Style investing

Wouters, T.

Publication date:
2006

Chapter 2

Review of the literature

2.1 Introduction

Academic researchers in finance are in the middle of a debate about the way people make decisions and how this should be modeled. In general, people make observations, process data and make judgments and decisions. The judgments and decisions have implications for individual portfolio compositions, the range of securities offered in the market, the character of earnings forecasts and the way in which securities are priced. In building models to study financial markets, assumptions have to be made about the decision making process of investors. In neo-classical finance, decision makers possess Von Neumann-Morgenstern preferences and use Bayesian techniques to make appropriate statistical judgments. Researchers working in the area of behavioral finance have produced evidence that people deviate from the Von Neumann-Morgenstern rationality and Bayesian rules. To reduce the amount of time and effort that is needed for the complex requirements of the decision-making process, people use rules of thumb or heuristics to simplify the decision-making process. However, relying on heuristics may result in biased decisions ignoring relevant information or processing irrelevant information. In this chapter, we discuss different frameworks from the rational and behavioral school. The aim of this chapter is to give a review of the relevant literature.

2.2 Market efficiency and anomalies

Fama (1970) defines an efficient market as one in which security prices reflect all available information. To make this definition empirically testable, efficient markets are specified in more detail in the efficient market hypothesis. This hypothesis states that none of the trading systems based on only currently available information can earn excess risk-adjusted average returns. If the efficient market hypothesis holds, investors cannot consistently beat the market. This implies that an average investor is better off following a passive strategy where he or she holds the market portfolio, instead of following an active strategy where he or she is wasting money and time by analyzing, picking and trading the 'right' securities.

Anomalies have been found over the last two decades that challenge the traditional view that securities are rationally priced and that the prices reflect all publicly available information. In the following paragraphs we outline some of the more salient findings in the literature. We first discuss two anomalies that challenge the weak form efficient market hypothesis. The weak-form efficient market hypothesis states that it is impossible to earn superior returns based on the knowledge of past returns and prices. De Bondt and Thaler (1985) give evidence of long-term return reversals. They divide stocks into losers and winners based on three year past returns. The five year post-formation returns show that losers outperform winners, which cannot be explained with any rational pricing model and is clearly at odds with weak-form efficiency. De Bondt and Thaler explain this by investors overreacting to past information. The second anomaly is the momentum effect. Jegadeesh and Titman (1993) show evidence of short-term trends. Investors, who follow strategies that buy stocks with a good performance and sell stocks with a bad performance over the last three to twelve months, will generate significant positive returns over three to twelve month holding period. They explain these results by investors buying past winners and selling past losers, which causes the prices of stocks move away from their long-run values temporarily. These results are also found for stocks outside the United States. Rouwenhorst (1998) finds short-term momentum in returns in twelve European countries.

The semi-strong form efficient market hypothesis has also faced empirical challenges over the past decades. This hypothesis states that investors cannot earn superior risk-adjusted profits using any publicly available information. Banz (1981) finds that the average returns of small caps are too high given their market beta and that the average returns of large caps are too low. Fama and French (1992) report more recent findings. They classify stocks into deciles based on market capitalization and measure the average return for each decile over the first year after formation. They find that the small caps outperformed the large caps by 0.74% on average per month over the period 1963 to 1990. Those studies demonstrate that small caps earn higher average returns than is predicted by the capital asset pricing model.

Evidence on corporate announcement effects also suggests a violation of the semi-strong form efficient market hypothesis. Bernard and Thomas (1989) divide stocks based on the size of the earnings surprise in the most recent earnings announcement into ten different portfolios. They analyze the performance of each portfolio during the sixty days after the earnings announcement and find that on average the (tenth) portfolio with good news outperforms the (first) portfolio with bad news by 4%. Bernard (1993) summarizes the studies that analyze the (under)reaction of stock prices to earnings announcements. Stocks with positive earnings surprises earn relatively high returns in the period prior to the earnings announcement and stocks with negative earnings surprises earn relatively low returns. In the post announcement period, stocks with higher earnings surprises also earn higher returns. This is after the portfolios have been formed, which means that the market underreacts to new information and slowly revises the company's stock price. This bias will be corrected in the following periods. This phenomenon is known as post-announcement drift.

Other publicly available variables that predict future returns are, for example, fundamentals scaled by price. An example of a scaled-price ratio is the book-to-market ratio. Companies with high book-to-market ratio are called value stocks. Companies with low book-to-market ratio are called growth stocks. US and international empirical studies show that value stocks generate higher returns than growth stocks, which cannot be explained on the basis of risk differentials using the capital asset pricing model (see table 2.1). Fama and French (1992, 1996) classify stocks into deciles based on the book-to-market ratio and calculate the average return for each decile the year after formation. They find that the average monthly return of value stocks is 1.53% higher than the average monthly return of growth stocks. The difference between the average return between value and growth stocks is called the value premium and cannot be explained with the market beta. Other measures like the price-to-earnings ratio generate similar results; the difference in average returns is 0.68% per month. Lakonishok, Shleifer and Vishny (1994) divide stocks based on the past 5-year sales growth and cash flow-to-price ratio into nine portfolios. The value portfolio (which is the portfolio with the highest cash flow-to-price ratio and the lowest past growth rate) outperformed the growth portfolio (portfolio with the lowest cash flow-to-price ratio and the highest past growth rate) with an average annual return of 10.7%. These results are also found for stocks outside the United States. Fama and French (1998) provide consistent evidence of the value premium for a broad sample of countries outside the US. In almost every country, the value portfolio generates a higher average return than the growth portfolio. These results hold up across a variety of value-growth determinants. Since the existence of the value premium is an important driver of our research, some major studies regarding this issue have been summarized in table 2.1.

Finally, the last group of anomalies shows that prices deviate from fundamental values and where mispricing can be established beyond any reasonable doubt. Froot and Dabora (1999) study Siamese twin stocks such as Royal Dutch and Shell. These two stocks are traded at different places, but should move together regarding that they have the same cash flow stream. They find that Royal Dutch is more sensitive to movements in the US market while Shell commoves with the UK market. A study by Cooper, Dimitrov and Rau (2001) shows that during the internet hype a corporate name change into dotcom related internet names lead to positive announcement returns on the order of 74% in the ten days surrounding the announcement. Their findings also indicate co-movements in stock prices that are not related to common fundamentals, since in their research also stocks of firms with non-internet related businesses highly benefited from a dotcom name change.

Table 2.1: Summary empirical evidence for the value premium

The table shows a variety of value-growth determinants: *P/E* measures the price-to-earnings ratio, *B/P* reflects the book-to-price ratio, *CF/P* measures the cash flow-to-price ratio, *D/P* measures the dividend-to-price ratio and *MV* reflects the market capitalization.

	Period	Market	Variable	Value premium
Basu (1977)	1957-1971	US: NYSE	*P/E*	6.96% per year
Rosenberg, Reid and Lanstein (1985)	1973-1984	US: NYSE	*B/P*	0.36% per month
Chan, Hamao and Lakonishok (1991)	1971-1988	Japan: Tokyo stock exchange	*B/M*, *E/P*, *CF/P*	1.1%, 0.4% and 0.8% per month
Fama and French (1992)	1963-1990	US: NYSE, AMEX and NASDAQ	*B/M*, *E/P*	1.53% and 0.68% per month
Capaul, Rowley and Sharpe (1993)	1981-1992	International	*B/P*	US:1.06%, Japan: 3.43%, UK: 1.09%, Europe: 1.30% and Global: 1.88%

[Table 2.1 continued]

Lakonishok, Shleifer and Vishny (1994)	1968-1990	US: NYSE and AMEX	*B/M*, *E/P* *CF/P* & 5-year sales growth rates	6.3%, 3.9% and 10.7% per year, respectively.
La Porta, Lakonishok, Shleifer and Vishny (1997)	1971-1992	US: NYSE, AMEX and NASDAQ	*B/M*	4.0% first post formation year
Fama and French (1998)	1975-1995	International	*B/M*, *E/P*, *CF/P* and *D/P*	Based on *B/M*: US: 6.8%, Japan: 9.85%, UK: 4.62%, Netherlands: 2.3%.
Chan, Karceski and Lakonishok (2000)	1970-1998	US: NYSE, AMEX and NASDAQ	*B/M* and *MV*	Small stocks: 6.5%, Large stocks: 3.3% per year

2.3 Asset return models related to market anomalies

In section 2.2, several anomalies are described that cannot be explained by a simple model of risk and return, such as the capital asset pricing model (CAPM). These anomalies may have different explanations. Boudoukh, Richardson and Whitelaw (1994) categorize the explanations into three camps: "loyalists, "revisionists" and "heretics". "Loyalists" defend the efficiency of stock markets by pointing to data mismeasurement or to market imperfections.
"Revisionists" defend the efficient market hypothesis with time-varying risk premiums. The third group, the "heretics", believes that the market is not rational and that psychological factors influence the pricing of securities. The heretics believe that profitable risk-adjusted trading strategies exist.
In our opinion, both the loyalists and revisionists can be grouped together since they both try to defend the efficient market hypothesis. This group can be labeled as rationalists and Heretics can be called the behavioralists.

2.3.1 Rationalists

The rationalists believe that the superior performance of investment strategies, for example the difference in performance between losers versus winners and value versus growth stocks, are not an anomaly. They assume that the efficient market hypothesis holds, and that stock returns cannot be predicted. Rationalists give two explanations for the cross-sectional differences in returns. The first explanation is that superior returns are a result by chance, which cannot be found outside the sample. Researchers who support this thought are for example Black (1993) and MacKinley (1995). A second explanation is that the superior returns are a result of exposure to a common risk factor that is initially ignored in analyzing the returns of stocks. Higher average returns are a result of higher risks. Researchers who support this view are Fama and French (1993) with their three-factor model, the time-varying risk model of Berk, Green and Naik (1999) and the model of Zhang (2000). In section 2.4, we describe three rational models that explain the cross-sectional differences in returns from a rational point of view.

2.3.2 Behavioralists

Behavioralists have suggested alternative explanations for stock market anomalies. They study the observed behavior of investors and focus in particular on those elements that deviate from rational behavior. A field of finance that proposes psychology-based theories to explain stock market anomalies is behavioral finance. Shleifer and Summers (1990) suggest that behavioral finance rests on two foundations. The first foundation is investor sentiment. Investors deviate from the maxims of economic rationality. This implies irrationalities in investors' behavior by forming beliefs, and in their preferences, or in how they make decisions, given their beliefs. This may lead to deviations between security prices and their fundamental values. Investor sentiment is an important fundament for behavioral finance.

The second foundation is limited arbitrage, which explains why inefficiencies in markets remain after the market is disrupted by irrational investors (indicated as noise traders). With unlimited arbitrage markets remain efficient even when some investors are irrational. The arbitrageurs will digest the large demand shocks by noise traders and markets will become efficient again. Arbitrage is limited and the reason is that arbitrage is risky. Many securities do not have perfect substitutes, and if they have, the prices do not necessarily converge directly to their fundamental values because of the existence of "noise trader risks" (De Long *et al.*, 1990, Shleifer and Vishny, 1997).

Shleifer (2000) categorizes the deviations of investors (investor sentiment) from the standard decision-making model in three broad classes: non-Bayesian expectation formation, attitudes towards risk, and sensitivity of decision making to the framing of problems.

The first class concentrates on beliefs or the way in which people process information. By predicting uncertain outcomes, investors show behavior different from Bayesian rationality. Instead, investors rely on a limited number of heuristics to assess probabilities and to evaluate sample outcomes. The heuristics may result in good decisions, but sometimes may lead to biased decisions caused by ignoring relevant information and/or processing irrelevant information. Tversky and Kahneman (1974) identify three heuristics that affect probability assessments and the evaluation of sample outcomes: representativeness, availability, anchoring and adjustment. In the next paragraphs the three heuristics are described, followed by some other sources of biases that influence the assessment of outcomes as well.

1. Representativeness suggests that people evaluate the probability of an event by the degree to which the event reflects similarities with comparable 'known' events. Thus when something looks the same they think that the probability of the event is also the same. This leads to errors because representativeness or similarity is influenced by other factors than that should affect the judgment of probabilities. An example is the sample size neglect (Kahneman and Tversky (1974)), which means that people fail to take the size of the sample into account. People will find the characteristics of a short sequence equally informative as the characteristics of a sequence generated by a random process. Thus they expect that the characteristics of the process that will be

represented, will not only be found in the entire sequence (global level), but also locally in small parts of the entire sequence. Another example is the base rate neglect, which means that people put to much weight on salient features and to little weight on the base rate probability. For example, if a detailed description about someone's personality matches up with the subject's experiences with people of a particular profession, the subject tends to overestimate the actual probability that the given individual belongs to that profession.

2. The availability heuristic generates biases that arise because people base the probability of an event on the ease with which instances or occurrences can be brought to mind. The ability to recall instances or occurrences depends also on the actual frequency, familiarity, salience, recency, imaginability and prominence of the occurrences. Events that are familiar, salient, recent, easy to imagine or prominent are judged as more frequently-occurring than events that do not have these qualities.
3. The anchoring and adjustment heuristic describes that biases arise, because people make estimates from an initial starting point (anchor) and adjust insufficiently to the final value. For example, in the absence of any solid information, past earnings are likely to act as anchors for making an earnings forecast for next year. In predicting next year's earnings it is easy to take last year's earnings and adding two percent to the number allowing for the circumstances of the present case. Anchoring and adjustment contains conservatism in updating, "the adjustment", in addition to the incorrect priors generated through the choice of the anchor. Conservatism implies that people fail to revise their beliefs in the face of new information to the same extent as Bayes' theorem (Edwards (1968)).
4. Other biases that influence asset pricing are generated by overconfidence (self-attribution), optimism and cognitive dissonance. Overconfidence and optimism are two psychological errors that often

occur simultaneously. People have a tendency to be overconfident about the precision of their knowledge. Odean (1998) shows that the degree of overconfidence varies among professions. It is strongest in professions that can easily shift the blame for mistakes on others or unforeseen circumstances. Overconfidence stems from two biases; the illusion of control and self-attribution. Illusion of control implies that people believe to be in control of a situation far more often than they really are. Self-attribution means that people attribute good outcomes to good personal skills and bad outcomes to bad luck. Optimism implies that people display unrealistic views of their prospects and abilities. For example, the planning fallacy (Buehler, Griffin and Ross, 1994): people consistently underestimate the time they need to complete tasks. The combination of overconfidence and optimism leads investors to overestimate their knowledge, understate risk and overstate their ability to control the situation. People want to reduce cognitive dissonance in order to avoid mental inconsistencies. Therefore they ignore information that suggests that they have made the wrong decisions and search for information that supports their decisions. Furthermore they surround themselves with people that made the same decisions or have the same opinions.

The second class described by Shleifer (2000) deals with the evaluation of financial and non-financial outcomes and in particular the risk attitudes of decision makers. An important ingredient of any model trying to explain asset prices or trading behavior is the assumption about investors' risk preferences. Finance theory is generally based on the expected utility model. This model assumes that investors' preferences satisfy Von Neumann and Morgenstern (1947) rationality. Von Neumann-Morgenstern rationality implies that investors assess gambles at the level of total wealth.

Kahneman and Tversky (1979) show with their prospect theory that investors look at gains and losses relative to some reference point. The reference point may change with the situation and is determined by the subjective perceptions and feelings of the individual. Another important feature is the S-shape of the utility function (in Kahneman and Tversky's terminology: value function), which is convex below the reference point and concave above. This means that the marginal utility of both losses and gains generally decreases with their magnitude and individuals are risk seeking with losses and risk averse with gains.In addition, the utility function is steeper for losses than for gains, what means that the response of individuals to losses is more extreme than the response to gains. This is called loss aversion.

The final feature of the prospect theory is that this theory treats preferences as a function of decision weights and assumes that these weights do not always correspond with probabilities. Specifically, prospect theory states that decision weights tend to overweight small probabilities and underweight moderate and high probabilities. The overweighting of small probabilities can give rise to risk seeking in choices involving sure losses and to risk aversion in choices involving sure gains. Kahneman and Tversky (1979) label the phenomenon where people place more weight on outcomes that are certain relative to outcomes that are merely probable, as the 'certainty effect'.

Finally, the last class of Shleifer (2000) describes that individuals use framing to make decisions. The form that is used to describe a decision problem is called a frame. Framing refers to the way that a problem is presented to the decision maker. In traditional finance, framing is transparent and the form of decision information is irrelevant for the decision process: only into substance is relevant. Earlier in this section, we saw how the prospect theory could explain why people make different decisions in situations with identical final wealth levels. According to the prospect theory, decision problems are analyzed by framing different outcomes. Kahneman and Tversky (1979) show that when people have to deal with choices in the face of risk and uncertainty, frame dependence is important. People tend to make different choices when problems are represented in different frames. For example, if a decision is framed in terms of losses, people tend to choose riskier outcomes whereas the same decision is framed in terms of gains people tend to avoid risks and choose the more certain outcome. In addition, choices are also made based on norms, habits and expectancies of the decision maker.

In summary, behavioralists study the observed behavior of investors and focus in particular on those elements that deviate from rational behavior. According to market efficiency not all participants are required to be rational. Only a small number is required to be rational, and they will drive the rest out of the market. Behavioralists argue that arbitrageurs (rational investors) cannot correct for the mistakes made by the irrational investors, because arbitrage is risky and therefore limited. This leads securities to be priced incorrectly subject to investor sentiments, which result in market inefficiencies (see list of anomalies described in section 2.2).

2.4 Rational models

Models such as CAPM cannot explain the previously discussed anomalies. As a consequence, several rational models have been developed in the last decade to explain these anomalies. In this section, we present three different rational models that are relevant for the following chapters, because they try to explain some of the most common anomalies (e.g. value premium) from a rational point of view. Both the three factor model of Fama and French

(1993) and the time-varying risk model of Berk, Green and Naik (1999) provide rational explanations for size and the book-to-market ratio in determining differences in stock returns. Fama and French create an empirical three-factor model where size and book-to-market ratio are systematic risk factors. Berk, Green and Naik create a time-varying risk model where size, book-to-market ratio and the number of growth options explain stock returns. The multi-period model of Zhang (2000) shows close resemblance with the time-varying risk model of Berk, Green and Naik by using real options. This model is not only based on the number of growth options but also takes the option to abandon into account.

2.4.1 Three-factor model

The three-factor model of Fama and French (1993) explains the expected excess return on a stock or portfolio with the sensitivity of its return to three factors. The factors are the excess return of the market portfolio, the difference between the returns on portfolios of small and large stocks (*SMB*), and the difference between the returns on portfolios of high and low book-to-market ratio stocks (*HML*). The expected excess return of portfolio *i* is:

$$E(R_i) - R_f = b_i \left[E(R_M) - R_f \right] + s_i E(SMB) + h_i E(HML), \quad (2.1)$$

where $E(R_M)$-R_f, $E(SMB)$ and $E(HML)$ are excess factor returns and the factor sensitivities b_i, s_i and h_i are the slopes obtained from a time-series regression of excess stock returns on excess factor returns,

$$R_i - R_f = a_i + b_i \left(R_M - R_f \right) + s_i SMB + h_i HML + \varepsilon_i, \quad (2.2)$$

Fama and French (1993) interpret the market, size and book-to-market equity as risk measures. These factors represent sources of systematic risk, which means that risk requires compensation in the form of higher expected returns. Fama and French (1993) interpret the risks captured by the book-to-market ratio of equity and size as proxies for distress. Following the economic interpretation of Chan and Chen (1991), distressed firms are defined as firms which are less efficiently run, have higher financial leverage and have lower accessibility to external funds. This leads to prices to be more sensitive to changes in the economy. Distressed firms are less likely to survive unfavorable economic conditions and are therefore riskier than other firms. Because distressed firms are riskier, the cost of capital is higher, which leads to higher expected returns. For example, firms with poor past earnings and high financial leverage (high loadings on *SMB* and *HML*) may have restrictive accessibility to external funding. This leads to cash flow problems. In times when economic conditions are poor, these firms are more likely to get into financial difficulty.

To explain the book-to-market ratio of equity in terms of risk Fama and French (1992) use two financial leverage measures, a measure of market leverage, *A*/*ME*, and a measure of book leverage, *A*/*BE* (where *A* is the book value of assets, *BE* is the book value of equity and *ME* is the market capitalization). Fama and French calculate both measures for the US stock market over the period 1963 to 1992. They show that both measures are statistically significant and close in absolute value, but with opposite signs. Both measures of leverage explain expected returns and the sum of the log of both measures, e.g. A/ME and A/BE, is expressed as the book-to-market ratio of equity, BE/ME. Weak firms with low earnings have high book-to-market ratios of equity and tend to have positive slopes on *HML*. Strong firms with high earnings have low book-to-market ratios of equity and negative slopes on *HML*.

One general feature of the three-factor model is that factor loadings (estimated slope coefficients) and not firm characteristics (size and BE/ME) determine average returns. Fama and French (1993) explain this with two empirical facts: first, high book-to-market stocks and small size stocks have high returns; second, within each portfolio (e.g. high book-to-market and small size), stocks covary with each other. Fama and French argue that these two findings occur because the book-to-market ratio and size proxy for financial distress, which lead to a risk premium. Daniel and Titman (1997) question whether stock returns are driven by risk based elements (factor loadings) or by firm characteristics. They cast doubt on the prediction that value stocks earn higher returns because such stocks have higher loadings on the book-to-market factor and not because they have high book-to-market ratios. They perform double sorts of stocks on both loadings on book-to-market and book-to-market ratios, and show that stocks with different loadings but the same book-to-market ratio do not differ in average returns. Lally (2004) argues that the sensitivity coefficients in the Fama and French model must be related to the firm's leverage as proposed by Modigliani and Miller (1958,1963). He shows that the empirical formulas developed by Fama and French (1997) to show the cost of equity through time, are inconsistent with the Modigliani and Miller propositions since they do not separate leverage from other factors that influence sensitivity coefficients.

Another shortcoming of the three-factor model is that it cannot explain momentum returns in the short run as documented by Jegadeesh and Titman (1993). Carhart (1997) adds a fourth factor, momentum, to the three-factor model. He sorts mutual funds into decile portfolios on one-year past returns and shows that the momentum factor and size factor can account for most of the variation in returns.

The final comment on the Fama-French model is that the three-factor model defines value and growth stocks with the book-to-market ratio (i.e. high and low book-to-market ratios). However, value and growth stocks can be defined with a variety of value-growth determinants (e.g. cash flow-to-price, earnings-to-price). However, Fama and French (1996) show that the three-factor model also captures the returns of portfolios formed on cash flow-to-price (CF/P) and earnings-to-price (E/P). Firms with low CF/P and E/P ratios have similar slopes as low BE/ME ratio stocks (negative loadings on *HML*), which mean low expected returns. Firms with high CF/P and E/P have similar slopes as high BE/ME ratios stocks (positive slopes on *HML*), which implies high expected returns.

2.4.2 Model with growth options

Berk, Green and Naik (1999) develop a model in which the valuation of cash flows and the firm's options to growth in the future leads to dynamically conditional expected returns. The changes in risk are related to investment opportunities. Firms that perform well tend to be firms that have exploited valuable investment opportunities.

The value of the firm is based on assets in place that generate current cash flows and on options to make positive net present value investments in the future. Each period existing assets may die off and new investment opportunities may arise. If a firm exploits those opportunities, the firm's systematic risk will change. For example, if a firm decides to invest in a low risk investment opportunity, its value will increase and the systematic risk will decrease. Because the systematic risk decreases over that period, the returns will also be lower on average. On the other hand, when the firm loses a low-risk asset, the systematic risk will increase, resulting in higher returns on average.

The central issue in the Berk, Green and Naik-model is the distinction between two kinds of assets: assets in place and assets that embody future growth opportunities with positive net present values. The value of the firm is:

$$P(t) = \sum_{j=0}^{t} V_j(t)\chi_j(t) + V^*(t), \tag{2.3}$$

where *P(t)* is the value of the firm at time *t*. The first term expresses the value of future cash flows of assets in place, where $\chi_j(t) = 0$ if the project has expired and $\chi_j(t) = 1$ if project *j* is still in operation. The second term presents the current value of future investment opportunities. Equation 2.3 can be expressed in terms of risk-adjusted discounted future cash flows of assets in place and cash flows from future investment opportunities:

$$P(t) = b(t)e^{\overline{C}-\beta(t)}D[r(t)] + Ie^{\overline{C}}J^*[r(t)], \tag{2.4}$$

where *b(t)* is the book value of assets that are currently alive, $\overline{C}$ is the parameter that controls the mean of the cash flows, *β(t)* the systematic risk of a projects cash flows, *D* is the value of a perpetual, riskless consol bond, *r(t)* is the one-period riskless, continuously compounded interest rate, *I* is the investment required to undertake a new project and J^* is the value of growth opportunities. The growth opportunities are in terms of European call options on pure discount bonds maturing at different dates.

In this model the expected return depends on the life cycle of the firm. Mature firms have different expected returns than developing growth firms. If there are no ongoing projects, the value of the firm will only consist of growth opportunities. In this case the value of the firm depends only on the interest rate and not on systematic risk (this is in line with the Black and Scholes option pricing model). If the number of ongoing projects goes to infinity, the growth component will become negligible. Then the expected return depends on systematic risk and interest rates. This means that changes in the systematic risk component are more important for mature firms and less important for growth firms. The firm's expected return can be derived from equation 2.4 and is given by the following equation:

$$E_t\left(1+R_{t+1}\right)=\frac{\pi D_e(r(t))}{D\left(r(t)\right)}+\pi e^{\bar{C}}\left[\frac{b(t)}{P(t)}\right]+Ie^{\bar{C}}\left[J_e^{*}[r(t)]-J^{*}[r(t)]\frac{\pi D_e(r(t))}{D\left(r(t)\right)}\right]\left[\frac{1}{P(t)}\right], \quad (2.5)$$

where D_e is the expected value of a perpetual, riskless consol bond. The first term of this equation reflects the effects of changing interest rates on the value of the cash flows produced by the assets in place. The consequence of higher interest rates is that future cash flows will be discounted at a higher rate, which leads to lower prices and to higher expected returns. Furthermore, higher interest rates also affect the current expectations about future values of systematic risk. In general, when the interest rate is high, a firm will undertake fewer projects. As a consequence firms will only accept projects with low systematic risk.

The second term is a proxy for the book-to-price ratio. The term $\pi e^{\bar{C}}$ reflects the project's expected cash flows which depreciate at rate π. Since, $\pi e^{\bar{C}}$ is a positive constant, the firm's expected return is positively correlated with the book-to-price ratio. Book-to-price ratio varies with changes in the systematic risk of the firm. The systematic risk of the firm changes, because each period existing cash flows can die off and new projects arrive. For example, when a firm adopts a new project with positive net present value that has low systematic risk, the firm creates value and lowers the average systematic risk of its cash flows in the next periods. From equation 2.4 it follows that if systematic risk decreases the current price will increase. Assuming that the expected cash flows are constant, a current price increase will result in lower expected returns in the future. This is consistent with classical finance where a positive relationship between systematic risk and expected returns exists. In addition, systematic risk changes with the fraction of existing assets relative to growth options, because systematic risk plays an important role in determining expected returns of mature firms, but is less important for growth firms (see explanation of equation 2.4).

The last term in the equation represents the value of the growth options. This is the difference between the expected value of bond options $J_e^*[r(t)]$ and the current value of bond options $J^*[r(t)]$ (accumulated at the risk-free rate) divided by its price.

The value of the growth options depends on changes in the interest rate. Changes in the interest rate lead to changes in the value of the underlying assets. Because the risk premium of an option depends on the value of the underlying assets (out-of-the money options have higher risk premia than in-the-money options), the value of the growth options also changes with changes in the interest rates. Because of the changes in the value of the options the expected returns can change even if the assets in place of the firm remain unchanged.

In a number of simulation experiments Berk *et al.* show that expected returns explain the important features of the cross-sectional and time series behavior of stock returns that is found in empirical research (e.g. explanatory power of book-to-market ratio, size and the momentum versus contrarian effects at different horizons). They compose an equally-weighted portfolio from simulated stock returns and run time-series regressions of the equally-weighted portfolio against the market beta, log of market value and log of book-to-market ratio. The results of the regressions show almost similar results in magnitude and direction as obtained by Fama and French (1992). Only when the market beta and market value are combined in one regression, the model shows different signs than the results of Fama and French show. Berk *et al.* also test the profitability of momentum and contrarian strategies. The model reproduces the patterns of contrarian returns at shorter horizons and momentum returns at intermediate horizons. A shortcoming of the model is that it predicts excess returns at longer horizons than empirically has been shown. This model predicts that the contrarian strategy is profitable at a horizon of twelve months or less, where Conrad and Kaul (1998) show empirically that contrarian strategies are only profitable at horizons of about three months or less. Furthermore, this model predicts the maximum profitability of momentum strategies is approximately at sixty months, while Jegadeesh and Titman (1993) show that momentum strategies reaches the maximum of profitability at horizons of about nine to twelve months. The difference between this model and three-factor model is that expected returns are not only explained by size and leverage alone, but also by growth options. This model shows that two firms with identical book-to-market ratio's and different growth potential can have different expected returns. For example, when a firm has no growth potential and its value depends only on the existing assets, the last term of equation 2.5 drops out and expected returns depend only on the book-to-market ratio. This brings the book-to-market ratio in a different perspective, because the book-to-market ratio is generally used as measure for growth potential. Daniel and Titman (2001) support empirically that only size and leverage cannot explain expected returns.

2.4.3 Model with options to expand and to discontinue operations

In section 2.4.1, we discussed the three-factor model where size and the book-to-market ratio are systematic risk factors. The model of Berk *et al.*

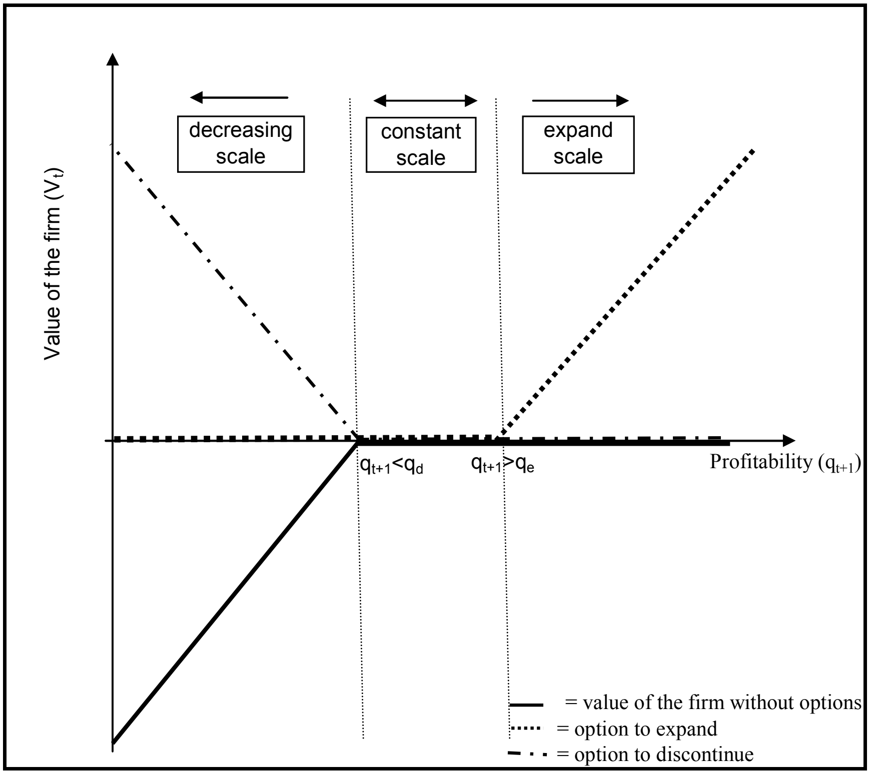
(1999) (section 2.4.2) added growth options to this model. In addition to the model of Berk *et al.*, Zhang's (2000) model allows firms not only to make rational choices to expand their operations when it is sufficiently profitable, but also to discontinue when it is sufficiently unprofitable. Zhang's model allows firms to make rational choices among investment or divestment alternatives (in terms of real options) whereby accounting signals are used to guide investment decisions. In addition, Zhang's model shows equity value as a nonlinear function of accounting variables, while the model of Berk *et al.* shows the equity value as a nonlinear stochastic process. The options to expand and to discontinue operations are determined by profitability and growth opportunities. It is a multi-period model where the valuation is based on three scenarios.

$$V_t = \frac{1}{R_f - 1} x_t^E + P_d(q_t) as_t + C_e(q_t) G, \tag{2.6}$$

where R_f equals one plus the risk free rate of return, x_t^E is the net value creation (or economic earnings), as_t is the bookvalue of assets in place, $P_d(.)$ is the put option to discontinue the operations, $C_e(.)$ the call option to expand the operations and G represents the growth potential. The first term in equation 2.6 shows the equity value if the firm operates at the same scale as the year before. The cash investment required at time t+1 is $ci_{t+1}=(1-\gamma)as_t$ (where γ represents the durability of the assets). The decisions whether to discontinue or to expand depend on future profitability, q_t. The option to abandon (second term in equation (2.6)) becomes in-the-money if the profitability to discontinue operations (q_d) are higher than the profitability to continue operations ($q_{t+1} < q_d \equiv (1-\gamma c_d)(R_f - 1)$. If the firm decides to discontinue operations, the value of the firm will depend on the assets in place and will be equal to $(1-c_d)as_t$, where c_d is the cost to discontinue. The option to expand (third term in equation 2.6) becomes in-the-money if the profitability to expand operations (q_e) is higher than the profitability to continue operations at the same scale ($q_{t+1} > q_e \equiv R_f - 1$).

Figure 2.1 visualizes the impact of profitability on the value of the firm. Under a specific level of q_{t+1} the put option becomes in the money which means that the value to discontinue is higher than the value to continue at the same scale ($q_{t+1}>q_d$).

Figure 2.1: Value of the firm according to model of Zhang (2000)



Above a specific level of q_{t+1} the call option becomes in the money which means that the value to expand is higher than the value to continue at the same scale as before ($q_{t+1}>q_e$). In order for the firm to rationally continue at its existing scale, profitability (q_t) will have to be above the profitability to discontinue and below the profitability to expand ($q_d \leq q_t \leq q_e$).

In this model investment decisions underlie value creation. Equity value is nonlinear in economic earnings and book value of assets in place. The model predicts that equity value is increasing with economic earnings for any level of book value of assets. In addition, for any level of economic earnings equity value is increasing with book value of assets for low profitable firms, equity value is insensitive for steady state firms, and equity value is decreasing with book value of assets for growth firms. The importance of book value of assets and economic earnings varies with profitability and growth potential. For low profitable firms, book value of assets is expected to dominate economic earnings, while for steady-state firms, economic earnings are predicted to dominate book value of assets. For growth firms, economic earnings and book value of assets are both important, and the impact of book value of assets increases with the magnitude of the growth potential. The model predicts that the impact of profitability and earnings growth rates on price-to-book is as follows: if profitability or growth opportunities are increasing, *P/B* increases. This implies that growth stocks are associated with high growth opportunities and high profitability and value stocks with low growth opportunities and low profitability. This is also shown by empirical studies, i.e. Lakonishok *et al.* (1994), which show that value stocks are stocks with low earnings growth rates and growth stocks are stocks with high earnings growth rates.

This model introduces an additional feature to the model by Berk *et al*. (1999). It has not only an option to expand but also an option to discontinue operations. Zhang's model shows that two firms with equal assets in place and equal growth potential can have different expected returns. For example, when a firm has no growth potential the last term of equation 2.6 drops out and the firm depends only on the existing assets and the option to abandon.

The main issue that merits discussion is whether the model of Zhang is able to explain anomalies such as the value premium and momentum in the short run. Chen and Zhang (2004) derive the return function from equation 2.6 and empirically test whether the model is able to explain the value premium. The return function shows that returns are explained by three factors, notably profitability-related information, capital investment and the change in growth opportunity. Chen and Zhang divide stocks into book-to-market deciles and perform regressions of each portfolio return against these three factors. They show that the difference in intercepts is of the same magnitude as the book-to-market effect found by Fama and French (1992). This implies that the book-to-market effect (value premium) is not subsumed by the model of Zhang.

In section 2.4, we have discussed three different rational models. The first two models try to explain market anomalies. The last model by Zhang is related to the model of Berk *et al*. in the way that he uses options to expand. In addition, Zhang adds options to abandon to the model. The first two models show that they are able to explain the value premium. Although the model by Zhang shows the relation between firm value and book-to-market ratio, it is not able to explain the value premium. The main issue in all of the three rational models is whether the models are able to explain anomalies such as the profitability of momentum strategies in the short run and contrarian strategies in the long run as is reported by empirical studies. In the next section, we present three behavioral models that are able to explain these two anomalies.

2.5 Behavioral models

With the growing number of empirical papers supporting the anomalies, behavioral models have been developed to translate the empirical evidence into theoretical frameworks. The majority of these models describe the interaction between informed versus uninformed investors and rational versus irrational investors. This section presents three behavioral models, which focus on investors' sentiment explicitly. These models try to explain how judgement biases of investors can produce overreaction to some events and underreaction to others. The models are based on two heuristics and two other sources of biases that influence the assessment of outcomes; representativeness, and anchoring and adjustment (conservatism), overconfidence and self-attribution bias. The general assumption in these models is that investors allocate funds based on past performance.

2.5.1 A model of investor sentiment

Barberis, Shleifer and Vishny (1998) construct a model of investor sentiment to reconcile the empirical findings of over- and underreaction with mental heuristics of investors. The overreaction evidence shows that over longer horizons security prices overreact to consistent patterns of new information pointing in the same direction. Securities with a long record of positive information tend to be overpriced. The underreaction evidence shows that new public information has a limited effect on prices. If there is a record of good information, prices keep trending up after the initial positive reaction. If there is a record of bad information, prices keep trending down after the initial negative reaction.

The model relates to two heuristics from the cognitive psychology to express under- and overreaction of new information, notably conservatism of Edwards (1968) and representativeness of Kahneman and Tversky (1974). Conservatism is a substitute for the anchoring and adjustment heuristic and means that investors slowly adjust their forecasts to new information (see section 2.2.1). Investors have prior views about a stock in question. When new information is revealed which is inconsistent with prior beliefs, investors adjust their expectations in the right direction, but by a smaller magnitude than the true normative rational Bayesian value. This implies an overweighing of the statistical base rate (probabilistic) information relative to the new statistical evidence. Because investors tend to stick to their prior beliefs, security prices are not adjusted sufficiently which results in underreaction of prices to earnings announcements. This can result in both positive and negative serial correlation in stock returns at short horizons. Overreaction is related to the representativeness heuristic where the recently perceived pattern is taken as representative for a persisting future pattern. For example, when a company has a consistent history of earnings growth over several years, investors may conclude that the past earnings growth history is representative for future growth potential. While, earnings growth may be nothing more than a random process, investors believe that the company belongs to a small distinctive population with high future growth potential. In overweighing past growth of the company, investors underweigh the statistical base rate (probabilistic) evidence of the small fraction of the population belonging to that high growth potential group. As a consequence investors tend to disregard the reality that a history of high earnings growth is unlikely to repeat itself. Therefore, investors overvalue the company and become disappointed when future information is revealed. This leads to overreaction in the long run.

The model consists of a representative, risk-neutral investor with a constant discount rate. There is only one security, which pays out all earnings as dividends. This means that the equilibrium price is the net present value of the future earnings as forecasted by the investor. In addition, prices reflect only the information that is contained in earnings. True earnings follow a random walk. The model represents an investor who does not realize that earnings follow a random walk. This investor believes in a world with two states, and for each state a different regime. In regime 1 earnings are mean-reverting and in regime 2 earnings are trending.

Regime 1:

	$y_{t+1} = y$[1]	$y_{t+1} = -y$
$y_t = y$	π_L	$1-\pi_L$
$y_t = -y$	$1-\pi_L$	π_L

Regime 2:

	$y_{t+1} = y$	$y_{t+1} = -y$
$y_t = y$	π_H	$1-\pi_H$
$y_t = -y$	$1-\pi_H$	π_H

In these matrices, π_L is small, between 0 and 0.5, and π_H is large, between 0.5 and 1. This implies that under regime 1 a positive shock is more likely to be reversed and under regime 2 a positive shock is more likely to be followed by another positive shock. Underreaction is caused if the investor puts more weight to regime 1 than to regime 2. In that case, when there is a positive earnings-shock the investor believes that earnings are mean-reverting in the next period. However, the positive earnings-shock is equally likely to be followed by a positive as a negative shock. Overreaction is caused if the investor puts more weight on regime 2 than regime 1. The investor believes that after a positive earnings-shock the chance of a new positive earnings-shock is higher than the chance of a negative earnings-shock. However, the investor does not realize that earnings follow a random walk and that the chance of positive or negative earnings surprises is equal.

[1] Where y_t is the shock to earnings at time *t*, which can take two values, $+y$ and $-y$

A Markov model is used to specify the underlying regime switching process. The current regime depends on what the regime was perceived to be prevailing last period. To value the security the investor has to forecast future earnings and uses the regime-switching model.

	$s_{t+1} = 1$	$s_{t+1} = 2$
$s_t = 1$	$1-\lambda_1$	λ_1
$s_t = 2$	λ_2	$1-\lambda_2$

If $s_t = 1$, the model is in the first regime, where the earnings shock is perceived to be generated by regime 1. If $s_t = 2$, the model is in the second regime, where the earnings shock is generated by regime 2. The parameters λ_1 and λ_2 are the transition probabilities from one state to the other state. The parameters, λ_1 and λ_2, are small which means that transitions from one state to another state in the investor's perception, the likelihood of regime switching is not high. Furthermore, λ_1 is smaller than λ_2, which implies that the investor believes that the occurrence of regime 1 is more likely than regime 2. The investor who forecasts earnings for the next period has to decide which regime is currently governing the earnings pattern. He observes the past series of earnings to decide which regime is generating earnings in the next period. If he has decided, he will use the transition probabilities to forecast the earnings change in the next period. The transition probabilities do not change in the investor's mind. Even after a long stream of earnings data he does not change the probabilities to a more random walk like model. At time *t*, he observes y_t and calculates q_t ($q_t = \Pr(s_t = 1 \mid y_t, y_{t-1}, q_{t-1})$), which is the probability that y_t is generated by model 1. The updating of q_{t+1} from q_t probability is based on Bayes Rule.

The model shows that if earnings shocks, y_{t+1}, have the same sign in period t+1 as period t, the probability, q_{t+1}, decreases. If earnings shocks have the opposite sign in period t+1 compared to period t the probability q_{t+1} increases.
The price of the security in this regime-switching model is:

$$P_t = \frac{N_t}{\delta} + y_t\left(p_1 - p_2 q_t\right), \tag{2.7}$$

where p_1 and p_2 are constants that depend on π_H, π_L, λ_1 and λ_2. The first term in the equation, $\frac{N_t}{\delta}$, is the stock's fundamental value, i.e. the price that would obtain if the investor used the "true" random walk process to forecast earnings changes. The second term, $y_t\left(p_1 - p_2 q_t\right)$, is the sentiment indicator which causes the price to deviate from its fundamental value. If on average there is underreaction, the stock price does not react sufficiently to the earnings shock, leading to a price beneath or above its fundamental value (depending on the nature of the earnings shock). Suppose the earnings shock is positive, the sentiment indicator will have to be negative resulting in a low value for p_1 relative to p_2q_t. If there is overreaction, the price is above (below) its fundamental value, which means that sentiment indicator will have to be positive (negative) resulting in a high (low) value of p_1 compared to p_2q_t.

The focus in this model is on learning about the time-series process of earnings shocks. Simulation results show that with an input of a random pattern of earnings changes, an output of a non-random pattern of stock returns is generated. This non-random pattern of returns, e.g. under- and overreaction pattern, is caused by two heuristics: representativeness and conservatism. Barberis *et al.* simulate for a large number of firms earnings, returns and prices. Then they form two portfolios on positive and negative earnings changes and calculate the difference in returns between the two portfolios in the year after formation. The differences in returns of the portfolios show post-earnings announcement drift and long-term reversals. In addition, they form portfolios on past performance in returns and on earnings-to-price ratios. They find momentum in the short run and cross-sectional forecasting power for scaled-price ratios.

Because this model assumes that a trend arises after a string of similar changes, the model neglects to explain the price drift after isolated information events (examined by event studies) such as a dividend cut, stock splits and stock issues. To explain such findings, it would be useful to extend the model with other kinds of news. In addition, it will be interesting to examine the effect of adding risk-averse arbitrageurs to the model. When arbitrageurs know the regime and the movements of noise traders, they can take advantage of the misperceptions. Because arbitrage is risky here, inefficiency will not be eliminated completely. This raises the following interesting question; to what extent can arbitrageurs bring the price closer to the fundamental value?

In addition, although the selected heuristics, representativeness and conservatism, are very plausible, they are not the only behavioral biases that can explain over- and underreaction. The next model uses other biases, cf. section 2.3.1, that can explain over- and underreaction.

2.5.2 Investor psychology and security market over- and underreaction

Daniel, Hirshleifer and Subramanyam (1998) provide a model that explains short-term momentum and long-term reversals, which is different from the approach of Barberis, Shleifer and Vishny (1998). The model by Daniel *et* *al.* emphasizes the roles of overconfidence and the self-attribution bias in the way investors react to private and public information. Overconfident investors are defined as investors who think that they are more able to value securities than they actually are, resulting in forecasting errors. The self-attribution bias determines the degree of overconfidence endogenously. When an investor receives a private signal, and a subsequent public signal is in agreement with initial private information, investors become even more confident. If the public signal contradicts with the initial private signal, overconfidence will fall, but not proportionally. This model describes two phases, the overreaction phase and the correction phase. The overreaction phase is the part of the impulse response prior to the peak and the correction phase is the time after the peak. In this model agents are divided into two groups. The first group represents uniformed, risk averse investors and the second group represents informed, risk neutral investors. We present the dynamic confidence model where a sequence of dates are described. At date 0 individuals begin with their endowments and identical prior beliefs, and trade solely for optimal risk-transfer purposes. At date 1 the overreaction phase starts. The informed investors receive a noisy private signal about the underlying security value and trade with the uniformed investors. The private signal is:

$$s_1 = \theta + \varepsilon, \tag{2.8}$$

where ε is normally distributed with variance σ_ε^2, which is independent of the signal for the terminal value of the risky security, θ. Overconfidence in the private signal causes the price to overreact to this new information at date 1.The implication for the price at date 1 is:

$$P_1 = E_C[\theta \mid \theta + \varepsilon] = \frac{\sigma_\theta^{\ 2}}{\sigma_\theta^2 + \sigma_C^2}(\theta + \varepsilon), \tag{2.9}$$

where E_c is the expectation of the informed traders' confident beliefs, σ_θ^2 is the variance of the risky security, and σ_C^2 reflects the investors' assessment of noise variance ($\sigma_C^2 < \sigma_\varepsilon^2$). At time 2, the noisy public signal arrives.

$$s_2 = \theta + \varepsilon^* \tag{2.10}$$

If the public signal at date 2 disconfirms the private signal, *sign* $(\theta + \varepsilon) \neq sign(s_2)$, then confidence decreases by little or remains constant. Because the signal is uninformative, the price does not move at time 2. However, if the public signal at date 2 has the same sign, *sign* $(\theta + \varepsilon) = sign(s_2)$, and therefore confirms his trade, the investor becomes more confident. The new price, calculated using the new level of assessed variance of ε, at date 2 is:

$$P_{2C} = E_C[\theta \mid \theta + \varepsilon, s_2] = \frac{\sigma_\theta^{\ 2}}{\sigma_\theta^2 + \sigma_C^2 + k}(\theta + \varepsilon), \tag{2.11}$$

where k determines the level of assessed noise variance. Because of the increase in investors' confidence, the investors' assessment of the noise variance decreases to σ_C^2-k ($0 < k < \sigma_C^2$). The continuing overreaction leads to positive autocorrelation during the initial overreaction phase ($cov(P_2-P_1,P_1-P_0)>0$). The overreaction of dates 1 and 2 must be reversed in the long run. In the correction phase, starting at date 3, the following public signal will be revealed:

$$s_3 = \theta + \eta \,, \tag{2.12}$$

where η is normally distributed with variance σ_p^2, which is independent of the risky security, θ. For simplicity, the second public signal does not cause overconfidence to be affected. Eventually, as more information is released and the public information becomes precise enough, the price will move to the full-information value θ. This process causes positive autocorrelation during the correction phase ($cov(P_3-P_2,P_4-P_3)>0$). Overreaction to private information and the underreaction to public information tend to produce short-term momentum in stock returns. Long-term reversals are the result of public information, which starts to dominate private information.

To explain the post-event abnormal price trends with the same sign as the average event-based return, two kinds of public events, non-selective and selective events, are distinguished. The non-selective event occurs with independence of the mispricing at date 2. The events can be characterized as news disclosed by outside sources, such as regularity events. The selective event at date 3 is an event whose occurrence depends on the mispricing at date 2. These events can be characterized as corporate events. For example, if a manager beliefs that the price of the security is undervalued, he will announce to repurchase shares. The public signal causes an intermediate price reaction that absorbs a fraction of the mispricing. This leads to the prediction of momentum for selective public events (although this is important for the model were confidence is constant, the dynamic confidence model shows that non-selective events can cause momentum). The magnitude of the mispricing (e.g. price/fundamental ratio) is positively correlated with the expected size and the probability of the selective event. For example, equity issues will tend to occur when price/fundamental ratios are high, because managers want to exploit the mispricing. When securities are underpriced, the probability of good-news selective events will increase.

A key difference with the model of Barberis *et al.* (1998) (section 2.5.1) concerns the basis of over- and underreaction. The model of Barberis *et al.* (1998) explains the drift in prices as a result of underreaction to new information. The model by Daniel *et al.* is different in the way, that it describes momentum as a result of overreaction to private information. Underreaction occurs at a later date, notably the correction phase.

In addition, this model suggests that because of overconfidence, investors buy additional stocks when the public signal is in agreement with their private signal. Odean (1999) shows that investors are inclined to purchase additional stocks that have declined in price after their initial purchase instead of stocks that have gone up.

2.5.3 Style investing

The previous two models have focused on over- and underreaction to information as a source of momentum in stock returns. While over- and underreaction to news has been documented empirically, it may not describe some important features of momentum; for example, price bubbles, in which prices continue to drift upward without much news, can also occur because investors are simply chasing a trend. Such price bubbles, which exhibit positive autocorrelations, are not well described by over- or underreaction to news about fundamentals. In this section, we describe a model by Barberis and Shleifer (2003) that exhibits the same properties as in price bubbles, but in this model, it is simply a reflection of style level phenomena. Barberis and Shleifer develop a model that explains the impact that style investing can have on financial markets and security valuation.

They combine style-based portfolio selection with a mechanism of how investors choose among styles. The model has two kind of investors, fundamental traders and switchers. The fundamental traders act as arbitrageurs that try to prevent the price of each asset to deviate too far from its expected final dividend. The investment policy of switchers is determined by two distinctive characteristics. Firstly, switchers classify assets into categories where they give each category a label. In this way, switchers try to simplify the information processing by making their decisions on a category level rather than an individual asset level. In this model, style investing is described as a production of life cycles of investment styles. The choice for a particular style depends on the relative past performance. Good fundamental news about the securities in a style is responsible for the start of a style[2]. When a style has a good past performance relative to other styles, switchers allocate their investments to that style and withdraw resources from other styles. If the style matures, good past performance is important to add new resources to a style. The style disappears when bad fundamental news arrives (or good news about the fundamentals of a competing style arrives) or due to arbitrage. A consequence of style investing is the emergence of life cycles in investment styles.

Barberis and Shleifer (2003) assume that each security fits to one style. There are *2n* risky assets in fixed supply, and a risk free asset and cash in perfectly elastic supply with zero net return. Suppose the world has only two kind of styles, *X* and *Y,* with securities in X and securities in Y. The switcher's demand of asset *i* in style *X* at time *t* is:

---

[2] A style can be defined as a classification of assets into a category with similar performance characteristics.

$$N_{i,t}^{S} = \frac{1}{n}\left[A_X + \sum_{k=1}^{t-1} \theta^{k-1} \frac{\left(\Delta P_{X,t-k} - \Delta P_{Y,t-k}\right)}{2} \right], \tag{2.13}$$

where n is the number of securities in style X, A_X is a constant, which represents the long run switchers target demand for style X, the parameter θ measures how far back investors look when they compare the past performance of styles, and $P_{X,t}$ and $P_{Y,t}$ are the average prices of a share across all assets in style X and Y. Furthermore, the switcher's demand of asset j in style Y at time t is

$$N_{j,t}^{S} = \frac{1}{n}\left[A_Y + \sum_{k=1}^{t-1} \theta^{k-1} \frac{\left(\Delta P_{Y,t-k} - \Delta P_{X,t-k}\right)}{2} \right] \tag{2.14}$$

The second kind of trader is the fundamental trader. The fundamental trader acts as an arbitrageur who does not want prices to diverge too far from their fundamental value. In contrast to switchers who base their expectations on past performance, fundamental traders base their expectations of the fundamental value. Assume that the fundamental traders have an amount W^F to allocate and have no constraints to their allocations, then they have to solve the following:

$$\max_{N_t} E_t^F \left(-\exp\left[-\gamma \left(W^F + N_t^{'} \left(P_{t+1} - P_t \right) \right) \right] \right), \tag{2.15}$$

where γ is the degree of risk aversion, N is the number of shares allocated to each risky asset, P_t is a price-vector for all assets, and E_t^F is the fundamental traders expectations at time t. Combining the switchers' demand for securities X and Y and the fundamental traders' demand, the price function is as follows:

$$P_t = D_t + \gamma V N_t^{\,s}, \tag{2.16}$$

where D_t is the dividend to be paid at time t and V is the covariance-matrix of returns. Equation 2.16 shows that the fundamental traders are not able to push back the price to its fundamental value D. This is supported in the empirical literature. The covariance matrix of returns in equation 2.16 can be simplified with additional assumptions:

$$V^{ij} = \begin{cases} \mathrm{cov}\left(\Delta P_{i,t+1}, \Delta P_{j,t+1}\right) = \sigma^2, i = j \\ \mathrm{cov}\left(\Delta P_{i,t+1}, \Delta P_{j,t+1}\right) = \sigma^2 \rho_1,\ i \neq j, i, j \text{ in the same style} \\ \mathrm{cov}\left(\Delta P_{i,t+1}, \Delta P_{j,t+1}\right) = \sigma^2 \rho_2, i \neq j\ i, j \text{ in different styles} \end{cases}, \tag{2.17}$$

where the variance of all asset returns is the same, the return correlation between assets in the same style is the same and the correlation between assets in different styles is the same. Substituting the covariance structure 2.17 into equation 2.16, the price of an asset i
in style X at time t is:

$$P_{i,t} = D_{i,t} + \frac{1}{\phi}\sum_{k=1}^{t-1} \theta^{k-1}\left(\frac{\Delta P_{X,t-k} - \Delta P_{Y,t-k}}{2}\right), \tag{2.18}$$

where,

$$\phi = \frac{n}{\gamma\sigma^2\left(1 - \rho_1 + n\left(\rho_1 - \rho_2\right)\right)}$$
$$\rho_1 = \text{correlation in the same style}$$
$$\rho_2 = \text{correlation between styles}$$

This model of style investing has a number of empirical predictions. According to this model, investors do not distinguish between stocks within a style. It may appear that fundamentally unrelated stocks are grouped into the same category, which leads to demand shocks across all assets in that style. The demand shock across all assets leads to a comovement in prices even if this is unrelated to the underlying fundamentals. This has consequences for the return correlation between assets in the same style and the return correlation between assets in different styles. When a style becomes popular, the return correlation between stocks in the same style will increase. In addition, fund inflow by one style drives resources out of competing styles, which leads to negative correlations among styles. Furthermore, the presence of style switchers leads to positively autocorrelated returns in the short run and negatively autocorrelated returns in the long run within the style. Good performance over the last period relative to other styles pushes the prices up again in the next period, inducing positive autocorrelation. This causes momentum in the short run. Eventually, the price is reversed in the long run, inducing negative autocorrelation. This leads to contrarian effects in the long run.

2.6 Summary and motivation for following chapters

Financial researchers are in the midst of the debate whether investors act rational and consider all available information in the decision-making process. In this respect, it is important to know the drivers behind stock valuation, in order to assess whether the anomalies outlined in section 2.2 result from efficient pricing of risk or from behavioral biases, such as the ones analyzed in behavioral finance. Behavioral finance argues that a plausible reason for the anomalies found is that agents are not fully rational. This limited-rational behavior generates two complementary sources for financial market inefficiency, notably: *investor sentiment*, which means that investors deviate from the maxims of economic rationality; and *limits of arbitrage* which argues that the arbitrageurs are not able or willing (due to their risk aversion) to digest the large demand shocks by noise traders.

While the patterns of aggregate stock market prices are not easy to understand from the rational point of view, different rational models have nonetheless been developed and can be tested against behavioral alternatives. In this chapter we discuss three rational and three behavioral models. The basic principal of rational models is that investors' decisions are made from a return-risk point of view. Fama and French (1993) suggest that risk can be analyzed in terms of three factors: the market premium, size and book-to-market. The three-factor model has been criticized by Daniel and Titman (1997) who cast doubt on the prediction that value stocks earn higher returns because such stocks have higher loadings on the book-to-market factor and not because they have high book-to-market ratios. In addition, the three factor model fails to explain long-term effects and momentum returns. Berk *et al.* (1996) develop non-linear models where returns are explained in terms of the option to expand. This model shows through simulation, that it can explain the value premium, the size-effect and the momentum effect. However, this model cannot reproduce momentum and contrarian effects at the horizons as found in empirical research. Furthermore, this model is difficult to test empirically. Zhang (2000) develops a model where the value of a firm is explained by the option to abandon and the option to expand. This model fails to explain anomalies.

Although these models assume that investors are rational, directly testing the validity of the assumption of rationality has not been general practice in economics. The rational school believes that not all participants are required to be rational in order to develop models that can explain stock returns. However, each of the three rational models described in section 2.4

fails to explain the momentum and contrarian effect such as is reported in empirical studies. Therefore, it may be important to start to investigate the validity of rationality. In the meantime, we should be skeptical about models on rational behavior that have not been further documented empirically.

The behavioral models assume that investors are exposed to biases from cognitive psychology, which influence their decision process. The models by Barberis *et al.* (1998) and Daniel *et al.* (1998) both focus on over- and underreaction to information as a source of momentum in stock returns. Momentum may also be caused by investors who follow trends. Prices go up and down without news about fundamentals. Such positive autocorrelation in prices that cannot be explained by under- or overreaction to new information, is described by the model of Barberis and Shleifer (2003). In this model, momentum is simply a reflection of style level phenomena.

Some critics argue that behavioral explanations for the empirical findings are (obviously) competing. For example, the models described in section 2.5 use different heuristics and biases from the cognitive psychology to explain the same phenomenon, notably momentum in the short run and mean-reversion in the long run. The model by Barberis *et al.* (1998) assumes that investors are exposed to representativeness and anchoring and adjustment while the model by Daniel *et al.* (1998) assumes that investors are exposed to overconfidence and self-attribution. In addition, Barberis and Shleifer (2003) emphasize the role of representativeness to explain momentum and contrarian effects. In addition, the models described in section 2.5 are only able to explain those anomalies that they are designed for. For example, the model by Barberis *et al.* fails to explain price drifts after isolated information events such as stock splits and dividends. The models by Barberis *et al.* and by Daniel *et al.* both fail to explain momentum due to investors who are chasing a trend instead of momentum as a result of the investors' reaction to new information.

Although the price pattern that is described in each model is the same, the explanations are different. In Daniel *et al.* the momentum phase features overreaction, as investors overreact to public signals that confirm their private information. The model of Barberis *et al.* (1998) works differently. Barberis *et al.* (1998) argue that the momentum phase reflects underreaction, as investors slowly react to new information. Barberis and Shleifer (2003) explain momentum as a result of investors who base their decisions on the style's past performance.

Fama (1998) reviews the studies of Barberis *et al.* (1998) and Daniel *et al.* (1998) and provide his view of under– and overreaction. He argues that underreaction appears as often as overreaction and that these instances cancel each other out. Therefore, market efficiency still remains. According to this view, market efficiency implies that prices coincide with fundamentals on average, but may deviate from each other due to chance. To explain the over– and underreaction evidence as a result of chance, Fama ignores that the circumstances where investors underreact differ from the circumstances where investors overreact. Notably, investors underreact to short-term information and overreact to long-term information.

The aim of this thesis is to contribute to the discussion between rationalists and behaviorists. In this chapter, we have described several models that are developed to explain anomalies. One way to compare the behavioral models is with empirical tests. The theoretical models are based on explicit and implicit assumptions about how beliefs of investors are formed. Hence, shedding some more light on these assumptions and testing the validity of these assumptions in practice is a prerequisite for evaluating these models. Based on the assessment of behavioral models, there are several directions deserving further research in order to obtain better insights in the decision-making process of investors:

1. Value premium
   With respect to the value premium two possibilities seem worth to investigate: first, the error-in-expectation hypothesis, and second, to examine the drivers behind investors' uncertainty. So far, both hypotheses have been explored by several empirical researchers in order to explain the value premium. None of these empirical studies have used a different approach for the classification of value and growth stocks. This line of research will be followed in the chapters 3 and 4.
2. Style investing
   Barberis and Shleifer (2003) create a model that is based on a demand-driven process. Stock returns are determined by investors who base their asset choice on a group level instead of an individual stock level. The investment process is in terms of investment cycles where the demand of a particular style is based on heuristics from the cognitive psychology. This model has been tested by several researches using in- and outflows of mutual funds and portfolios of individual investors. So far, social effects such as collective preferences have been neglected in these studies. We use different variables that reflect collective preferences of investors and investors' sentiment over time to test the popularity issue. We will pursue this direction in chapter 5.

Because of the competing behavioral explanations for some of the empirical facts, we believe that the assumptions of behavioral models need empirical scrutiny. In what follows we will therefore concentrate largely on the drivers behind behavioral models.