-0.024	0.033	0.052	0.089	0.101	0.116	0.151	0.228	0.362	2.30	0.039
F36M	-0.129	-0.040	0.024	0.077	0.103	0.155	0.196	0.223	0.264	0.392	0.521	3.47	0.004
H12M	0.185	0.127	0.065	0.074	0.040	0.032	0.020	-0.013	-0.033	-0.106	-0.290	-3.74	0.004
H36M	0.597	0.506	0.400	0.351	0.241	0.200	0.146	0.045	-0.083	-0.341	-0.938	-8.57	0.000
MV	6,501	7,103	5,388	4,389	3,155	2,434	2,110	1,835	1,177	454	-6.046	-5.75	0.000
LIQ	0.233	0.372	0.188	0.316	0.154	0.141	0.140	0.116	0.108	0.097	-0.136	-11.89	0.000
BETA	1.150	1.116	1.085	1.009	0.981	0.925	0.890	0.866	0.891	0.859	-0.291	-5.79	0.000
VOL	0.176	0.154	0.143	0.134	0.130	0.126	0.122	0.123	0.130	0.161	-0.016	-1.07	0.303
ACC	-0.015	-0.021	-0.027	-0.027	-0.024	-0.034	-0.026	-0.031	-0.034	-0.040	-0.024	-2.23	0.044
TASG	0.260	0.242	0.209	0.172	0.175	0.133	0.118	0.098	0.083	0.011	-0.249	-8.39	0.000
NSI	0.092	0.076	0.069	0.049	0.054	0.045	0.043	0.045	0.038	0.024	-0.068	-6.13	0.000
CEG	0.323	0.277	0.227	0.181	0.187	0.132	0.112	0.082	0.055	-0.110	-0.433	-8.57	0.000
NOA	0.577	0.619	0.717	0.706	0.748	0.774	0.742	0.794	0.753	0.726	0.149	3.03	0.010
ROA	0.032	0.010	0.022	0.030	0.036	0.032	0.032	0.025	0.009	-0.014	-0.047	-1.76	0.101
EP	0.089	0.181	0.253	0.321	0.394	0.476	0.569	0.691	0.882	1.677	1.588	8.89	0.000
CP	0.038	0.042	0.045	0.052	0.054	0.060	0.066	0.079	0.086	0.183	0.146	3.96	0.002
SP	0.058	0.066	0.080	0.092	0.102	0.118	0.134	0.151	0.180	0.331	0.273	3.94	0.002
NEST	8.838	9.123	8.987	8.872	8.378	7.791	7.621	6.910	5.917	4.595	-4.243	-8.70	0.000
SDEST	0.275	0.239	0.206	0.190	0.227	0.234	0.220	0.206	0.311	0.452	0.176	2.38	0.034
EEPSG	0.427	0.374	0.332	0.315	0.288	0.269	0.250	0.242	0.246	0.255	-0.171	-3.71	0.003
EERRD	0.088	0.086	0.104	0.089	0.090	0.093	0.081	0.104	0.077	-0.197	-0.285	-0.86	0.409
EERR	0.219	0.188	0.198	0.196	0.205	0.217	0.216	0.221	0.258	0.402	0.184	2.81	0.016
LEERD	0.062	0.060	0.056	0.053	0.058	0.049	0.054	0.063	0.027	0.044	-0.018	-0.34	0.742

Table IX: Panel B – Fama-MacBeth Annual Return Regressions based on Book to Market

Intercept	**BM**	**MV**	**LIQ**	**TASG**	**EERRD**
0.062	0.053	-0.001			
(0.73)	(1.83)[1]	(-0.12)			
0.093	0.043		-0.402		
(1.51)	(2.07)[1]		(-2.76)[2]		
0.075	0.044			-0.138	
(1.36)	(1.91)[1]			(-4.31)[3]	
0.127	0.042				-0.425
(3.02)	(1.59)				(-14.81)[3]

Table X
Portfolios and regressions for Cash flow to Price

Panel A shows characteristics of portfolios sorted on Cash flow to Price. The portfolios are created annually according to the methodology in section 3.4.1 and the variable abbreviations are listed in Table I, section 3.3.3. The numbers in the table below are the portfolio averages for the 1994-2008 period. The upper variable is the sorting variable, on which the portfolios are sorted from low to high. The right side of the table reports the difference between the highest and lowest portfolio for each variable and the statistical significance of the resulting number. Panel B displays Fama-Macbeth regressions created in line with the methodology in section 3.4.3. Every second line represents a regression with the regression variables in the column headings and the statistical significance reported in brackets below the coefficients.

Table X: Panel A - Characteristics of portfolios sorted on Cash flow to Price. 1994-2008

	Low			**Variable sorted portfolios**				**High**			**Significance**		
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**H-L**	**T-stat**	**P**
CP	0.020	0.043	0.061	0.076	0.091	0.108	0.128	0.156	0.201	0.471	0.451	5.95	0.000
F6M	-0.101	-0.057	-0.056	-0.041	-0.030	-0.030	-0.022	-0.018	-0.014	-0.022	0.079	1.36	0.196
F12M	-0.102	-0.025	-0.005	0.015	0.026	0.045	0.062	0.071	0.077	0.117	0.220	2.39	0.033
F24M	-0.170	-0.028	0.036	0.067	0.091	0.120	0.139	0.159	0.178	0.235	0.404	3.11	0.008
F36M	-0.177	-0.027	0.078	0.144	0.176	0.214	0.256	0.264	0.305	0.387	0.564	4.74	0.000
H12M	0.103	0.087	0.093	0.089	0.080	0.069	0.057	0.073	0.039	0.032	-0.071	-0.90	0.385
H24M	0.452	0.439	0.397	0.364	0.337	0.280	0.251	0.222	0.174	0.019	-0.433	-3.75	0.002
MV	3,255	6,474	5,510	4,339	4,586	4,571	3,616	3,369	2,572	1,737	-1.517	-2.32	0.038
LIQ	0.240	0.385	0.173	0.144	0.147	0.126	0.116	0.112	0.115	0.121	-0.119	-6.21	0.000
BETA	1.345	1.184	1.019	0.955	0.894	0.848	0.811	0.756	0.766	0.846	-0.499	-7.48	0.000
VOL	0.178	0.146	0.125	0.116	0.112	0.109	0.107	0.106	0.113	0.144	-0.034	-3.31	0.016
ACC	-0.005	-0.009	-0.014	-0.020	-0.021	-0.028	-0.024	-0.035	-0.040	-0.053	-0.049	-6.95	0.000
TASG	0.318	0.247	0.196	0.158	0.140	0.116	0.121	0.111	0.095	0.065	-0.253	-7.21	0.000
NSI	0.088	0.061	0.052	0.036	0.028	0.024	0.028	0.029	0.022	0.045	-0.042	-3.22	0.007
CEG	0.306	0.234	0.239	0.162	0.176	0.126	0.129	0.142	0.092	0.030	-0.276	-3.89	0.002
NOA	0.782	0.712	0.741	0.727	0.733	0.731	0.754	0.759	0.757	0.706	-0.076	-0.65	0.526
ROA	0.037	0.095	0.098	0.095	0.092	0.091	0.078	0.067	0.060	0.033	-0.04	-0.26	0.797
EP	0.023	0.033	0.040	0.048	0.055	0.061	0.066	0.077	0.083	0.156	0.133	4.70	0.000
BM	0.330	0.328	0.365	0.421	0.469	0.518	0.569	0.650	0.765	1.182	0.852	6.13	0.000
SP	0.596	0.641	0.811	1.010	1.252	1.500	1.659	2.019	2.542	5.684	5.088	4.19	0.001
NEST	7.737	8.977	9.315	8.723	8.452	8.169	8.299	8.269	8.514	6.946	-0.790	-1.11	0.286
SDEST	0.340	0.146	0.116	0.103	0.117	0.101	0.101	0.148	0.192	0.295	-0.044	-0.72	0.483
EEPSG	0.663	0.434	0.353	0.320	0.294	0.269	0.250	0.223	0.180	0.156	-0.508	-10.06	0.000
EERRD	0.198	0.121	0.095	0.088	0.081	0.081	0.076	0.074	0.074	0.059	-0.139	-2.41	0.033
EERR	0.331	0.234	0.192	0.180	0.184	0.184	0.192	0.200	0.233	0.297	-0.035	-1.65	0.124
LEERD	0.140	0.072	0.045	0.056	0.048	0.049	0.037	0.034	0.033	0.018	-0.122	-3.19	0.008

Table X: Panel B – Fama-MacBeth Annual Return Regressions based on Cash flow to Price

Intercept	**CP**	**BM**	**MV**	**ACC**	**LIQ**	**TASG**	**EERRD**
0.160	0.053	0.031					
(2.01)	(2.24)[2]	(0.98)					
0.182	0.052	0.006	-0.004				
(1.94)	(2.29)[2]	(0.66)	(-0.55)				
0.139	0.050			-0.261			
(1.69)	(1.91)			(-4.78)[3]			
0.207	0.052						
(2.59)	(1.74)						
0.187	0.051				-0.295		
(2.18)	(2.42)				(-2.64)[2]		
0.161	0.048					-0.115	
(1.99)	(1.93)					(-4.82)[3]	
0.207	0.052						-0.420
(2.59)	(1.74)						(-14.40)[3]

Table XI

Portfolios and regressions for Earnings to Price

Panel A shows characteristics of portfolios sorted on Earnings to Price. The portfolios are created annually according to the methodology in section 3.4.1 and the variable abbreviations are listed in Table I, section 3.3.3. The numbers in the table below are the portfolio averages for the 1994-2008 period. The upper variable is the sorting variable, on which the portfolios are sorted from low to high. The right side of the table reports the difference between the highest and lowest portfolio for each variable and the statistical significance of the resulting number. Panel B displays Fama-Macbeth regressions created in line with the methodology in section 3.4.3. Every second line represents a regression with the regression variables in the column headings and the statistical significance reported in brackets below the coefficients.

Table XI: Panel A - Characteristics of portfolios sorted on Earnings to Price. 1994-2008

	Low		Variable sorted portfolios					High			Significance		
	1	2	3	4	5	6	7	8	9	10	H-L	T-stat	P
EP	0.010	0.023	0.033	0.041	0.049	0.056	0.064	0.074	0.090	0.264	0.255	7.02	0.000
F6M	-0.088	-0.061	-0.033	-0.046	-0.033	-0.038	-0.002	-0.010	-0.029	-0.036	0.052	1.42	0.178
F12M	-0.067	-0.038	0.025	0.015	0.040	0.024	0.062	0.061	0.062	0.080	0.146	2.50	0.027
F24M	-0.073	-0.036	0.049	0.066	0.107	0.097	0.155	0.118	0.136	0.151	0.224	2.31	0.038
F36M	-0.041	-0.017	0.109	0.134	0.197	0.170	0.243	0.234	0.256	0.275	0.316	3.06	0.009
H12M	0.107	0.110	0.120	0.102	0.094	0.062	0.085	0.060	0.036	-0.011	-0.118	-2.00	0.067
H36M	0.386	0.434	0.476	0.393	0.389	0.348	0.335	0.325	0.305	0.258	-0.128	-1.41	0.181
MV	3,517	5,385	5,956	5,907	4,747	4,879	4,732	3,542	3,486	2,391	-1.125	-1.22	0.245
LIQ	0.211	0.488	0.180	0.153	0.143	0.139	0.129	0.120	0.120	0.132	-0.078	-4.84	0.000
BETA	1.146	1.117	1.011	0.911	0.837	0.794	0.757	0.738	0.755	0.870	-0.277	-5.42	0.000
VOL	0.155	0.137	0.125	0.115	0.106	0.106	0.099	0.100	0.108	0.138	-0.016	-1.77	0.100
ACC	-0.019	-0.014	-0.025	-0.023	-0.024	-0.030	-0.024	-0.028	-0.024	-0.010	0.009	1.23	0.241
TASG	0.261	0.233	0.191	0.172	0.145	0.138	0.119	0.116	0.122	0.114	-0.147	-6.21	0.000
NSI	0.071	0.059	0.048	0.036	0.026	0.020	0.017	0.014	0.012	0.024	-0.047	-4.12	0.001
CEG	0.270	0.236	0.220	0.178	0.161	0.145	0.146	0.154	0.156	0.191	-0.080	-1.78	0.099
NOA	0.762	0.747	0.747	0.725	0.723	0.736	0.717	0.721	0.721	0.682	-0.080	-1.73	0.107
ROA	0.042	0.075	0.088	0.092	0.095	0.094	0.089	0.109	0.120	0.154	0.111	3.30	0.006
BM	0.404	0.386	0.388	0.425	0.442	0.511	0.529	0.583	0.673	0.943	0.539	4.63	0.000
CP	0.076	0.078	0.084	0.099	0.103	0.117	0.130	0.148	0.177	0.277	0.201	6.63	0.000
SP	0.913	0.906	0.922	1.172	1.202	1.314	1.428	1.702	2.228	3.428	2.516	5.43	0.000
NEST	8.168	9.068	9.661	9.273	8.854	8.994	8.576	8.002	7.583	6.557	-1.611	-2.30	0.038
SDEST	0.250	0.116	0.115	0.092	0.081	0.070	0.069	0.086	0.126	0.173	-0.076	-1.89	0.081
EEPSG	0.746	0.498	0.397	0.335	0.301	0.265	0.230	0.201	0.160	0.071	-0.675	-16.37	0.000
EERRD	0.166	0.147	0.093	0.096	0.063	0.092	0.055	0.076	0.106	0.059	-0.107	-1.73	0.109
EERR	0.359	0.262	0.205	0.195	0.168	0.175	0.151	0.179	0.215	0.313	-0.046	-1.66	0.123
LEERD	0.097	0.093	0.044	0.063	0.041	0.046	0.022	0.035	0.058	0.026	-0.071	-1.44	0.175

Table XI: Panel B – Fama-MacBeth Annual Return Regressions based on Earnings to Price

Intercept	EP	BM	MV	EERRD	TASG	LIQ
0.157	0.036	0.023				
(2.28)	$(2.56)^2$	(1.56)				
0.148	0.034	0.024	0.001			
(1.85)	$(2.43)^2$	(1.43)	(0.12)			
0.217	0.044			-0.425		
(2.45)	(1.74)			$(-14.41)^3$		
0.181	0.042				-0.127	
(2.53)	$(2.46)^2$				$(-4.24)^3$	
0.184	0.037					-0.379
(2.50)	$(2.61)^2$					$(-2.82)^2$

Table XII
Portfolios and regressions for Sales to Price

Panel A shows characteristics of portfolios sorted on Sales to Price. The portfolios are created annually according to the methodology in section 3.4.1 and the variable abbreviations are listed in Table I, section 3.3.3. The numbers in the table below are the portfolio averages for the 1994-2008 period. The upper variable is the sorting variable, on which the portfolios are sorted from low to high. The right side of the table reports the difference between the highest and lowest portfolio for each variable and the statistical significance of the resulting number. Panel B displays Fama-Macbeth regressions created in line with the methodology in section 3.4.3. Every second line represents a regression with the regression variables in the column headings and the statistical significance reported in brackets below the coefficients.

Table XII: Panel A - Characteristics of portfolios sorted on Sales to Price. 1994-2008

	Low		Variable sorted portfolios					High			Significance		
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**H-L**	**T-stat**	**P**
SP	0.068	0.205	0.345	0.493	0.676	0.908	1.215	1.677	2.591	15.79	15.73	3.35	0.005
F6M	-0.116	-0.097	-0.047	-0.037	-0.041	-0.030	-0.039	-0.023	-0.042	-0.033	0.083	1.15	0.271
F12M	-0.148	-0.084	-0.009	0.016	0.022	0.030	0.041	0.074	0.061	0.102	0.250	1.99	0.068
F24M	-0.212	-0.093	0.006	0.063	0.082	0.092	0.106	0.146	0.153	0.223	0.434	2.41	0.031
F36M	-0.244	-0.062	0.026	0.131	0.163	0.191	0.172	0.261	0.242	0.396	0.640	3.64	0.003
H12M	0.023	0.062	0.058	0.064	0.056	0.038	0.054	0.023	0.026	-0.021	-0.044	-0.40	0.679
H24M	0.333	0.420	0.298	0.277	0.204	0.162	0.196	0.093	0.062	-0.129	-0.462	-2.74	0.017
MV	2,170	5,153	5,751	5,031	4,323	4,219	3,001	2,228	1,413	743	-1.427	-3.84	0.002
LIQ	0.235	0.219	0.174	0.306	0.137	0.133	0.124	0.274	0.116	0.123	-0.112	-6.30	0.000
BETA	1.300	1.171	1.079	1.014	0.947	0.901	0.869	0.858	0.871	0.932	-0.368	-7.51	0.000
VOL	0.199	0.165	0.143	0.132	0.124	0.119	0.121	0.127	0.133	0.165	-0.035	-2.57	0.023
ACC	-0.027	-0.022	-0.032	-0.026	-0.031	-0.031	-0.028	-0.030	-0.032	-0.036	-0.009	-0.88	0.394
TASG	0.311	0.259	0.190	0.153	0.129	0.108	0.100	0.081	0.071	0.044	-0.268	-4.66	0.000
NSI	0.131	0.075	0.065	0.055	0.047	0.036	0.033	0.032	0.030	0.051	-0.081	-6.34	0.000
CEG	0.329	0.303	0.206	0.134	0.095	0.104	0.101	0.086	0.033	-0.009	-0.338	-3.62	0.003
NOA	0.617	0.788	0.722	0.735	0.733	0.713	0.729	0.699	0.689	0.614	-0.003	-0.03	0.974
ROA	-0.145	0.028	0.035	0.043	0.043	0.041	0.040	0.031	0.029	-0.001	0.143	5.59	0.000
BM	0.252	0.296	0.365	0.413	0.465	0.504	0.573	0.671	0.781	1.238	0.986	6.80	0.000
CP	0.035	0.054	0.076	0.092	0.105	0.116	0.134	0.153	0.183	0.336	0.301	4.49	0.001
EP	0.039	0.036	0.044	0.051	0.060	0.060	0.065	0.078	0.086	0.177	0.138	4.12	0.001
NEST	6.537	8.984	8.952	8.830	8.870	8.765	7.907	6.969	6.152	5.225	-1.312	-3.65	0.003
SDEST	0.538	0.357	0.255	0.235	0.178	0.140	0.162	0.150	0.175	0.255	-0.282	-3.54	0.004
EEPSG	0.613	0.408	0.333	0.294	0.282	0.267	0.275	0.279	0.285	0.281	-0.332	-6.64	0.000
EERRD	0.140	0.141	0.089	0.083	0.078	0.066	0.071	0.091	0.123	0.116	-0.024	-0.50	0.628
EERR	0.314	0.244	0.214	0.193	0.179	0.177	0.179	0.229	0.264	0.278	-0.036	-1.33	0.208
LEERD	0.085	0.085	0.055	0.042	0.039	0.045	0.036	0.048	0.064	0.073	-0.012	-0.29	0.775

Table XII Panel B – Fama-MacBeth Annual Return Regressions based on Sales to Price

Intercept	SP	BM	MV	EERRD	TASG	LIQ
0.042	0.038	0.017				
(0.76)	(1.66)	(1.29)				
0.049	0.038	0.014	-0.002			
(0.56)	(1.68)	(1.05)	(-0.21)			
0.089	0.025			-0.421		
(2.69)	(1.09)			(-15.30)[3]		
0.048	0.036				-0.126	
(0.95)	(1.67)				(-4.74)[3]	
0.071	0.039					-0.350
(1.28)	(1.98)					(-2.92)[2]

The market multiples, even though highly correlated with balance sheet measures of growth such as growth in assets and capital expenditures, seems to capture the value/growth effect in a different way because the return effect is not only affecting the high growth portfolios, but is present throughout the sample from the highest to the lowest portfolio. Asset growth is highly significant when regressed together with the market multiples, and the significance of the multiples does not drop considerably. Contrary to the balance sheet growth sorted portfolios, there is a strong connection between the market multiples and stock liquidity. Thus, the effect of the market multiples on returns is difficult to gauge without taking liquidity into account, as this measure also has a strong connection to stock returns. However, the fact that both the liquidity measure and the market multiples tend not to drop considerably in significance when run in the same regressions indicates that the effects are largely independent.

Although the market multiples tend to have fundamentally the same variable characteristics for the same portfolios making up the cross-section, there are also some differences. The market values of the portfolios sorted on book-to-market are clearly declining for every portfolio going from low to high book-to-market. The market values of stocks in the low book-to-market portfolio are on average more than ten times the average value in the high portfolio. This relationship is far weaker for the other market multiples. The correlation with historical returns also seems to be stronger for BM than for the other multiples. CP is different to the other market multiples in that it has a strong relation to accruals. This is quite logical since low accrual firms tend to have higher cash flows and vice versa.

When it comes to the source of the stock return effects, there is a fundamental difference between the BM and SP variables and the other multiples. Both the CP and EP effect on stock returns could easily be explained by overoptimism from the investor's part, because the upward bias in the earnings estimates seems to increase as the returns in the cross-section decrease. For BM and SP however, the portfolio sorts fail to provide an analyst misestimation explanation for the return distribution among the portfolios, as the differences in estimate errors are minuscule compared to the large return spread between the high and low portfolios. This is interesting as most other variables in this thesis with explanatory power for the cross-section of stock returns seem to be related to analyst errors. Thus, with the analyst estimate error approach, as the portfolio sorts provides indications that the effects on stock returns of most of the studied variables are due mispricing, it fails to provide a mispricing explanation for the BM and the SP effect. In the Fama-MacBeth regressions however, the impact of the market multiples on returns do not differ considerably when run in the same regression as the overoptimism variable.

From the Fama-MacBeth regressions can also be seen that all market multiples except EP just fall short of being significant when asset growth is included, but that most of them seem to do slightly better when regressed together with liquidity.

All of the four market multiples seem to be related to future returns, both in the short and the longer term. The statistic significance of the multiples also seems to increase gradually with time. Even though the book-to-market ratio is by far the most central market multiple in research literature, and the earnings-to-price the most dominant on Wall Street, there is no evidence from the portfolios that these multiples have stronger correspondence with returns than cash flow-to-price and sales-to-price. If anything, quite the contrary, as the hedge return spreads seem to be larger for the latter. CP has stronger predictive abilities than BM on 12 months forward returns as it does not lose much significance when regressed together with BM, while BM becomes highly insignificant. The same goes for EP, even though the results are not conclusive with the full sample regressions. In any case, there is nothing obvious in the data suggesting that BM should be the dominant market multiple in asset pricing research from a stock predictability perspective. The CP ratios strong stock return relation compared to BM is consistent with Lakonishok et al. (1994) findings using data from before the coverage period of this thesis.

4.6 Balance Sheet Accounting Based Measures

Below follows a detailed review of the characteristics of portfolios sorted by Accruals, Asset Growth, Capital Expenditures, Net Operating Assets, Net Stock Issues.

Table XIII
Portfolios and regressions for Accruals

Panel A shows characteristics of portfolios sorted on Accruals. The portfolios are created annually according to the methodology in section 3.4.1 and the variable abbreviations are listed in Table I, section 3.3.3. The numbers in the table below are the portfolio averages for the 1994-2008 period. The upper variable is the sorting variable, on which the portfolios are sorted from low to high. The right side of the table reports the difference between the highest and lowest portfolio for each variable and the statistical significance of the resulting number. Panel B displays Fama-Macbeth regressions created in line with the methodology in section 3.4.3. Every second line represents a regression with the regression variables in the column headings and the statistical significance reported in brackets below the coefficients.

Table XIII: Panel A - Characteristics of portfolios sorted on Accruals. 1994-2008

	Low			Variable sorted portfolios				High			Significance		
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**H-L**	**T-stat**	**P**
ACC	-0.218	-0.100	-0.071	-0.053	-0.039	-0.025	-0.010	0.009	0.039	0.167	0.385	35.10	0.000
F6M	-0.078	-0.018	-0.031	-0.006	-0.039	-0.044	-0.048	-0.040	-0.081	-0.095	-0.017	-0.99	0.339
F12M	-0.001	0.060	0.032	0.058	0.022	0.009	0.007	0.011	-0.027	-0.065	-0.064	-3.27	0.006
F24M	0.016	0.114	0.090	0.131	0.064	0.058	0.050	0.048	-0.025	-0.060	-0.076	-1.77	0.100
F36M	0.069	0.167	0.182	0.241	0.132	0.147	0.094	0.105	0.059	-0.012	-0.081	-1.41	0.181
H12M	0.001	0.044	0.077	0.047	0.020	0.050	0.055	0.014	-0.006	0.033	0.032	1.13	0.281
H36M	-0.097	0.084	0.165	0.187	0.171	0.196	0.210	0.215	0.258	0.340	0.436	7.53	0.000
MV	1,918	4,586	5,121	5,381	5,535	5,142	5,290	3,548	2,275	1,174	-745	-2.33	0.036
LIQ	0.167	0.158	0.145	0.140	0.140	0.144	0.137	0.155	0.177	0.194	0.027	2.44	0.030
BETA	1.083	1.018	0.970	0.986	0.935	0.960	0.930	1.026	1.054	1.150	0.067	1.50	0.157
VOL	0.175	0.142	0.132	0.122	0.120	0.123	0.122	0.133	0.143	0.175	0.001	0.10	0.925
TASG	0.072	0.093	0.122	0.094	0.108	0.128	0.108	0.161	0.195	0.296	0.223	8.64	0.000
NSI	0.059	0.044	0.034	0.036	0.032	0.033	0.038	0.053	0.055	0.105	0.046	6.03	0.000
CEG	0.020	0.052	0.086	0.091	0.108	0.088	0.103	0.163	0.228	0.307	0.286	6.56	0.000
NOA	0.513	0.643	0.648	0.645	0.699	0.700	0.669	0.745	0.732	0.919	0.406	8.22	0.000
ROA	-0.113	-0.007	-0.020	0.017	0.022	0.029	0.027	0.024	0.031	0.034	0.148	7.89	0.000
BM	0.616	0.538	0.548	0.557	0.551	0.533	0.530	0.512	0.515	0.565	-0.051	-0.83	0.424
EP	0.074	0.059	0.058	0.057	0.058	0.056	0.056	0.058	0.062	0.095	0.021	1.03	0.324
CP	0.160	0.140	0.151	0.139	0.125	0.117	0.112	0.106	0.109	0.120	-0.040	-1.93	0.076
SP	2.193	1.769	1.720	1.576	1.424	1.417	1.498	1.326	1.488	2.037	-0.157	-1.34	0.202
NEST	6,682	8,518	8,734	9,519	9,140	8,690	8,232	7,774	7,029	5,944	-0,738	-2,47	0.028
SDEST	0.381	0.288	0.252	0.193	0.275	0.207	0.191	0.222	0.245	0.177	-0.203	-3.65	0.003
EEPSG	0.361	0.304	0.289	0.280	0.272	0.285	0.285	0.308	0.336	0.382	0.020	0.47	0.649
EERRD	0.100	0.077	0.077	0.043	0.064	0.076	0.078	0.107	0.138	0.191	0.091	2.63	0.022
EERR	0.218	0.197	0.188	0.171	0.151	0.169	0.164	0.185	0.253	0.280	0.062	3.68	0.003
LEERD	0.054	0.042	0.063	0.022	0.035	0.036	0.040	0.078	0.085	0.141	0.088	3.85	0.002

Table XIII: Panel B – Fama-MacBeth Annual Return Regressions based on Accruals

Intercept	ACC	BM	MV	EERRD	TASG	CP
0.054	-0.286	0.055				
(0.95)	(-4.54) [3]	(2.48)				
0.021	-0.270	0.061	0.004			
(0.23)	(-4.29) [2]	(2.16)	(0.38)			
0.095	-0.169			-0.498		
(2.82)	(-1.58)			(-13.92) [3]		
0.033	-0.134				-0.159	
(0.66)	(-2.38)[2]				(-4.24) [3]	
0.139	-0.261					0.050
(1.69)	(-4.78) [3]					(1.91)[1]

Table XIV
Portfolios and regressions for Asset Growth

Panel A shows characteristics of portfolios sorted on Asset Growth. The portfolios are created annually according to the methodology in section 3.4.1 and the variable abbreviations are listed in Table I, section 3.3.3. The numbers in the table below are the portfolio averages for the 1994-2008 period. The upper variable is the sorting variable, on which the portfolios are sorted from low to high. The right side of the table reports the difference between the highest and lowest portfolio for each variable and the statistical significance of the resulting number. Panel B displays Fama-Macbeth regressions created in line with the methodology in section 3.4.3. Every second line represents a regression with the regression variables in the column headings and the statistical significance reported in brackets below the coefficients.

Table XIV: Panel A - Characteristics of portfolios sorted on Asset Growth. 1994-2008

	Low			Variable sorted portfolios				High			Significance		
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**H-L**	**T-stat**	**P**
TASG	-0.356	-0.081	-0.016	0.025	0.062	0.104	0.156	0.235	0.391	1.003	1.359	15.06	0.000
F6M	-0.036	-0.040	-0.028	-0.025	-0.019	-0.029	-0.041	-0.049	-0.082	-0.147	-0.110	-2.77	0.016
F12M	0.066	0.057	0.052	0.048	0.043	0.036	0.021	-0.006	-0.044	-0.175	-0.242	-3.38	0.005
F24M	0.125	0.133	0.152	0.128	0.099	0.107	0.062	0.023	-0.044	-0.235	-0.360	-3.84	0.002
F36M	0.166	0.220	0.248	0.215	0.206	0.189	0.141	0.087	0.026	-0.200	-0.366	-3.64	0.003
H12M	-0.081	0.009	0.052	0.059	0.055	0.058	0.075	0.089	0.047	0.030	0.111	2.85	0.014
H36M	-0.485	-0.127	0.048	0.133	0.185	0.248	0.337	0.443	0.576	0.692	1.177	12.44	0.000
MV	1,209	2,506	4,044	4,454	4,733	4,861	4,442	3,524	2,712	2,097	888	3.16	0.008
LIQ	0.236	0.128	0.115	0.119	0.125	0.133	0.150	0.179	0.211	0.245	0.009	0.09	0.932
BETA	1.190	1.034	0.915	0.841	0.862	0.900	0.966	0.999	1.082	1.151	-0.039	-0.57	0.579
VOL	0.206	0.149	0.123	0.112	0.112	0.118	0.127	0.135	0.150	0.182	-0.024	-1.91	0.079
ACC	-0.066	-0.051	-0.044	-0.036	-0.029	-0.025	-0.023	-0.021	-0.008	0.006	0.073	4.72	0.000
NSI	0.052	0.028	0.017	0.016	0.020	0.025	0.033	0.052	0.095	0.205	0.152	11.90	0.000
CEG	-0.326	-0.125	-0.061	0.045	0.103	0.128	0.183	0.283	0.391	0.733	1.059	19.43	0.000
NOA	0.295	0.493	0.580	0.602	0.639	0.644	0.674	0.707	0.825	1.514	1.218	7.10	0.000
ROA	-0.212	-0.021	0.022	0.035	0.056	0.060	0.064	0.076	0.045	-0.065	0.146	3.85	0.002
BM	0.695	0.708	0.690	0.625	0.616	0.522	0.485	0.435	0.383	0.385	-0.310	-8.91	0.000
EP	0.166	0.121	0.068	0.062	0.072	0.067	0.061	0.057	0.060	0.050	-0.116	-2.42	0.031
CP	0.179	0.178	0.162	0.191	0.143	0.128	0.118	0.105	0.099	0.085	-0.094	-7.48	0.000
SP	2.084	2.210	2.250	4.073	1.761	1.562	1.434	1.246	1.530	1.297	-0.787	-1.42	0.179
NEST	5.322	7.234	8.511	8.713	8.681	8.428	8.828	8.520	8.071	7.141	1.819	4.70	0.000
SDEST	0.604	0.345	0.219	0.219	0.179	0.139	0.142	0.212	0.231	0.325	-0.278	-4.25	0.001
EEPSG	0.224	0.276	0.251	0.255	0.258	0.276	0.296	0.333	0.366	0.456	0.232	6.89	0.000
EERRD	0.072	0.053	0.051	0.058	0.076	0.094	0.077	0.118	0.132	0.186	0.114	2.73	0.018
EERR	0.259	0.213	0.193	0.182	0.174	0.199	0.199	0.222	0.252	0.317	0.058	1.82	0.093
LEERD	0.016	0.027	0.023	0.028	0.042	0.056	0.029	0.076	0.087	0.120	0.103	2.59	0.024

Table XIV: Panel B – Fama-MacBeth Annual Return Regressions based on Asset Growth

Intercept	TASG	BM	MV	EERRD
0.075	-0.138	0.044		
(1.36)	(-4.31) [3]	(1.91) [1]		
0.075	-0.133	0.042	0.000	
(0.88)	(-4.04) [2]	(1.59)	(-0.04)	
0.118	-0.093			-0.401
(4.97)	(-2.47) [2]			(-15.57) [3]

Table XV
Portfolios and regressions for Capital Expenditure Growth

Panel A shows characteristics of portfolios sorted on Capital Expenditure Growth. The portfolios are created annually according to the methodology in section 3.4.1 and the variable abbreviations are listed in Table I, section 3.3.3. The numbers in the table below are the portfolio averages for the 1994-2008 period. The upper variable is the sorting variable, on which the portfolios are sorted from low to high. The right side of the table reports the difference between the highest and lowest portfolio for each variable and the statistical significance of the resulting number. Panel B displays Fama-Macbeth regressions created in line with the methodology in section 3.4.3. Every second line represents a regression with the regression variables in the column headings and the statistical significance reported in brackets below the coefficients.

Table XV: Panel A - Characteristics of portfolios sorted on Capital Expenditures Growth. 1994-2008

	Low			Variable sorted portfolios				High			Significance		
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**H-L**	**T-stat**	**P**
CEG	-1.346	-0.512	-0.243	-0.073	0.062	0.188	0.333	0.518	0.805	1.704	3.050	26.83	0.000
F6M	-0.038	-0.040	-0.026	-0.043	-0.036	-0.041	-0.043	-0.063	-0.097	-0.127	-0.089	-2.06	0.060
F12M	0.052	0.053	0.037	0.019	0.028	0.014	0.017	-0.009	-0.062	-0.132	-0.184	-2.82	0.014
F24M	0.089	0.123	0.106	0.100	0.068	0.052	0.071	0.028	-0.072	-0.156	-0.245	-3.21	0.007
F36M	0.129	0.207	0.177	0.163	0.170	0.134	0.129	0.111	0.001	-0.110	-0.239	-3.01	0.010
H12M	0.018	0.035	0.064	0.057	0.043	0.045	0.040	0.006	-0.002	-0.002	-0.021	-0.35	0.733
H24M	-0.373	-0.100	0.014	0.166	0.186	0.275	0.339	0.370	0.442	0.500	0.873	11.92	0.000
MV	548	1,177	2,471	4,248	3,674	3,719	2,981	2,170	1,566	804	256	2.69	0.019
LIQ	0.147	0.134	0.140	0.140	0.308	0.152	0.159	0.262	0.187	0.200	0.053	6.62	0.000
BETA	1.149	1.093	1.002	1.017	1.012	1.001	1.029	1.060	1.094	1.103	-0.046	-0.66	0.523
VOL	0.191	0.157	0.139	0.130	0.129	0.128	0.135	0.143	0.157	0.179	-0.013	-1.21	0.247
ACC	-0.046	-0.041	-0.038	-0.035	-0.033	-0.032	-0.033	-0.014	-0.014	0.000	0.046	3.91	0.002
NSI	0.076	0.054	0.039	0.037	0.033	0.044	0.049	0.057	0.095	0.126	0.050	3.79	0.002
TASG	-0.032	0.040	0.071	0.092	0.111	0.150	0.183	0.221	0.292	0.497	0.528	10.47	0.000
NOA	0.438	0.543	0.612	0.648	0.654	0.674	0.693	0.770	0.781	1.150	0.712	6.14	0.000
ROA	-0.120	-0.052	0.032	0.016	0.026	0.029	0.023	0.028	-0.003	-0.062	5.879	2.24	0.043
BM	0.712	0.656	0.658	0.572	0.526	0.513	0.507	0.491	0.467	0.444	-0.268	-5.16	0.000
EP	0.119	0.073	0.077	0.077	0.058	0.082	0.064	0.066	0.064	0.063	-0.056	-1.17	0.264
CP	0.287	0.149	0.152	0.147	0.129	0.130	0.127	0.117	0.111	0.103	-0.184	-1.36	0.196
SP	3.049	9.161	2.607	1.808	1.529	2.234	1.580	1.396	3.905	1.097	-1.951	-2.34	0.036
NEST	5.093	6.552	8.104	9.366	8.857	8.611	8.380	7.483	7.026	5.721	0.628	2.34	0.036
SDEST	0.515	0.342	0.242	0.222	0.209	0.180	0.221	0.232	0.349	0.357	-0.158	-1.87	0.084
EEPSG	0.358	0.318	0.316	0.283	0.292	0.291	0.323	0.332	0.370	0.437	0.079	2.72	0.018
EERRD	0.121	0.065	0.053	0.069	0.069	0.094	0.106	0.130	0.171	0.183	0.061	1.25	0.234
EERR	0.302	0.242	0.206	0.183	0.189	0.196	0.212	0.240	0.279	0.334	0.033	0.92	0.374
LEERD	0.052	0.030	0.030	0.045	0.039	0.056	0.051	0.076	0.110	0.134	0.081	3.71	0.003

Table XV: Panel B – Fama-MacBeth Annual Return Regressions based on Capital Expenditures Growth

Intercept	**CEG**	**BM**	**MV**	**EERRD**	**TASG**
0.060	-0.047	0.053			
(1.01)	(-2.96)[2]	(2.13)[1]			
0.083	-0.046	0.049	-0.005		
(0.93)	(-2.89)[2]	(1.67)	(-0.45)		
0.084	-0.012			-0.433	
(2.79)	(-0.89)			(-16.76)[3]	
0.037	-0.032				-0.151
(0.71)	(-2.11)[1]				(-5.01)[3]

Table XVI
Portfolios and regressions for Net Operating Assets

Panel A shows characteristics of portfolios sorted on Net Operating Assets. The portfolios are created annually according to the methodology in section 3.4.1 and the variable abbreviations are listed in Table I, section 3.3.3. The numbers in the table below are the portfolio averages for the 1994-2008 period. The upper variable is the sorting variable, on which the portfolios are sorted from low to high. The right side of the table reports the difference between the highest and lowest portfolio for each variable and the statistical significance of the resulting number. Panel B displays Fama-Macbeth regressions created in line with the methodology in section 3.4.3. Every second line represents a regression with the regression variables in the column headings and the statistical significance reported in brackets below the coefficients.

Table XVI: Panel A - Characteristics of portfolios sorted on Net Operating Assets. 1994-2008

	Low			Variable sorted portfolios				High			Significance		
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**H-L**	**T-stat**	**P**
NOA	-0.153	0.275	0.434	0.541	0.622	0.689	0.755	0.834	0.958	2.014	2.167	9.17	0.000
F6M	-0.042	-0.054	-0.046	-0.031	-0.020	-0.021	-0.030	-0.055	-0.069	-0.129	-0.087	-2.47	0.028
F12M	-0.012	-0.012	0.019	0.033	0.050	0.048	0.050	0.012	0.002	-0.097	-0.085	-1.58	0.138
F24M	-0.017	0.011	0.070	0.088	0.129	0.111	0.104	0.074	0.068	-0.096	-0.079	-1.00	0.334
F36M	0.018	0.063	0.136	0.175	0.224	0.203	0.199	0.151	0.145	-0.032	-0.050	-0.66	0.518
H12M	-0.004	0.003	0.025	0.054	0.066	0.063	0.042	0.060	0.040	0.036	0.040	0.70	0.495
H36M	-0.051	-0.047	0.093	0.144	0.181	0.236	0.204	0.293	0.379	0.549	0.600	6.80	0.000
MV	933	1,989	4,127	6,644	5,603	4,133	3,788	2,615	2,416	2,407	1.473	6.09	0.000
LIQ	0.191	0.275	0.161	0.153	0.137	0.133	0.127	0.134	0.150	0.190	-0.001	-0.04	0.970
BETA	1.327	1.192	1.121	1.045	0.949	0.823	0.811	0.832	0.883	0.971	-0.356	-3.99	0.002
VOL	0.212	0.177	0.150	0.130	0.122	0.113	0.116	0.121	0.128	0.148	-0.064	-6.15	0.000
ACC	-0.081	-0.044	-0.047	-0.040	-0.037	-0.031	-0.025	-0.014	-0.009	0.018	0.099	7.71	0.000
NSI	0.093	0.048	0.027	0.023	0.027	0.026	0.028	0.036	0.057	0.167	0.074	9.39	0.000
CEG	0.078	0.019	0.039	0.049	0.025	0.035	0.075	0.140	0.253	0.665	0.587	8.99	0.000
TASG	0.053	0.012	0.051	0.057	0.063	0.075	0.101	0.147	0.238	0.718	0.665	15.62	0.000
ROA	-0.254	-0.54	0.004	0.040	0.050	0.051	0.061	0.076	0.067	0.036	0.290	10.14	0.000
BM	0.390	0.493	0.522	0.543	0.582	0.627	0.662	0.604	0.598	0.483	0.093	3.65	0.003
EP	0.162	0.087	0.066	0.063	0.063	0.071	0.064	0.059	0.069	0.061	-0.101	-1.89	0.081
CP	0.100	0.116	0.170	0.131	0.140	0.146	0.157	0.133	0.140	0.111	0.011	1.43	0.178
SP	1.144	1.877	2.432	1.961	2.312	3.455	1.702	1.485	1.403	1.614	0.470	0.95	0.361
NEST	5.303	6.745	8.145	9.112	9.121	8.744	8.532	8.100	8.237	7.891	2.588	10.65	0.000
SDEST	0.558	0.364	0.305	0.239	0.230	0.140	0.182	0.158	0.184	0.208	-0.350	-7.25	0.000
EEPSG	0.415	0.339	0.303	0.290	0.260	0.261	0.263	0.291	0.320	0.386	-0.029	-1.01	0.332
EERRD	0.083	0.108	0.072	0.079	0.049	0.068	0.077	0.118	0.111	0.147	0.064	1.75	0.106
EERR	0.294	0.244	0.233	0.201	0.170	0.173	0.189	0.211	0.230	0.249	-0.045	-2.01	0.067
LEERD	0.060	0.036	0.038	0.049	0.021	0.042	0.040	0.078	0.060	0.094	0.034	1.42	0.182

Table XVI: Panel B – Fama-MacBeth Annual Return Regressions based on Net Operating Assets

Intercept	NOA	BM	MV	EERRD	TASG
0.111	-0.065	0.058			
(1.96)	(-2.43)[2]	(2.29)[2]			
0.105	-0.064	0.058	0.001		
(1.18)	(-2.53)[2]	(1.98)[1]	(0.09)		
0.090	-0.007			-0.420	
(2.67)	(-0.35)			(-15.11)[3]	
0.111	-0.065				0.058
(1.96)	(-2.43)[2]				(2.29)[2]

Table XVII
Portfolios and regressions for Net Stock Issues

Panel A shows characteristics of portfolios sorted on Net Stock Issues. The portfolios are created annually according to the methodology in section 3.4.1 and the variable abbreviations are listed in Table I, section 3.3.3. The numbers in the table below are the portfolio averages for the 1994-2008 period. The upper variable is the sorting variable, on which the portfolios are sorted from low to high. The right side of the table reports the difference between the highest and lowest portfolio for each variable and the statistical significance of the resulting number. Panel B displays Fama-Macbeth regressions created in line with the methodology in section 3.4.3. Every second line represents a regression with the regression variables in the column headings and the statistical significance reported in brackets below the coefficients.

Table XVII: Panel A - Characteristics of portfolios sorted on Net Stock Issues. 1994-2008

			Low	Variable sorted portfolios						High	Significance		
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**H-L**	**T-stat**	**P**
NSI	8	-0.014	-0.001	0.002	0.007	0.013	0.024	0.049	0.124	0.499	0.612	37.22	0.000
F6M	-0.011	-0.003	-0.038	-0.052	-0.039	-0.054	-0.056	-0.075	-0.087	-0.120	-0.109	-2.61	0.022
F12M	0.076	0.072	0.061	0.022	0.049	0.017	0.000	-0.033	-0.049	-0.107	-0.182	-3.36	0.005
F24M	0.147	0.152	0.160	0.104	0.130	0.093	0.053	-0.042	-0.068	-0.165	-0.312	-3.52	0.004
F36M	0.232	0.242	0.258	0.192	0.215	0.167	0.125	0.029	-0.055	-0.159	-0.390	-5.00	0.000
H12M	0.018	0.030	-0.001	-0.016	0.003	0.005	0.051	0.046	0.064	0.095	0.077	1.22	0.246
H36M	0.135	0.106	-0.001	-0.017	0.126	0.172	0.271	0.354	0.309	0.160	0.025	0.29	0.773
MV	5,823	7,371	3,937	2,736	3,654	2,314	1,839	2,093	1,841	1,906	-3.916	-3.48	0.004
LIQ	0.137	0.206	0.306	0.106	0.129	0.147	0.173	0.206	0.220	0.207	0.071	3.63	0.003
BETA	0.896	0.859	0.784	0.904	0.971	0.984	1.051	1.107	1.114	1.134	0.238	3.28	0.006
VOL	0.112	0.113	0.134	0.134	0.131	0.137	0.151	0.165	0.178	0.203	0.091	9.50	0.000
ACC	-0.046	-0.041	-0.031	-0.032	-0.031	-0.027	-0.037	-0.024	-0.018	-0.012	0.034	4.84	0.000
TASG	0.034	0.063	0.056	0.068	0.090	0.110	0.137	0.200	0.252	0.407	0.372	11.28	0.000
CEG	0.043	0.042	0.056	0.028	0.066	0.106	0.137	0.204	0.262	0.322	0.279	4.62	0.000
NOA	0.608	0.635	0.659	0.654	0.675	0.675	0.606	0.684	0.747	1.056	0.449	4.16	0.001
ROA	0.070	0.080	0.100	0.030	0.016	0.016	0.010	-0.006	-0.061	-0.266	-0.337	-4.08	0.001
BM	0.581	0.597	0.754	0.682	0.581	0.538	0.509	0.464	0.430	0.457	-0.124	-3.54	0.004
EP	0.078	0.066	0.089	0.087	0.063	0.059	0.055	0.057	0.051	0.112	0.034	1.05	0.312
CP	0.141	0.134	0.172	0.145	0.138	0.125	0.121	0.107	0.112	0.123	-0.017	-1.48	0.163
SP	1.653	1.782	4.295	2.062	1.970	1.478	1.439	1.307	2.004	1.257	-0.396	-2.47	0.028
NEST	9.690	9.518	7.274	7.373	7.980	7.649	7.302	7.277	7.117	6.172	-3.518	-5.27	0.000
SDEST	0.152	0.154	0.221	0.231	0.223	0.241	0.287	0.322	0.317	0.378	0.225	6.06	0.000
EEPSG	0.242	0.245	0.257	0.268	0.297	0.326	0.363	0.369	0.419	0.434	0.192	9.75	0.000
EERRD	0.054	0.063	0.063	0.095	0.074	0.091	0.094	0.127	0.160	0.142	0.088	4.24	0.001
EERR	0.168	0.170	0.178	0.215	0.215	0.211	0.224	0.252	0.271	0.263	0.096	5.72	0.000
LEERD	0.018	0.032	0.035	0.052	0.035	0.050	0.046	0.084	0.106	0.118	0.100	4.57	0.001

Table XVII: Panel B – Fama-MacBeth Annual Return Regressions based on Net Stock Issues

Intercept	**NSI**	**BM**	**MV**	**EERRD**	**TASG**
0.079	-0.207	0.051			
(1.35)	(-3.54)[2]	(2.32)[2]			
0.084	-0.213	0.041	-0.004		
(1.01)	(-3.77)[2]	(1.65)	(-0.45)		
0.087	-0.101			-0.423	
(2.95)	(-1.83)[1]			(-14.74)[3]	
0.050	-0.138				-0.146
(1.03)	(-2.24)[2]				(-3.93)[2]

From the portfolio analysis it can be seen that the hedge portfolio formed by taking a long position in the low portfolio and a short position in the high portfolio has generated significantly positive returns for asset growth, accruals, net stock issues and capital expenditures growth, while there seems to be some effect also for net operating assets and return on assets, yet not statistically significant.

The asset growth return spread is, both in terms of magnitude and statistic significance, the most pronounced among the factors covered in this thesis, and when run in regressions, it is significant when run together with all of the studied variables. Just as Cooper et al. (2008) observed through their research, the results displayed in Table XIV indicates that high asset growth is related to very high long-term past returns, while low asset growth firms have had very disappointing returns over the last 36 months. This is a common theme for most of the balance sheet related variables, making the effect of these variables clearly related to the 36 months stock return reversal effect. It is also clear that the balance sheet measures are very strongly related to each other, making it questionable whether all the variables really have an effect per se or if they merely are proxies for other variables. The very weak performance of the highest NOA portfolio, for example, is more than justified by the very high average asset growth for stocks in this portfolio. There are some differences in the stock return patterns for the different variables though.

While the asset growth, accruals and net operating assets effects are clearly driven by very poor performance for the highest portfolio, the capital expenditures effect; and especially the net stock issues effect; is less so, as some return effect seems to exist also for the medium and low portfolios. There are also differences in the durations of excess returns. While the effect of net operating assets and accruals only exist over the first 6-12 months after portfolio formation, asset growth and capital expenditures growth have an effect on returns over 24 months, and net stock issues and return on assets excess returns seem to persist over at least 36 months.

The hedge portfolio excess returns cannot, possibly with the exception for net operating assets, be explained by observed betas for the stocks in the portfolios as the beta difference between the high and low is either insignificant or in the “wrong” direction. There is however a very strong correspondence between analyst earnings estimation errors and the excess returns. Almost without exception does the deviation of a portfolio from the average return of the sample correspond nicely with a similar difference in the analyst over-optimism measure compared to the other portfolios. A good example of this is the asset growth effect, where the effect on returns is caused by increasingly worse returns for the highest portfolios while there is no major difference in returns for the medium and low portfolios. Exactly the same pattern emerges for analyst overestimations of earnings. Cooper et al. (2008) analyzed the stock returns following earnings announcements and came to the conclusion that the asset growth effect on stock returns was likely a result of mispricing because the returns after earnings releases where significantly lower for high than for low asset growth firms. This paper further enhances the evidence that the asset growth effect is due to mispricing, as it documents that analysts systematically overestimate future earnings for high asset growth stocks. However, only the earnings of the top asset growth portfolios seems to be systematically misestimated relative to the other portfolios, as the difference between estimated earnings and actual outcomes is basically the same for all portfolios except for the high ones. The same pattern can be seen in the returns.

When regressed together with the overoptimism measure, the significance of the balance sheet measures drops considerably. Yet, while net operating assets becomes very insignificant after adjusting for overoptimism; asset growth, net stock issues and accruals just makes it over the 5 % significance limit.

The results from Table XIV and XVII show that analysts both expected higher growth rates and systematically overestimated those growth rates for the historical high asset growth and net stock issues portfolios. This is consistent with the Lakonishok et al. (1994) hypothesis that investors tend to erroneously extrapolate historical trends into future estimates, resulting in overly optimistic views on high growth firms and overly pessimistic views on low growth firms. This hypothesis does not, however, provide an explanation for why there is only a strong effect for the highest asset growth portfolios and not for the medium to low ones.

There is a clear difference between asset growth and accruals when it comes to expectations about future earnings, which could possibly indicate different sources of the effects of these variables on stock returns. While asset growth is strongly related to higher expected future growth; accruals is not. The high historical growth of the high accruals portfolio is not translated into higher expectations about future earnings growth. Yet, the actual future earnings of these stocks are still lower than expected. Thus, it seems that the overoptimism and subsequent low returns of the high accruals portfolio is not caused by investors extrapolating historical growth trends. The evidence from the portfolio sorts is more supportive of the Sloan (1996) thesis that the bad performance of high accruals firms is caused by investors focusing too much on earnings, missing that the accruals component of earnings is less persistent than the cash flow component.

4.7 Additional Analyst Related Data

Table XVIII
Portfolios Sorted on Analyst on Number of Earnings Estimates.

Panel A shows characteristics of portfolios sorted on Number of Earnings Estimates. The portfolios are created annually according to the methodology in section 3.4.1 and the variable abbreviations are listed in Table I, section 3.3.3. The numbers in the table below are the portfolio averages for the 1994-2008 period. The upper variable is the sorting variable, on which the portfolios are sorted from low to high. The right side of the table reports the difference between the highest and lowest portfolio for each variable and the statistical significance of the resulting number. Panel B displays Fama-Macbeth regressions created in line with the methodology in section 3.4.3. Every second line represents a regression with the regression variables in the column headings and the statistical significance reported in brackets below the coefficients.

Table XVIII: Panel A - Characteristics of portfolios sorted on Number of Earnings Estimates. 1994-2008

	Low			**Variable sorted portfolios**				**High**			**Significance**		
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**H-L**	**T-stat**	**P**
NEST	1.005	1.724	2.588	3.517	4.636	6.018	7.842	10.43	14.42	23.07	22.07	37.26	0.000
F6M	-0.070	-0.061	-0.091	-0.086	-0.084	-0.071	-0.060	-0.047	-0.025	-0.022	0.049	1.49	0.159
F12M	0.007	0.014	-0.034	-0.050	-0.059	-0.041	-0.022	-0.008	0.015	0.025	0.019	0.50	0.627
F24M	0.072	0.066	-0.008	-0.044	-0.061	-0.016	-0.011	0.023	0.053	0.093	0.021	0.38	0.712
F36M	0.161	0.120	0.058	0.030	-0.015	0.036	0.035	0.076	0.114	0.177	0.016	0.20	0.848
H12M	0.003	-0.004	-0.003	0.012	0.041	0.034	0.037	0.046	0.054	0.058	0.055	1.47	0.166
H36M	0.020	0.055	0.109	0.128	0.272	0.313	0.325	0.341	0.347	0.370	0.350	5.26	0.000
MV	261	303	418	553	760	1,067	1,641	2,828	6,435	24163	23902	13.78	0.000
LIQ	0.190	0.122	0.754	0.302	0.173	0.348	0.202	0.206	0.214	0.230	0.041	0.52	0.613
BETA	0.921	0.945	0.967	1.027	0.991	1.031	1.056	1.049	1.017	1.000	0.079	1.56	0.143
VOL	0.165	0.158	0.150	0.152	0.145	0.138	0.131	0.120	0.105	0.098	-0.068	-18.83	0.000
ACC	-0.024	-0.030	-0.029	-0.027	-0.014	-0.025	-0.027	-0.028	-0.036	-0.041	-0.017	-2.74	0.017
TASG	0.083	0.116	0.151	0.195	0.193	0.206	0.203	0.187	0.162	0.151	0.068	2.59	0.023
NSI	0.064	0.064	0.056	0.069	0.061	0.055	0.053	0.043	0.033	0.025	-0.039	-3.45	0.004
CEG	0.077	0.099	0.143	0.205	0.209	0.242	0.225	0.164	0.166	0.158	0.081	2.03	0.064
NOA	0.592	0.649	0.675	0.714	0.734	0.704	0.732	0.758	0.798	0.754	0.162	3.36	0.005
ROA	-0.031	-0.011	-0.016	-0.014	0.011	0.011	0.031	0.052	0.061	0.067	0.099	9.76	0.000
BM	0.705	0.607	0.552	0.514	0.488	0.460	0.446	0.432	0.415	0.373	-0.332	-4.65	0.000
EP	0.084	0.066	0.065	0.067	0.056	0.056	0.054	0.052	0.053	0.051	-0.033	-4.21	0.001
CP	0.171	0.138	0.134	0.118	0.116	0.109	0.112	0.112	0.113	0.108	-0.063	-3.00	0.010
SP	2.169	1.681	1.563	1.560	1.311	1.192	1.259	1.202	1.095	0.808	-1.361	-7.88	0.000
SDEST	0.279	0.305	0.392	0.331	0.256	0.301	0.215	0.174	0.208	0.142	-0.135	-1.78	0.100
EEPSG	0.449	0.386	0.365	0.344	0.334	0.327	0.313	0.298	0.272	0.256	-0.193	-1.78	0.100
EERRD	0.207	0.145	0.119	0.127	0.117	0.122	0.098	0.081	0.067	0.035	-0.172	-3.00	0.011
EERR	0.371	0.299	0.280	0.262	0.251	0.226	0.211	0.192	0.166	0.152	-0.219	-3.07	0.010
LEERD	0.081	0.081	0.074	0.088	0.076	0.065	0.049	0.052	0.035	0.011	-0.071	-4.03	0.002

Table XVIII: Panel B – Fama-MacBeth Annual Return Regressions based on number of earnings estimates

Intercept	NEST	BM	MV	EERRD
0.028	0.003	0.057		
(0.46)	(1.78)[1]	(1.94)		
0.000	0.002	0.057	0.005	
(0.00)	(1.39)	(1.81)[1]	(0.42)	
0.093	-0.001			-0.433
(2.61)	(-0.66)			(-15.30)[3]

The analyst coverage is, not surprisingly, strongly related firm size. On average, the higher the market value, the more analysts are covering the company. It can be seen that higher analyst coverage greatly decreases the uncertainty about future earnings, as the standard deviation of earnings forecasts is clearly dependent on the number of issued forecasts. The same pattern is visible for average individual estimates errors, which also tend to decrease with the number of forecasts.

So far, the results have made evident that analyst's earnings forecasts are overoptimistic on average and that this overoptimism is a strong driver of stock returns causing part of the cross-sectional differences in returns of variable sorted portfolios. However, the relation between overoptimism and returns is not uncomplicated as the level of overoptimism is, as can be seen clearly in table XVIII, highly related to the analyst coverage. While the estimates for companies with low analyst coverage tend to be way too optimistic, the estimates for companies with high coverage are almost right on average. Unlike what is the case for most of the variable sorted portfolios, this overoptimism does not seem to be closely related to returns as the returns do not correspond to the number of estimates in a linear fashion.

There are two possible explanations for why the overoptimism in analyst coverage sorted portfolios is not translated into differences in returns. Firstly, the very low coverage portfolios might be subject to higher risk than the rest, leading to high risk premiums offsetting the negative returns following earnings disappointments. This risk is not visible by observing beta measures as stock betas are unrelated to analyst coverage. However, the lower coverage portfolios seem to be more risky in terms of volatility. The firms in these portfolios are also less profitable, have had lower historical firm growth and stock returns, and have higher book-to-market ratios as well as lower stock liquidity (at least in dollar terms). All of these factors might lead to risk premiums. However, it is entirely possible that only extreme values on these variables command risk premiums, thus losing their effect in the middle of the cross-section where instead the overoptimism measure takes the upper hand, leading to increasingly higher returns from the middle portfolios to the highest. This explanation would be consistent with the return patterns of the sorted portfolios in table XVIII.

Another possible explanation for the ineffectiveness of the overoptimsm measure on the stock returns of analyst coverage sorted portfolios is that investors do not base their small stock investment decisions on analyst forecasts. As previously stated, it is plausible that analysts on average overestimate the earnings of small stocks because there is no incentive for them to deliver negative forecasts. Knowing that, the rational investor should be critical to analyst buy recommendations for small companies, and if they are, the overoptimism measure will be a less effective predictor of stock returns. It is also possible that small company investors tend to base their investment decisions more on their personal feelings than on expected earnings numbers published in analyst reports.

Table XIX
Portfolios and regressions for Expected Earnings Growth

Panel A shows characteristics of portfolios sorted on Expected Earnings Growth. The portfolios are created annually according to the methodology in section 3.4.1 and the variable abbreviations are listed in Table I, section 3.3.3. The numbers in the table below are the portfolio averages for the 1994-2008 period. The upper variable is the sorting variable, on which the portfolios are sorted from low to high. The right side of the table reports the difference between the highest and lowest portfolio for each variable and the statistical significance of the resulting number. Panel B displays Fama-Macbeth regressions created in line with the methodology in section 3.4.3. Every second line represents a regression with the regression variables in the column headings and the statistical significance reported in brackets below the coefficients.

Table XIX: Panel A - Characteristics of portfolios sorted on Expected Earnings Growth. 1994-2008

	Low			Variable sorted portfolios				High			Significance		
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**H-L**	**T-stat**	**P**
EEPSG	-0.184	0.069	0.169	0.229	0.278	0.331	0.396	0.488	0.648	1.118	1.303	27.19	0.000
F6M	-0.030	-0.015	-0.011	-0.018	-0.013	-0.025	-0.044	-0.061	-0.071	-0.087	-0.056	-1.17	0.263
F12M	0.010	0.041	0.045	0.052	0.049	0.034	0.011	-0.013	-0.035	-0.072	-0.083	-1.70	0.113
F24M	0.037	0.084	0.124	0.125	0.099	0.080	0.045	0.013	-0.060	-0.074	-0.112	-1.59	0.136
F36M	0.129	0.184	0.207	0.219	0.164	0.138	0.103	0.055	-0.026	-0.050	-0.179	-2.14	0.052
H12M	-0.213	-0.051	0.004	0.052	0.083	0.119	0.156	0.233	0.268	0.270	0.483	8.02	0.000
H36M	0.195	0.228	0.300	0.300	0.383	0.444	0.554	0.654	0.682	0.499	0.303	4.57	0.001
MV	5,090	7,267	8,591	8,823	9,237	6,229	4,835	4,203	2,994	2,369	-2.721	-4.49	0.001
LIQ	0.180	0.139	0.132	0.136	0.139	0.166	0.442	0.209	0.224	0.225	0.045	3.01	0.010
BETA	0.935	0.753	0.805	0.872	0.919	0.932	0.956	1.079	1.155	1.142	0.206	2.82	0.014
VOL	0.118	0.092	0.091	0.090	0.096	0.105	0.116	0.128	0.141	0.150	0.032	4.72	0.000
ACC	-0.025	-0.030	-0.030	-0.024	-0.027	-0.017	-0.011	-0.015	-0.012	-0.018	0.007	1.05	0.314
TASG	0.155	0.124	0.124	0.128	0.154	0.173	0.220	0.243	0.294	0.268	0.113	3.52	0.004
NSI	0.024	0.018	0.011	0.012	0.015	0.025	0.043	0.061	0.071	0.084	0.060	4.67	0.000
CEG	0.202	0.172	0.141	0.120	0.159	0.196	0.238	0.269	0.300	0.299	0.097	1.65	0.123
NOA	0.717	0.728	0.729	0.710	0.728	0.742	0.788	0.800	0.812	0.800	0.083	2.48	0.027
ROA	0.103	0.105	0.095	0.099	0.110	0.102	0.098	0.098	0.089	0.062	-0.041	-4.85	0.000
BM	0.583	0.514	0.461	0.428	0.416	0.390	0.394	0.386	0.383	0.429	-0.154	-7.86	0.000
EP	0.085	0.071	0.065	0.059	0.056	0.052	0.048	0.042	0.036	0.025	-0.059	-21.71	0.000
CP	0.166	0.142	0.126	0.113	0.109	0.101	0.100	0.094	0.088	0.083	-0.082	-13.25	0.000
SP	1.368	1.313	1.401	1.277	1.202	1.157	1.152	1.082	0.991	1.016	-0.352	-3.94	0.002
NEST	10.57	10.84	10.67	10.55	10.71	10.17	9.344	8.877	8.342	7.924	-2.651	-3.58	0.003
SDEST	0.340	0.079	0.062	0.051	0.053	0.054	0.060	0.072	0.095	0.152	-0.188	-3.13	0.008
EERRD	0.164	0.099	0.084	0.064	0.059	0.060	0.073	0.076	0.099	0.196	0.032	0.87	0.401
EERR	0.321	0.185	0.160	0.139	0.154	0.159	0.198	0.243	0.294	0.381	0.060	1.36	0.198
LEERD	0.040	0.038	0.039	0.030	0.037	0.042	0.056	0.050	0.087	0.147	0.107	2.59	0.024

Table XIX: Panel B – Fama-MacBeth Annual Return Regressions based on Expected Earnings Growth

Intercept	**EEPSG**	**BM**	**MV**	**EERRD**	**TASG**
0.095	-0.034	0.036			
(1.98)	(-1.59)	(1.35)			
-0.019	-0.038	0.038	0.012		
(-0.28)	(-1.63)	(1.34)	(1.73)		
0.108	-0.062			-0.438	
(3.24)	(-2.14)[1]			(-14.72)[3]	
0.065	-0.049				-0.143
(1.57)	(-2.10)[1]				(-3.72)[2]

Although not highly significant, expected growth in earnings is negatively related to future returns. Expected growth is strongly related to the market multiples, which is far from surprising because the multiples are commonly used as indicators of the market's perception of the growth opportunities of companies. While CP and EP are linearly related to expected earnings growth, BM is not. In fact, if the lowest growth expectations portfolios are excluded, there is actually a slight increase in BM ratios with increasing growth expectations. Although this may seem puzzling, the lower profitability for the highest growth firms might be the explanation.

Analysts tend to overestimate the earnings of the high expected growth firms, but also, to a smaller degree; the earnings of the lowest growth firms. The portfolios in the middle have the lowest levels of overoptimism as well as the lowest estimate errors in general. The overoptimism pattern seems to be well reflected in the returns, as the middle portfolios have the highest average returns over a 6-24 month period.

The Fama-MacBeth regressions between expected earnings growth and other variables show interesting results. While the expected earnings growth measure loses significance when combined with BM, it does not seem to lose any predictive ability when combined with asset growth, and more puzzlingly; overoptimism. Thus, even though there is a tendency for higher earnings growth portfolios to contain stocks with overestimated future earnings, this does not seem to be the driver for the lower returns for these stocks.

Table XX
Portfolios and regressions for Expected Earnings Dispersion

Panel A shows characteristics of portfolios sorted on Expected Earnings Dispersion. The portfolios are created annually according to the methodology in section 3.4.1 and the variable abbreviations are listed in Table I, section 3.3.3. The numbers in the table below are the portfolio averages for the 1994-2008 period. The upper variable is the sorting variable, on which the portfolios are sorted from low to high. The right side of the table reports the difference between the highest and lowest portfolio for each variable and the statistical significance of the resulting number. Panel B displays Fama-Macbeth regressions created in line with the methodology in section 3.4.3. Every second line represents a regression with the regression variables in the column headings and the statistical significance reported in brackets below the coefficients

Table XX: Panel A - Characteristics of portfolios sorted on Expected Earnings Dispersion. 1994-2008

		Low			Variable sorted portfolios				High			Significance	
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**H-L**	**T-stat**	**P**
SDEST	0.020	0.024	0.033	0.044	0.058	0.081	0.115	0.172	0.284	1.708	1.687	11.27	0.000
F6M	-0.015	-0.025	-0.036	-0.046	-0.073	-0.058	-0.077	-0.084	-0.099	-0.128	-0.113	-1.81	0.094
F12M	0.012	0.023	0.006	-0.002	-0.022	-0.016	-0.052	-0.043	-0.069	-0.115	-0.126	-1.52	0.152
F24M	0.041	0.059	0.045	0.020	-0.011	0.018	-0.064	-0.006	-0.062	-0.103	-0.144	-1.22	0.243
F36M	0.092	0.109	0.093	0.101	0.043	0.081	-0.013	0.048	0.004	-0.055	-0.147	-1.47	0.166
H12M	0.151	0.130	0.124	0.121	0.091	0.060	0.024	0.001	-0.101	-0.203	-0.353	-3.61	0.003
H36M	0.480	0.463	0.466	0.421	0.418	0.369	0.247	0.219	-0.017	-0.114	-0.594	-5.86	0.000
MV	9,995	8,670	6,064	4,300	3,722	4,491	3,530	3,264	1,785	934	-9.061	-9.87	0.000
LIQ	0.141	0.566	0.161	0.174	0.179	0.198	0.193	0.202	0.199	0.199	0.058	4.90	0.000
BETA	0.739	0.798	0.865	0.928	1.029	1.133	1.171	1.168	1.223	1.365	0.626	5.44	0.000
VOL	0.092	0.096	0.102	0.112	0.120	0.131	0.140	0.152	0.164	0.190	0.098	9.54	0.000
ACC	-0.021	-0.022	-0.023	-0.023	-0.021	-0.027	-0.031	-0.035	-0.043	-0.052	-0.031	-6.09	0.000
TASG	0.168	0.167	0.178	0.188	0.189	0.194	0.205	0.188	0.174	0.175	0.008	0.16	0.873
NSI	0.032	0.028	0.038	0.039	0.038	0.045	0.050	0.066	0.069	0.097	0.064	6.76	0.000
CEG	0.211	0.169	0.199	0.192	0.231	0.196	0.204	0.202	0.186	0.132	-0.079	-0.90	0.387
NOA	0.762	0.760	0.758	0.783	0.756	0.767	0.704	0.689	0.698	0.674	-0.088	-0.76	0.462
ROA	0.099	0.103	0.094	0.078	0.065	0.047	0.015	-0.016	-0.069	-0.145	-0.244	-16.77	0.000
BM	0.376	0.390	0.420	0.433	0.433	0.460	0.491	0.508	0.516	0.529	0.153	4.14	0.000
EP	0.054	0.054	0.055	0.054	0.053	0.058	0.054	0.058	0.070	0.057	0.003	0.52	0.614
CP	0.101	0.105	0.106	0.110	0.109	0.111	0.118	0.125	0.138	0.135	0.034	5.55	0.000
SP	1.033	1.082	1.181	1.230	1.208	1.321	1.273	1.376	1.295	1.389	0.357	1.23	0.240
NEST	9.482	10.62	10.22	9.761	9.028	8.700	8.367	8.882	8.084	6.535	-2.947	-11.45	0.000
EEPSG	0.278	0.277	0.286	0.302	0.327	0.335	0.355	0.369	0.303	-0.248	-0.525	-2.59	0.022
EERRD	0.045	0.042	0.060	0.069	0.106	0.145	0.171	0.189	0.163	0.303	0.258	3.26	0.007
EERR	0.107	0.128	0.148	0.186	0.216	0.272	0.320	0.385	0.471	0.878	0.771	7.98	0.000
LEERD	0.030	0.025	0.027	0.043	0.070	0.087	0.099	0.085	0.034	-0.056	-0.085	-0.23	0.820

Table XX: Panel B – Fama-MacBeth Annual Return Regressions based on Expected Earnings Dispersion

Intercept	**SDEST**	**BM**	**MV**	**NSI**	**ACC**	**CP**
0.049	-0.064	0.051				
(0.83)	(-0.78)	(1.59)				
-0.078	-0.032	0.058	0.019			
(-0.99)	(-0.42)	(1.68)	(2.21)[1]			
0.012	-0.070			-0.281		
(0.27)	(-0.83)			(-.97)		
-0.011	-0.075				-0.336	
(-0.25)	(-0.85)				(-2.35)[2]	
0.160	0.008					0.062
(1.75)	(0.11)					(2.11)[1]

The portfolios with high dispersions in analyst forecasts are characterized by small to medium sized companies with weak historical returns, lower-than-average analyst following, high volatilities and betas, low profitability and bad expected future prospects implied by high book-to-market ratios and low expected earnings growth (for the highest dispersion portfolio). These stocks are very likely to be mispriced and despite low investor expectations, analysts do not seem to be pessimistic enough since future earnings estimates are inflated well above the actual future outcomes. Although not strongly significant, there seems to be a negative relationship between the dispersion in earnings forecasts and future returns, in line with (Diether, Malloy, & Scherbina, 2002), but this is fully attributable to the first 6 months after portfolio formation. That the effect of earnings dispersion is only short term is consistent with the values of the earnings overestimation variables. While the companies with high dispersion tend to have overestimated earnings forecasts (EERRD), this effect disappears within the 6 months subsequent to portfolio formation (LEERD). Thus, it seems that the analysts realize that the earnings estimates are two high, revise them down to more realistic levels leading to low returns for the downgraded companies. After the downgrade, the high dispersion portfolios seem to have approximately correct earnings estimates, resulting in no extra dispersion effect on returns.

4.8 Estimate Overoptimism and the Predictability of Stock Returns

The results from this thesis clearly indicate that analysts tend to be overly optimistic on average, in particular in companies where the analyst following is low. It is also clear that analyst earnings misestimations have been a leading driver of American stock prices during the 1994-2009 time period. Portfolios formed on the degree of realized misestimations lead to high excess returns for those companies that performed better than expected and low returns for the worse performers. This observation in itself is not at all controversial, and fully consistent with the idea of efficient capital markets because the investor cannot form portfolios on the misestimation variable until after the returns have already occurred. The intriguing part is that the degree of overoptimism seems to be predictable to a considerable extent. Having examined a large number of variables for which previous research have shown predictive abilities on future return, it is clear that a substantial part of their effect comes, for many of them, from the ability to predict earnings surprises. There is considerable disagreement among researchers whether the variables included in this thesis predict returns because they are proxies for risk exposure or if the effects are created by behavioral bias resulting in mispricing. The evidence that the variables can predict erroneous earnings estimates, clearly gives support to the latter camp.

There are considerable differences among the studied variables with regards to relations to the misestimate measure. Even though there is no doubt that most variables that predict returns are related to earnings misestimates, the results are inconclusive for a few variables. The imprecise nature of the misestimate measure makes it impossible to say exactly how much of the predictability in returns that comes from exposure to overoptimism. Thus, ranking the variables based on correlation with overoptimistic perceptions about the future is an incredibly imprecise exercise, but yet meaningful due to the important implications the results lead to in terms of possible reasons behind the predictive abilities of the variables.

XXIII contains information about the degree of correspondence with the overoptimism measure for the variables that exhibited predictive abilities on returns. The table contains two parts; the first one is based on hedge returns of portfolios sorted on each of the variables, the second one contains pairwise correlations between each of the variables and earnings overoptimism and regression data based on univariate regressions with overoptimism as the dependent variable and each of the variables as independent variable.

The portfolio side of the table contains three parts. The first one displays hedge returns and t-statistics for portfolios sorted on each of the variables. This is done by using the same method as previously, but the results differ because the stocks now need to have values for the overoptimism measure to be included. The middle part of the table shows how much of the hedge return could be explained by the overoptimism characteristics of the portfolios. To obtain this number, a regression is first run explaining the effect of overoptimism on 12 months future stock returns. Then, the overoptimism measure of each portfolio year is entered into the regression, resulting in an expected return based on the average earnings misestimation impact on returns. This expected return is then used to create the hedge portfolio. In the third part of the table, the hedge return expected from the overoptimism regression (middle part) is deducted from the initial hedge return (first part); to get the hedge return adjusted for earnings estimation errors.

Table XXIII
The impact of overoptimistic analyst earnings estimates

Table XXIII: Panel A – Portfolio hedge returns and regressions

	Portfolios						Regressions and correlations		
	Before adjustment		Overoptimism		After adjustment		Correlation	Regression	
	H-L	T-stat	H-L	T-stat	H-L	T-stat	Correlation	T-stat	R^2
LIQ	-0.147	-1.90	-0.050	-2.75	-0.107	-1.50	0.079	7.77	0.006
BETA	-0.037	-0.77	-0.061	-2.75	0.006	0.14	0.058	5.62	0.003
ACC	-0.084	-2.42	-0.032	-2.51	-0.057	-1.70	0.055	4.28	0.003
TASG	-0.121	-2.59	-0.051	-4.31	-0.079	-1.91	0.089	9.75	0.008
NSI	-0.101	-2.70	-0.040	-5.30	-0.068	-1.95	0.052	5.71	0.003
CEG	-0.068	-1.60	-0.031	-1.92	-0.041	-1.29	0.072	7.14	0.005
BM	0.128	1.63	-0.009	-0.82	0.146	1.90	-0.011	-1.19	0.000
EP	0.137	1.79	0.036	1.98	0.085	1.20	-0.06	-6.58	0.004
CP	0.197	2.14	0.043	2.73	0.172	1.83	-0.079	-8.84	0.006
SP	0.133	1.33	0.011	0.65	0.132	1.38	-0.027	-3.00	0.001
SDEST	-0.010	-0.27	-0.072	-2.49	0.071	2.62	0.154	17.2	0.024

The results from table XXIII are largely consistent with the results from what was indicated in the earlier portfolio sorts and regressions. Thus, it is now clear that both the second and the third hypothesis of this thesis; that the studied variables can predict stock returns, and that the effect of analyst misestimations of future earnings is related to this return predictive effect in a systematic way; are true. Misestimated earnings can explain a considerable part of the asset growth, capital expenditures growth, liquidity and earnings-to-price as more than 25 % of the hedge portfolio returns disappear after the adjustments. They also have highly significant t-statistics when ran in regressions with earnings overoptimism, and their correlations with this measure are all in the upper half when compared to the other variables. This provides support for mispricing explanations for those effects. The hedge returns of net stock issues and accruals are also clearly changed after the adjustment, but they are less correlated in the regressions. Conversely, cash flow-to-price is among the strongest correlated factors in the regressions, but the adjustment effect on the hedge portfolio is weaker. The beta effect on return changes direction after the adjustment. This, taken together with evidence from the portfolio sorts that high beta portfolios have variable characteristics that predicts low returns, suggest that the “pure beta effect” on returns might be positive; as the CAPM states; but that beta is positively correlated with overpricing characteristics caused by exogenous factors, creating lower returns because of subsequent earnings disappointments.

Book-to-market, sales-to-price and expected earnings growth are the variables least correlated to earnings misestimations and the impact on their hedge returns of the return adjustment is also miniscule. For book-to-market and sales-to-price, these results are in line with the results from the portfolio analysis which also did not show any indications of relation between these variables and overoptimism. While the significance of these variables decreased the slightly when including overoptimism in Fama-MacBeth regressions, expected earnings growth was still as significant. The portfolio sorts for earnings growth showed however that there might be a relation between the level of overoptimism and returns for the portfolios, although not visible in the hedge portfolios. Despite some slightly conflicting results, the overall impression is that there is no strong relation between these three variables and earnings overoptimism.

It is easy to argue that the book-to-market, sales-to-price and earnings growth variables impact on returns all are the same effect since all three can be seen as measures of the value/growth characteristics of companies. It is puzzling though that cash flow-to price and earnings-to-price, also proxies for the value/growth effect, clearly seem to be related to overoptimism. Although not proving that it is the case, the fact that the misestimate test failed, points at risk related explanations for the book-to-market effect, which further provides support for the FF3M model as an adjustment model in asset pricing research.

5. Conclusion

The empirical results indicate that asset growth and other balance sheet related factors do generally exhibit the strongest relations to returns in terms of statistical significance, but while their effect is primarily driven by the extreme portfolios, the market multiples are related to returns throughout the cross-section in a more linear fashion. Liquidity seems to have a much stronger economic significance on future returns in this paper than what has been observed in previous research, and the return spread between the most and least liquid stocks is much larger than what could possibly be explained by liquidity risk premiums. Rather, it seems like the liquidity measure can help locate overhyped and forgotten stocks as the most liquid stocks seems to have highly glamorous characteristics but low future returns and the least liquid stocks boring characteristics but high future returns.

By providing a comprehensive empirical review of the characteristics of stock return predicting variables in terms of their relation to returns, their relation to each other and their relation to analyst misestimations of future earnings; several interesting observations have been made. This thesis firstly makes clear that stock returns are to a considerable extent driven by corrections to erroneous forecasts about the future of individual firms. This is not controversial in itself as it is consistent with an efficient market. What is more intriguing, and also the paramount finding of this thesis; is that the analyst earnings estimate errors are predictable, not only in magnitude but also in direction, and that these errors correspond nicely with the stock return predictive abilities of most of the studied variables. Although not ruling out that part of the relations between the studied factors and stock returns is related to risk, the results provides strong support for the view that stock returns are predictable because of systematic mispricing, which can be detected by observing values on some of the variables studied in this text.

Although confirming that most of the covered variables predicting stock returns are strongly related to analyst overoptimism patterns, a few similar value/growth related variables, most notably the book-to-market ratio, do not seem to be related to estimate errors in a systematic way. While this finding does not prove that the excess returns predicted by book-to-market are not due to mispricing, it provides some support for the view that the ratio is a proxy for risk, thus supporting its inclusion in the FF3M model even though other variables are stronger predictors of stock returns. To include the size factor in the model is not supported by the data in the 1994-2008 period however, as market value has basically no explanatory power on returns when combined with book-to-market.

Another interesting finding made possible by combining analyst overoptimism and variable characteristics data is that the short term stock return momentum effect does not exist after adjusting for overoptimistic earnings forecasts. When regressed together with this measure, the return effect turns negative, resulting in a pattern that does not seem to be different to the longer term reversal effect. Thus, the short-term momentum effect seems to exist solely because of too high expectations about the future, and when this is adjusted for, high historical returns leads to lower future returns, even in the shorter term. Further evidence show that the momentum effect (unadjusted for overoptimism) decreases with higher analyst coverage, indicating that the effect is dependent on the speed of information flow, in line with the findings of Hong and Stein. While providing out of sample support for the Hong and Stein results, this paper also confirms that the effect is actually due to lags in earnings estimates revisions, and not due to risk exposure or correlations with other factors.

Beta is negatively related to future stock returns, contrary to what is predicted by the CAPM. However, just as in the case of the momentum effect, this seems to be solely attributable to investors being overly optimistic about high beta firms and overly pessimistic about low beta firms. When the beta effect on returns is adjusted for overoptimistic earnings forecasts, the relation with returns is virtually zero. Although not turning significantly positive after the adjustment, this has important implications since it shows that the limited beta effect on returns observed in this paper as well as in previous research may be caused by mispricing generated noise. While the overoptimism measure used in this thesis is quite unsophisticated as it only takes into account the effect of errors in two years ahead earnings forecasts, it is possible that beta could be a significantly positive predictor of returns if adjusted using a more effective misestimate measure.

The implications of the predictability of future earnings surprises are important for investors as they, if they can predict these surprises, can use the knowledge to generate excess returns.

6. Bibliography

Anderson, C. W., & Garcia-Feijóo, L. (2006). Empirical Evidence on Capital Investment, Growth Options, and Security Returns. *Journal of Finance 61* , 171-194.

Ang, A., Hodrick, R. J., Xing, Y., & Zhang, X. (2006). The Cross-Section of Volatility and Expected Returns. *Journal of Finance 61* , 259-299.

Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics 9* , 3-18.

Barbee, W. C., Jeong, J.-G., & Mukherji, S. (2008). Relations between portfolio returns and market multiples. *Global Finance Journal 19* , 1-10.

Barberis, N., Shleifer, A., & Vishny, R. (1998). A model of investor sentiment. *Journal of Financial Economics 49* , 307-343.

Basu, S. (1983). The relationship between earnings yield, market value, and return for NYSE common stocks. *Journal of Financial Economics 12* , 129-156.

Campbell, J. Y., & Vuolteenaho, T. (2004). Bad Beta, Good Beta. *The American Economic Review 94* , 1249-1275.

Carhart, M. M. (1997). On Persistence in Mutual Fund Performance. *Journal of Finance 52* , 57-82.

Cooper, M. J., Gulen, H., & Schill, M. J. (2008). Asset Growth and the Cross-Section of Stock Returns. *Journal of Finance 63* , 1609-1651.

Daniel, K., & Titman, S. (1997). Evidence on the Characteristics of Cross Sectional Variation in Stock Returns. *Journal of Finance 52* , 1-33.

Daniel, K., Hirshleifer, D., & Subrahmanyam, A. (1998). Investor Psychology and Security Market Under- and Overreactions. *Journal of Finance 53* , 1839-1885.

Datar, V. T., Naik, N. Y., & Radcliffe, R. (1998). Liquidity and stock returns: An alternative test. *Journal of Financial Markets 1* , 203-219.

De Bondt, W. F., & Thaler, R. H. (1987). Futher Evidence on Investment Overreaction and Stock Market Seasonality. *Journal of Finance 62* , 557-581.

Diether, K. B., Malloy, C. J., & Scherbina, A. (2002). Differences of Opinion and the Cross Section of Stock Returns. *Journal of Finance 57* , 2113-2141.

Easterwood, J. C., & Nutt, S. R. (1999). Inefficiency in Analysts' Earnings Forecasts: Systematic Misreaction or Systematic Optimism? *Journal of Finance 54* , 1777-1797.

Fairfield, P., Scott, W., & Yohn, T. (2003). Accrued Earnings and Growth: Implications for Future Profitability and Market Mispricing. *The Accounting Review 78* , 353-371.

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

Fama, E. F., & French, K. R. (2008). Dissecting Anomalies. *Journal of Finance 63* , 1653-1678.

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

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

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

Haugen, R. A., & Baker, N. L. (2010). Case Closed. In J. B. Guerard, *Handbook of Portfolio Construction: Contemporary Applications of Markowitz Techniques* (pp. 601-620). Springer.

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

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

Hong, H., Lim, T., & Stein, J. C. (2000). Bad News Travels Slowly: Size, Analyst Coverage, and the Profitability of Momentum Strategies. *Journal of Finance 55* , 265-295.

Huang, A. G. (2009). The cross section of cashflow volatility and expected stock returns. *Journal of Empirical Finance 16* , 409-429.

Ikenberry, D., Lakonishok, J., & Vermaelen, T. (1995). Market underreaction to open market share repurchases. *Journal of Financial Economics 39* , 181-208.

Ince, O. S., & Porter, B. R. (2006). Individual Equity Return Data From Thomson Datastream: Handle With Care! *The Journal of Financial Research 31* , 463-479.

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance 48* , 65-92.

La Porta, R. (1996). Expectations and the Cross-Section of Stock Returns. *Journal of Finance 51* , 1715-1742.

La Porta, R., Lakonishok, J., Shleifer, A., & Vishny, R. (1997). Good News for Value Stocks: Further Evidence on Market Efficiency. *Journal of Finance 52* , 859-874.

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

Lara, J. M., Osma, B. G., & Noguer, B. G. (2006). Effects of Database Choice on International Accounting Research. *Abacus 42* , 426-454.

Lintner, J. (1965). Security Prices, Risk, and Maximal Gains From Diversification. *Journal of Finance 20* , 587-615.

Loughran, T., & Ritter, J. R. (1995). The New Issues Puzzle. *Journal of Finance 50* , 23-51.

Nagel, S. (2005). Short sales, institutional investors and the cross-section of stock returns. *Journal of Financial Economics 78* , 277-309.

Pontiff, J., & Woodgate, A. (2008). Share Issuance and Cross-sectional Returns. *Journal of Finance 63* , 921-945.

Rosenberg, B., Reid, K., & Lanstein, R. (1985). Persuasive evidence of market inefficiency. *Journal of Portfolio Management 11* , 9-17.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of Market Equilibrium under Conditions of Risk. *Journal of Finance 19* , 425-442.

Sloan, R. G. (1996). Do stock prices fully reflect information in accruals and cash flows about Future Earnings? *The Accounting Review 71* , 289-315.

Subrahmanyam, A. (2010). The Cross-Section of Expected Stock Returns: What Have We Learnt from the Past Twenty-Five Years of Research? *European Financial Management 16* , 27-42.

Titman, S., Wei, J., & Xie, F. (2004). Capital Investments and Stock Returns. *Journal of Financial and Quantitative Analysis 39* , 677-700.

Appendices

Appendix A
12 months returns of variable sorted portfolios

Panel A shows characteristics of portfolios sorted on each of the variables below. The portfolios are created annually according to the methodology in section 3.4.1 and the variable abbreviations are listed in Table I, section 3.3.3. The numbers in the table below are the portfolio averages for the 1994-2008 period. All variables below are sorted from low to high; with the returns displayed in each cell. The right side of the table reports the difference between the highest and lowest portfolio for each variable and the statistical significance of the resulting number.

12 months returns of variable sorted portfolios. 1994-2008

	Low			Variable sorted portfolios				High			Significance	
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**H-L**	**T-stat**	**P**
H36M	0.092	0.014	0.051	0.027	0.045	0.018	0.028	-0.003	-0.009	-0.085	-0.177	-3.08	0.009
MV	0.092	0.024	0.020	-0.022	-0.055	-0.041	-0.028	-0.019	0.002	0.015	-0.078	-1.34	0.203
LIQ	0.096	0.056	0.043	0.031	0.010	0.009	-0.033	-0.069	-0.107	-0.149	-0.244	-2.92	0.012
BETA	0.030	0.065	0.060	0.040	0.051	0.051	0.039	0.036	0.013	-0.044	-0.074	-1.44	0.174
VOL	0.075	0.060	0.057	0.030	0.016	0.013	0.001	-0.030	-0.038	-0.040	-0.115	-1.37	0.194
ACC	-0.001	0.060	0.032	0.058	0.022	0.009	0.007	0.011	-0.027	-0.065	-0.064	-3.27	0.006
TASG	0.066	0.057	0.052	0.048	0.043	0.036	0.021	-0.006	-0.044	-0.175	-0.242	-3.38	0.005
NSI	0.076	0.072	0.061	0.022	0.049	0.017	0.000	-0.033	-0.049	-0.107	-0.182	-3.36	0.005
CEG	0.052	0.053	0.037	0.019	0.028	0.014	0.017	-0.009	-0.062	-0.132	-0.184	-2.82	0.014
NOA	-0.012	-0.012	0.019	0.033	0.050	0.048	0.050	0.012	0.002	-0.097	-0.085	-1.58	0.138
ROA	-0.077	-0.065	0.000	0.022	0.037	0.031	0.043	0.039	0.036	0.019	0.096	1.48	0.163
BM	-0.087	-0.059	-0.033	-0.011	-0.002	0.026	0.039	0.043	0.065	0.106	0.194	2.07	0.059
EP	-0.067	-0.038	0.025	0.015	0.040	0.024	0.062	0.061	0.062	0.080	0.146	2.5	0.027
CP	-0.102	-0.025	-0.005	0.015	0.026	0.045	0.062	0.071	0.077	0.117	0.220	2.39	0.033
SP	-0.148	-0.084	-0.009	0.016	0.022	0.030	0.041	0.074	0.061	0.102	0.250	1.99	0.068
NEST	0.007	0.014	-0.034	-0.050	-0.059	-0.041	-0.022	-0.008	0.015	0.025	0.019	0.5	0.625
SDEST	0.012	0.023	0.006	-0.002	-0.022	-0.016	-0.052	-0.043	-0.069	-0.115	-0.126	-1.52	0.152
EEPSG	0.010	0.041	0.045	0.052	0.049	0.034	0.011	-0.013	-0.035	-0.072	-0.083	-1.7	0.113
EERRD	0.341	0.193	0.129	0.075	0.012	-0.048	-0.099	-0.177	-0.267	-0.447	-0.788	-10.55	0.000
EERR	0.052	0.054	0.047	0.045	0.034	-0.002	-0.028	-0.088	-0.162	-0.241	-0.293	-3.68	0.003
LEERD	0.347	0.252	0.185	0.124	0.071	0.007	-0.052	-0.105	-0.206	-0.373	-0.720	-10.14	0.000

Appendix B
Fama-MacBeth Regressions – 1 Year Stock Returns, 1994-2008

Average slopes, intercepts and their t-statistics from annual univariate regressions with the variables below as independent variables and one year future logged return as the dependent variable. Adj. R^2 is the average adjusted R^2 for the annual regressions and obs. the number of average annual observations. For a detailed explanation on variable definitions and regression methodology refer to section 3.4.3.

	Regression slope					Intercept		
Factor	t-stat	Coef.	Adj. R^2	Prob	Correct	t-stat	Coef.	Obs
TASG	-4.5	-0.170	0.022	0.001	15 (13)	0.82	0.041	1,882
ACC	-3.42	-0.243	0.003	0.005	14 (5)	0.07	0.004	958
NSI	-3.28	-0.240	0.009	0.006	12 (8)	0.51	0.025	1,969
CEG	-2.81	-0.061	0.015	0.015	13 (7)	0.25	0.013	1,591
LIQ	-2.66	-0.381	0.025	0.020	13 (10)	0.96	0.050	1,646
H36M	-2.59	-0.048	0.015	0.022	12 (9)	0.38	0.019	1,236
EP	2.56	0.049	0.024	0.024	12 (7)	2.42	0.174	1,396
CP	2.31	0.059	0.046	0.038	10 (9)	2.06	0.168	1,601
NOA	-2.2	-0.055	0.009	0.046	12 (7)	0.89	0.049	1,867
BM	2.17	0.056	0.028	0.049	9 (9)	1.06	0.059	1,869
EEPSG	-2.09	-0.047	0.010	0.057	10 (7)	0.76	0.033	1,289
SP	2.02	0.045	0.055	0.064	8 (8)	0.58	0.030	1,909
H6M	1.79	0.078	0.014	0.097	10 (10)	-0.12	-0.006	2,043
H24M	-1.62	-0.038	0.015	0.129	10 (6)	0.32	0.016	1,784
VOL	-1.51	-0.339	0.025	0.155	8 (7)	1.38	0.067	1,929
ROA	1.43	0.001	0.017	0.176	10 (7)	0.18	0.009	1,89
NEST	1.31	0.002	0.007	0.213	8 (7)	-0.56	-0.033	1,712
BETA	-1.25	-0.030	0.012	0.233	11 (5)	1.28	0.061	1,374
MV	-1.08	-0.009	0.012	0.300	10 (7)	0.61	0.050	2,122
SDEST	-1.02	-0.090	0.015	0.326	8 (8)	-0.27	-0.013	1,384
ROE	0.93	0.000	0.009	0.369	8 (5)	0.11	0.006	1,905
H3M	0.9	0.061	0.013	0.384	10 (7)	0.02	0.000	2,078
H1M	-0.23	-0.027	0.013	0.822	7 (3)	-0.09	-0.005	2,088
H12M	0	0.000	0.000	1.000	9 (6)	0.15	0.008	1,947