



Lappeenranta-Lahti University of Technology

School of Business and Management

Strategic Finance and Business Analytics

**The impact of holding period length, composite value measures, and firm size on value  
strategy performance in the Finnish stock markets**

**Master's thesis, 2020**

Author: Samuli Parhamaa

1<sup>st</sup> supervisor: Eero Pätäri

2<sup>nd</sup> supervisor: Timo Leivo

## ABSTRACT

**Author:** Samuli Parhamaa  
**Title:** The impact of holding period length, composite value measures, and firm size on value strategy performance in the Finnish stock markets  
**Faculty:** LUT, School of Business and Management  
**Master's program:** Strategic Finance and Business Analytics  
**Year:** 2020  
**Master's thesis:** Lappeenranta University of Technology  
71 pages, 7 tables, 5 figures and 7 appendices  
**Examiners:** Professor Eero Pätäri and Post-Doctoral Researcher Timo Leivo  
**Keywords:** value anomaly, value strategy, stocks, stock market, Finnish stock market, size effect, size anomaly, anomaly

This thesis examines the impact of holding period length, composite value measures, and firm size on value strategy performance in the Finnish stock markets during 1996-2020. The valuation ratios utilized in portfolio formation are E/P, B/P, D/P, FCF/P, and EBITDA/EV. In addition, a total of 10 composite value measures are selected, of which five are two-composite value measure portfolios, and the remaining five are three-composite value measure portfolios. Transaction costs and taxation procedures were omitted from the analysis, which would affect the resulting outcome in a real-life scenario.

With the support of existing literature, this study finds evidence of value anomaly. The value portfolios outperform the market and comparable growth portfolios almost without exception and with statistical significance. The best long-term performance is mainly obtained with the individual FCF/P value portfolio in terms of raw returns and most risk-adjusted metrics. The results exhibit some enhanced results with composite value measures when examining bull and bear markets separately. Generally, the combinatory portfolios are most capable of amplifying the value premiums of the individual ratios. Extending the holding period length gradually deteriorated the value strategy performance against the market on average. However, in some specific portfolios, slight enhancements could be obtained with extended holding periods. While some academics have reported that value anomaly is another form of size anomaly, this study found no evidence of firm size being a significant factor in explaining value portfolio returns.

## TIIVISTELMÄ

<b>Tekijä:</b>	Samuli Parhamaa
<b>Otsikko:</b>	Pitoajan, tunnuslukukombinaatioiden ja yrityskoon vaikutus arvostrategian suoriutumiseen Suomen osakemarkkinoilla
<b>Tiedekunta:</b>	LUT, School of Business and Management
<b>Maisteriohjelma:</b>	Strategic Finance and Business Analytics
<b>Vuosi:</b>	2020
<b>Pro gradu -tutkielma:</b>	Lappeenranta University of Technology 71 sivua, 7 taulukkoa, 5 kuviota and 7 liitettä
<b>Tarkastajat:</b>	Professori Eero Pätäri ja Tutkijatohtori Timo Leivo
<b>Hakusanat:</b>	arvoanomalia, arvostrategia, osakkeet, osakemarkkinat, Suomen osakemarkkinat, kokoanomalia, anomalia

Tässä tutkimuksessa tarkastellaan portfolion pitoajan, tunnuslukukombinaatioiden, ja yrityskoon vaikutusta arvostrategian suoriutumiseen Suomen osakemarkkinoilla vuosina 1996-2020. Portfolioiden muodostuksessa käytetyt tunnusluvut ovat E/P, B/P, D/P, FCF/P ja EBITDA/EV. Lisäksi tunnusluvuista muodostetaan yhteensä 10 kombinaatioportfoliota, joista viisi ovat kahden tunnusluvun yhdistelmäportfolioita ja loput viisi ovat kolmen tunnusluvun yhdistelmäportfolioita. Tarkastelussa ei oteta huomioon kaupankäyntikuluja eikä verotusta, joilla todellisuudessa olisi lopputuleman kannalta vaikutusta.

Olemassa olevan kirjallisuuden tukemana myös tämän tutkimuksen tulokset puoltavat arvoanomalian olemassaoloa. Arvoportfoliot ylisuoriutuvat markkinaan ja kasvuportfolioihin nähden lähes poikkeuksetta ja tilastollisesti merkitsevästi. Merkittävimmät pitkän aikavälin tuotot saavutetaan pääasiassa FCF/P arvoportfoliolla vuosituotoilla ja riskikorjatusti mitattuna. Tarkasteltaessa nousu- ja laskukausia erikseen, tuottoja pystytään joissain määrin parantamaan tunnuslukukombinaatioilla. Yleisesti ottaen suurin osa tunnuslukukombinaatioista pystyy kasvattamaan yksittäisten tunnuslukujen arvopreemioita. Portfolioiden pitoaikaa pidentämällä arvoportfolioiden keskimääräinen suoriutuminen markkinoihin nähden asteittaisesti laski, mutta tiettyjen yksittäisten portfolioiden tapauksessa suoriutumista pystyttiin kuitenkin hieman parantamaan pidemmällä pitoajoilla. Osassa akateemista kirjallisuutta arvoanomalian on väitetty olevan pienyhtiöanomalian ilmenemismuoto, mutta näiden tutkimustulosten pohjalta yhtiökoko ei ole merkitsevä muuttuja selittämään arvoportfolioiden tuottoja.

# Table of Contents

<b>1. Introduction.....</b>	<b>1</b>
1.1 Literature review .....	2
1.1.1 Value Anomaly .....	2
1.1.2 Size Anomaly .....	3
1.2 Objectives and research questions.....	5
1.3 Limitations .....	6
1.4 Study structure .....	6
<b>2. Theoretical Framework .....</b>	<b>7</b>
2.1 Efficient market hypothesis.....	7
2.2 Capital Asset Pricing Model (CAPM).....	9
2.3 Value anomaly .....	10
2.3.1 Background of value investing .....	11
2.3.2 Explanations of value anomaly .....	11
2.4 Causes of size anomaly.....	12
2.5 Valuation ratios .....	14
2.5.1 E/P ratio.....	14
2.5.2 B/P ratio.....	15
2.5.3 D/P ratio .....	17
2.5.4 FCF/P ratio.....	18
2.5.5 EBITDA/EV ratio.....	20
2.5.6 Composite value measures .....	21
<b>3. Data and Methodology .....</b>	<b>23</b>
3.1 Data.....	23
3.2 Portfolios .....	23
3.3 Sharpe ratio .....	27
3.4 SKASR (Skewness and Kurtosis Adjusted Sharpe Ratio).....	28
3.5 Three-factor model.....	29
3.6 Statistical testing .....	30
3.6.1 Jobson-Korkie-Memmel Z-test.....	31
3.6.2 Welch's t-test .....	31
<b>4. Results.....</b>	<b>33</b>
4.1 Performance of Individual and composite value measure portfolios.....	33
4.1.1 Full sample period .....	33
4.1.2 Bull market periods .....	39
4.1.3 Bear market periods.....	44
4.2 Performance of the value strategy with extended holding periods .....	48
4.3 Size effect.....	55
<b>5. Conclusions.....</b>	<b>58</b>
5.1 Summary.....	58
5.2 Concluding remarks .....	61
<b>References.....</b>	<b>62</b>

## Appendices

Appendix 1. Results of the Fama-French 3-factor model. ....	72
Appendix 2. All one-year holding period quartile portfolios during the full sample period (1996-2020) – Return, risk, and performance metrics. ....	73
Appendix 3. All two-year holding period quartile portfolios during the full sample period (1996-2020) – Return, risk, and performance metrics. ....	76
Appendix 4. All three-year holding period quartile portfolios during the full sample period (1996-2020) – Return, risk, and performance metrics. ....	79
Appendix 5. All four-year holding period quartile portfolios during the full sample period (1996-2020) – Return, risk, and performance metrics. ....	82
Appendix 6. All one-year holding period quartile portfolios during bull market periods – Return, risk, and performance metrics. ....	85
Appendix 7. All one-year holding period quartile portfolios during bear market periods – Return, risk, and performance metrics. ....	88

## List of Tables

Table 1. Correlation between the returns of value portfolios based on individual valuation ratios. ....	26
Table 2. One-year holding period - return, risk, and performance metrics of value and growth portfolios during the full sample period (1996-2020). ....	37
Table 3. One-year holding period - return, risk, and performance metrics of value and growth portfolios during bull market periods. ....	41
Table 4. One-year holding period - return, risk, and performance metrics of value and growth portfolios during bear market periods. ....	46
Table 5. Performance differences of an average value portfolio against the market and an average growth portfolio with one-, two-, three- and four-year holding periods during <i>the full sample period</i> (1996-2020). ....	48
Table 6. One- to four-year holding periods - return, risk, and performance metrics of all value portfolios during the full sample period (1996-2020). ....	51

Table 7. Three-factor versus two-factor alphas for measuring size effect in value & growth portfolios. ....	56
---	----

## List of Figures

Figure 1. Comparison of compounding excess returns of different market indices. ....	25
Figure 2. Compounding returns of average quartile portfolios of one-year holding period during the full sample period (1996-2020). ....	34
Figure 3. Relationship between risk and return during the full sample period (1996-2020). ....	35
Figure 4. Relationship between risk and return during the bull market periods. ....	40
Figure 5. Relationship between risk and return during bear market periods. ....	44

## List of Abbreviations

AMEX	American Stock Exchange
B/P	Book-to-Price Ratio
CAPM	Capital Asset Pricing Model
CF/P	Cash Flow-to-Price Ratio
D/P	Dividend-to-Price Ratio
EBITDA/EV	Earnings Before Interest, Taxes, Depreciation, and Amortization-to-Enterprise Value Ratio
EBIT/EV	Earnings Before Interest and Taxes-to-Enterprise Value Ratio
EMH	Efficient Market Hypothesis
E/P	Earnings-to-Price Ratio
EV	Enterprise Value
EWMP	Equally Weighted Market Portfolio
FCF/P	Free Cash Flow-to-Price Ratio
FF3	Fama and French Three-Factor Model
HML	High-Minus-Low Factor
ME/BE	Market Value of Equity-to-Book Value of Equity Ratio
NASDAQ	National Association of Securities Dealers Automated Quotations
NYSE	New York Stock Exchange
OMXHCAPGI	OMX Helsinki Weight Limited Gross Market Index
OMXHGI	OMX Helsinki Gross Market Index
P/E	Price-to-Earnings Ratio
P/B	Price-to-Book Ratio

Q1	Value Portfolio	
Q2	2nd Quartile Portfolio	
Q3	3rd Quartile Portfolio	
Q4	Growth Portfolio	
QMJ	Quality-Minus-Junk Factor	
SKAD	Skewness and Kurtosis Adjusted Deviation	
SKASR	Skewness and Kurtosis Adjusted Sharpe Ratio	
SMB	Small-Minus-Big Factor	
S/P	Sales-to-Price Ratio	
2A	E/P & B/P	(2-Composite Value Measure Portfolio)
2B	B/P & EBITDA/EV	(2-Composite Value Measure Portfolio)
2C	D/P & FCF/P	(2-Composite Value Measure Portfolio)
2D	D/P & EBITDA/EV	(2-Composite Value Measure Portfolio)
2E	FCF/P & EBITDA/EV	(2-Composite Value Measure Portfolio)
3A	E/P, B/P & D/P	(3-Composite Value Measure Portfolio)
3B	B/P, D/P & FCF/P	(3-Composite Value Measure Portfolio)
3C	B/P, D/P & EBITDA/EV	(3-Composite Value Measure Portfolio)
3D	B/P, FCF/P & EBITDA/EV	(3-Composite Value Measure Portfolio)
3E	D/P, FCF/P & EBITDA/EV	(3-Composite Value Measure Portfolio)

## 1. Introduction

A long-persisted paradigm in the financial world has been the assumption of efficient markets, which refers to investors' inability to obtain greater returns than the market offers (Fama, 1970). Regardless, investors have challenged this paradigm by seeking abnormal returns with different investing strategies across time. To reach this goal, they aim to exploit certain identifiable market inefficiencies - anomalies - to their advantage. Any successful exploitation would be considered a violation of efficient markets. Nevertheless, numerous anomaly related studies have shown evidence against it, setting the functionality of the efficient markets to questionable light.

This study aims to find evidence of value anomaly and size effect in the Finnish stock markets, which both are well-known market inefficiencies with studies from multiple decades globally. The origins of value investing go back to the thirties when Graham and Dodd (1934) published the book *Security Analysis*, a cornerstone for value investors, and an inspiration for countless academic studies. In their book, Graham and Dodd provide methods to detect and act on stock pricing errors caused by investors' tendency to extrapolate firms' financial performance too far into the future. The temporal financial distress of these firms would lead to plummeting stock prices. Thus a well-informed investor could buy the stocks cheap and wait until the pricing errors correct, expecting excess returns. Although the method was already invented in the thirties, one of the first studies attempting to find abnormal returns with cheaply priced firms were exhibited decades later by Nicholson (1960). He found that the returns of firms with low price-to-earnings ratios (P/E) tend to perform better than firms with high P/E ratios. Since then, value anomaly has gained plenty of advocates with studies extended to multiple decades and different global markets (e.g., Basu, 1977; Fama and French, 1992; 1998; 2012; Lakonishok, Shleifer & Vishny, 1994).

The size effect was captured by Banz (1981). He found that, on average, firms with low market capitalization generated better risk-adjusted returns than firms with large market capitalization. Using the term "size anomaly" refers explicitly to small-capitalization firms' excess returns over large capitalization firms. Academic literature from the recent decades have presented mixed views on the anomaly's existence: Studies from the eighties have mostly concluded that size

anomaly is evident (Reinganum, 1981; Banz & Breen, 1986), while studies from the late '90s and early 2000s have shown that the anomaly has disappeared or inverted (Dimson & Marsh 1999; Chan, Karceski & Lakonishok, 2000).

## **1.1 Literature review**

Deviations from market efficiency have been shown to occur frequently, and the intensity varies across time and markets. Next, we will discuss how the value strategy has generated excess returns against corresponding growth portfolios with robust historical evidence. Moreover, it is discussed how the size anomaly has fluctuated from one extreme to another, causing vigorous debate among researchers.

### **1.1.1 Value Anomaly**

Presumably, one of the first studies concerning value anomaly was published by Nicholson (1960). The study examined U.S. stocks and a sample period during and after the second world war suggesting low P/E firms generating higher returns than high P/E firms measured with raw returns. Basu (1977) concurs with Nicholson showing similar results: High E/P ratio firms generated larger risk-adjusted returns than low E/P firms. Fama and French (1992) studied value premium in the U.S. stock markets with a broader sample, including three American stock exchanges - NYSE, AMEX, and NASDAQ – using decile portfolios in the evaluations. They suggested that during 1963-1990 firms with high ME/BE (market value of equity-to-book value of equity) ratios produced higher returns than firms with low ME/BE ratios and having them also explaining a significant part of the cross-section of average stock returns. A later study of Lakonishok et al. (1994) provided parallel implications strengthening the evidence of the value premium in U.S. stocks. Fama and French (1998) expanded their former study to 13 different stock exchanges globally, including B/P, CF/P, E/P, and D/P ratios. Their findings showed consistent results in favor of firms with high ratios compared to low ratios: In 12 out of 13 markets, the value portfolios exceeded the growth portfolios on average. In a more recent study, Fama & French (2012) expanded their previous studies to the 21st century examining the stock markets of the U.S.,

Europe, Asia, and the Pacific region during 1989-2011. These findings did not differ from the previous, indicating perseverance of the value premium.

In the Finnish stock markets, evidence of the value anomaly has been displayed by Pätäri and Leivo (2009). They tested portfolios' performance with six individual valuation ratios (E/P, EBITDA/EV, CF/P, D/P, B/P, and S/P). Moreover, they introduced eight different composite value measures aiming to enhance the returns of single ratio portfolios. Their portfolio performance tests implied that the existence of the value premium is evident during 1993-2008. The majority of value portfolios outperformed both the market and comparable growth/glamour portfolios, indicating strong evidence in favor of value stocks. The composite value measure portfolios showed better performance to a certain degree when measured with risk-adjusted metrics. (Pätäri & Leivo, 2009) During the same year, Leivo, Pätäri, and Kilpiä (2009) published a similar study using a different period of 1991-2006 with reinforcing results. The same year Leivo and Pätäri (2009) expanded their previous research, including extended holding periods up to 5 years. The authors showed implications of performance enhancements with longer holding periods on some of the portfolios. Thus, the implications of enhanced results with composite value measures and extended holding periods give motivation for this study to see if similar results could be found.

### **1.1.2 Size Anomaly**

While the value anomaly has gained plenty of supporting evidence on its dispositional perseverance across time and different stock markets, researches have not been unanimous with the size effect. The anomaly tends to vary depending on the research period and market: Some researchers have exhibited clear indications that firms with small market capitalization can consistently generate higher returns than large-capitalization firms. Others have suggested the anomaly's reversal by supporting the underperformance of small firms. Moreover, some researchers have argued that the size effect or its reversal cannot be captured at all.

Banz (1981) belongs presumably to one of the first advocates of the size effect. He wanted to expand the view of cheaply priced stocks of value strategy by examining the impact of firms'

market value on the cross-section of average stock returns. The study used a 40-year period consisting of firms listed on the New York stock exchange. His main conclusion was that small-cap firms could generate greater risk-adjusted returns compared to large-cap firms on average during the entire sample period of 1936-1975. However, he remarked that after dividing the whole period into 10-year sub-periods, the anomaly's magnitude started to deviate substantially. Reinganum (1981) added AMEX to the sample in addition to NYSE. His findings supported Banz's evidence by suggesting that the decile with the smallest firms outperformed the largest decile by 1,77 % per month. Reinforcing results were also shown by Fama and French (1992), who examined three U.S stock exchanges - NYSE, AMEX, and NASDAQ. They reported that stock returns and firm size have a negative relationship together, implying that decreasing firm size would increase stock returns on average. The decile portfolios used in their study showed that the small-cap decile portfolio performed 0,63 % better per month than the largest decile. Signs of the anomaly were found from the United Kingdom stock markets in the '80s (Dimson & Marsh 1986). However, the paradigm began to shift to the opposite at the end of the 20<sup>th</sup> century.

Dimson and Marsh (1999) inverted their view on the size anomaly in the U.K. markets in the late '90s. They found that the anomaly began to shift during the '80s and had disappeared, or rather switched to the opposite during the '90s: They showed that large-cap firms had started to generate larger returns than small-cap firms. A year later, Chan et al. (2000) obtained similar results from the U.S. stock markets, setting the size effect's existence into a questionable light.

Van Dijk (2011) critiqued the studies that advocated the size effect's disappearance by stating that the size premium is still positive in the U.S. markets. However, the author points out that more empirical research is needed to make conclusions about the robustness of the anomaly's existence. Moreover, he remarked that the theories around the anomaly have not been developed adequately and have not been tested systematically.

In a more recent study, Asness, Frazzini, Israel, Moskowitz, and Pedersen (2018) suggested the reappearance of the anomaly on a global level when firm quality-related determinants were accounted for in their evaluation. The main results were obtained by using Fama and French's (1992) three-factor model (FF3) combined with Quality-Minus-Junk factor (QMJ). The sample

period used was extremely long, beginning from 1929 until 2012, and the sample consisted of 24 different markets across the world. Their main finding was that when the QMJ factor - the difference between "quality" firms and "junk" firms – was used together with the FF3, the size anomaly could be detected from 23 out of the 24 markets systematically from the entire period and sub-periods.

Pätäri and Leivo (2009) aimed to explain the outperformance of value stocks with size effect in the Finnish stock market by using a two-factor model, which is an extended version of Capital Asset Pricing Model (CAPM) introduced by Sharpe (1964), Lintner (1965). Alongside market risk, the 2-factor alpha includes SMB-factor (Small-Minus-Big), which captures the return differences between small-cap and large-cap firms. However, their main conclusion was that firm size could not explain the value premium in the Finnish stock markets.

## **1.2 Objectives and research questions**

This thesis aims to find evidence on both value and size anomaly in the Helsinki stock exchange during 1996-2020. The research questions are formulated as follows:

1. *Do value portfolios outperform the market and growth portfolios consistently?*
2. *Do composite value measures enhance the performance of the value portfolios?*
3. *Do extended holding periods enhance the performance of the value portfolios?*
4. *Does size anomaly explain the returns of value portfolios?*

The Finnish stock market is a particularly interesting research subject due to its abnormal behavior in comparison with major global markets. Pätäri and Leivo (2009) suggest the following: "It suffers from intermittent 'periphery syndrome' caused by the herding behavior of international institutional investors who cash their equity positions first from the furthest stock markets during the turbulent times." The authors also explain that the liquidity of the Finnish stocks is relatively low, which leads to drops larger than more developed markets and possibly larger pricing errors. Moreover, the trading volume during bull markets is also relatively low,

causing higher volatility. These determinants combined would potentially drive more pricing errors and more opportunities for investors to obtain abnormal returns. (Pätäri & Leivo, 2009)

### **1.3 Limitations**

The study concentrates strictly on firms listed on the Helsinki stock exchange. All the stocks quoted on both the main list and the First North list are included. The time period is restricted to the last 24 years of stock data in order to fit 1-, 2-, 3- and 4-year holding periods equally. If a firm has more than one series of shares, only the one with the highest liquidity is included. Lastly, this thesis does not account for any transaction costs, and therefore the reader must be aware that in a real-life scenario, the returns would potentially be slightly lower.

### **1.4 Study structure**

In the first chapter, we overviewed the main subjects of this study – value and size anomaly. In chapter 2, the theories relevant to this study are introduced and addressed in more depth. Chapter 3 addresses the data and the methodologies used in the research process. Chapter 4 presents the overall results to obtain answers to the research questions and take a more in-depth look at the portfolio performances during the entire sample period and bull and bear markets. Chapter 5 is the final chapter, where the conclusions are presented.

## 2. Theoretical Framework

This section discusses the financial theories that aim to explain investor behavior and price formation in the financial markets. Most financial theories assume that stock markets are efficient, and the behavior of market participants is entirely rational. However, in practice, these assumptions seem not always to hold, and therefore market inefficiencies may lead to anomaly-based strategies to succeed (Schwert, 2002).

### 2.1 Efficient market hypothesis

One of the essential theories within the financial domain is the efficient market hypothesis (EMH), first introduced by Fama (1970). EMH states that investors cannot make excess returns in relation to the market since all available information is reflected in asset prices as soon as new information arises. In order to have perfectly efficient markets, EMH assumes that there are no transaction costs, information is free and easy to obtain, investors' expectations are homogenous, and investors are fully rational (Findlay & Williams, 2000). The form of efficiency depends on the available information. Fama recognizes three forms of efficient markets: 1) weak form, 2) semi-strong form, and 3) strong form efficiency (Fama, 1970).

*The weak form* states that the only information available is the historical asset prices (Fama, 1970). This form is closely related to the random-walk theory, which suggests that asset prices follow random patterns. Hence, if random-walk is present, investors could not predict future prices based on historical returns, and thus making excess returns should be impossible. (Malkiel, 2003)

With *the semi-strong form*, asset prices reflect all publicly available information. This form closely follows a real-world scenario, where investors use news and companies' financial statements to determine asset prices. As soon as new public information arises, investors take this into account by pricing the new information immediately. (Fama, 1970)

If markets followed *the strong-form*, the asset prices would reflect all information. Thus, in addition to public information, all private information would also be reflected in the prices immediately. (Fama, 1970)

Since EMH assumes perfect market efficiency, pricing errors should not occur (Fama, 1970). If this assumption holds, investors should prefer not to try to “beat the market” but to be content with a low-cost market index. In reality, investors are not fully rational, their expectations do not align perfectly, and different transaction and information costs occur. Therefore, opportunities arise due to pricing errors, which do not correct immediately, but rather stay present long enough for investors to exploit. (Malkiel, 2003) The studies addressed in the literature review (chapter 1.1) make the hypothesis of efficient markets questionable since they show that errors in the market prices can prevail even after the publications about market errors arise.

Value and size anomalies are not the only inefficiencies that researchers have detected in recent decades. For example, Jegadeesh and Titman (1993) found that stocks with excellent past performance tend to continue their positive trend in the future, and poorly performed stocks seem to continue their poor performance. This phenomenon is called momentum anomaly, where the stock prices tend to follow their recent trend. Thus, predictions of future prices could be made, which already challenges the weak-form market efficiency, which assumes that stock prices follow random patterns. An extreme example of a market failure would be the Dot-Com bubble at the end of the '90s and the early 2000s. Cooper, Dimitrov, and Rau (2001) found that firms that added an internet-related term to their company's name soared on the order of 53 % on average in the next five days after the name announcement.

Due to the apparent evidence on anomalies and the critique towards the original EMH, Fama (1991) corrected his view on the efficient markets. He states that such information and transaction costs do exist that would cancel the assumption of truly efficient markets. He concludes that although the markets cannot be considered truly efficient, the original hypothesis can be utilized as a baseline for other theories. (Fama, 1991)

## 2.2 Capital Asset Pricing Model (CAPM)

Highly influenced by the modern portfolio theory of Harry Markowitz (1952), the capital asset pricing model (CAPM) of Sharpe (1964) and Lintner (1965) is one of the most fundamental tools in finance used to describe the linear relationships between systematic risk and expected asset returns. The simplicity of the model is one reason for its wide use among asset managers and in the financial literature. The CAPM is built upon specific prerequisites: It assumes that investors are risk-averse and seek to maximize returns, no transaction costs or taxations exist, they view securities with homogenous expectations of the return probability distribution, and all investors can borrow and lend at the risk-free rate of interest (Black, Jensen & Scholes, 1972). Its basic form aims to explain an asset's price movement with a single explanatory variable, market returns (minus risk-free rate). The CAPM equation is formulated in the following way:

$$E(r_i) = r_f + \beta_i[E(r_m) - r_f] \quad (1)$$

where

$E(r_i)$  = expected return of an asset,

$r_f$  = risk-free rate,

$\beta_i$  = asset beta estimate, which measures systematic risk and

$E(r_m)$  = expected return of the market.

The beta coefficient is calculated by dividing the covariance of the asset and market returns with the market variance:

$$\beta = \frac{Cov(r_i, r_m)}{Var(r_m)} \quad (2)$$

The beta estimate of the asset simply describes the asset's price behavior in relation to a relevant market index. The CAPM can also be translated into a linear regression format, which produces the same result for the beta-estimate. The beta's interpretation is intuitive: For example, if the beta of an asset is 2 and the market moves up by 1 %, the asset price is expected to rise by 2 % on average. Vice versa, a 1 % decline in the market would result in a 2 % decline in the asset price on average.

However, the theory is claimed to have weaknesses, thus causing vigorous debate among academics. According to Fama and French (2004), the model's explanatory power is too weak to provide reliable results due to too many simplifying assumptions. Thus the market itself is not sufficient enough to explain the cross-section of asset returns. Black, Jensen, and Scholes (1972) argue that the CAPM beta estimates are not consistent. The authors explain that the expected excess returns of an asset are not strictly proportional to its beta. While examining the size effect, Banz (1981) found that the CAPM is misspecified by showing that the size effect is not linear in the market value. Roll (1977) argues that the CAPM theory is not testable without a comprehensive index that includes information of every individual asset class. He adds that choosing a misspecified market index causes drastic changes in the beta estimate.

To counter the original CAPM's insufficiencies, academics have created several extensions to increase the explanatory power by adding several new factors to the equation. This thesis utilizes the three-factor model (FF3) formulated by Fama and French (1992). It introduces two new explanatory variables: HML- and SMB-factor. The HML factor stands for high-minus-low, which is the difference in high book-to-market stock returns versus low book-to-market stock returns. The SMB factor stands for small-minus-big, which is the difference between the returns of small-cap and large-cap firms. Fama and French found that these two factors seem to reflect the unknown key variables left by the CAPM and show high explanatory power over the cross-section of stock returns. (Fama & French, 1992) Their later studies have demonstrated the model's apparent success in the U.S. stock markets (Fama & French, 1993; 1996). Although international evidence of the model's explanatory power is mostly mixed, academics have shown that the model is, regardless, a substantial improvement from the original CAPM (Harvey & Siddique, 2000; Connor & Sehgal, 2001; Fama & French, 1998).

### **2.3 Value anomaly**

As discussed in chapter 1.1.1, value anomaly is a widely detected phenomenon. Next, we will further discuss the background of value investing and the reasons why value anomaly exists.

### **2.3.1 Background of value investing**

The origins of value investing go back to the '30s when Columbia University professors Benjamin Graham and David Dodd published *Security Analysis* (1934). The book mainly addresses the concept of purchasing equities cheaper than their intrinsic values in order to minimize the risks of losses. The intrinsic value refers to a thorough fundamental analysis used to value equities. Later, Graham published *Intelligent Investor* (1949), which mainly focuses on margin of safety – the difference between an equity's price and its intrinsic value. These two concepts – intrinsic value & margin of safety – encompass the fundamental idea behind value investing. The work of Graham and Dodd belongs presumably to one of the greatest influences in value anomaly related studies. The authors did not coin the terms value premium or value anomaly themselves. Instead, they are concepts used in academic financial literature (e.g., Ball, 1992; Zhang, 2005).

The important point is that academic studies mainly do not address the thorough fundamental analysis of individual firms. Instead, they compile stocks to quantile portfolios based on relative valuation metrics, for example, E/P, B/P, or CF/P. (e.g., Basu, 1977; Fama and French, 1992; Lakonishok et al., 1994). The valuation ratios utilized in this study are introduced in chapter 2.5.

### **2.3.2 Explanations of value anomaly**

Academics have found several explanations for value anomaly. There are currently at least three identified categories: 1) risk-based explanations, 2) irrational investor behavior, and 3) data snooping bias. (Pätäri & Leivo, 2017)

Concerning the *risk-based theories*, Fama and French (1995) argue that value premium is compensation for the risk that is otherwise missed by asset pricing models such as CAPM. Doukas, Kim, and Pantzalis (2004) suggested a similar explanation, stating that standard deviation measures in analysts' earnings-per-share forecasts compensate for the risk reflected by the abnormal returns of value stocks. Petkova and Zhang (2005) argue that value stocks tend to be more vulnerable to financial crises than growth stocks. It is suggested that relatively cheap stock valuation is apparent in firms that have more expensive production scaling during economic

cycles, and thus the inflexibility exposes value stocks to financial distress, leading to undervaluation (Zhang, 2005).

*Irrational investor behavior* was an essential part of value investing already in *Security Analysis* by Graham and Dodd (1934), who believed that excess returns could be obtained with mispriced securities. This led to creating the theory of overreaction hypothesis decades later (De Bondt & Thaler 1985), which is closely related to the extrapolation of historical data too far into the future. That is to say, the markets assume that a company's good or bad financial situation will continue in the future, and therefore this is reflected in the asset prices leading to possible stock mispricing. (Lakonishok et al., 1994; Fama & French, 1995) Conrad (1995) argues that the value premium is due to the undervaluation of distressed stocks and overvaluation of growth stocks. He concludes that when pricing errors correct, the undervalued stocks tend to generate higher returns than glamour stocks. The overreaction assumption is reinforced by Daniel, Hirshleifer, and Subrahmanyam (2001), who state that the overreaction is caused by overconfidence, and overreactions to private signals cause extreme B/P ratios. Bartov and Kim (2004) suggest that B/P anomaly is more significant in stocks mostly held by unsophisticated small investors with little analyst coverage than firms owned by major institutions with a large following of market participants.

Some academics argue that value anomaly exists due to *data snooping*, meaning that statistically significant patterns are purposely found through data manipulation or other forms of misuse (Black 1993; Conrad, Cooper & Kaul, 2003). However, considering the advocating evidence from the recent decades around the world, it is more than unlikely that all of the studies would come to similar conclusions due to biased data selection (Markowitz & Xu, 1994; Guerard Jr., Xu & Gultekin, 2012).

## **2.4 Causes of size anomaly**

The size effect is another example of markets' inefficiencies. However, as it is already shown in chapter 1.1.2, the anomaly has gained plenty of mixed evidence among academics. Nevertheless,

there are possible explanations for its occurrence. Van Dijk (2011) finds that size anomaly is caused by 1) risk, 2) investor behavior, and 3) liquidity.

Asness et al. (2018) explain that if size itself is not a *risk*, standard asset pricing models such as CAPM predict that size has no effect when controlling for risk exposures. It is suggested that size can correlate with expected returns merely because size is mainly measured by market value, which leads to size influencing time-varying risk premia. Thus, the required return is higher on riskier firms, leading to firm size being lower. (Berk, 1995) Garleanu, Panageas & Yu (2012) argue that one reason for the size effect is that if small firms have more risky growth options, these kinds of firms are relatively riskier than firms with less risky growth options, leading to lower market value. Vassalou and Xing (2004) show that there is a significant correlation between small firms and credit risk. Similar results have been found by Campbell, Hilscher, and Szilagyi (2008), which show that firms with a risk of bankruptcy have high explanatory power with the size factor.

Concerning *investor behavior*, Shleifer and Vishny (1997) state, it is more challenging to generate arbitrage profits with small firms, and therefore these firms are more likely to become mispriced. Barberis, Shleifer, and Vishny (2005) argue that some market participants overreact to the growth expectations of small firms, which also leads to mispricing. Merton (1987) explains that size anomaly exists due to insufficient information available from small firms and adds that less known firms include more expected returns. Potentially, the lack of information could be explained by lesser information demand compared to larger firms, and thus small firms require substantially more justification for financing (Lakonishok et al., 1994). Insufficient information is also related to delays in price changes: Hou and Moskowitz (2005) suggest that since obtaining information from small firms can be more difficult, delays in price changes may occur. The authors suggest that the price delays capture the majority of size anomaly.

Acharya and Pedersen (2005) explain that small firms are less *liquid* than large firms leading to more liquidity risk. Thus, these firms have larger expected returns, which consequently leads to lower market value. Most size-related studies do not account for transaction costs or bid-ask spread. It is critiqued that size anomaly is captured due to the omission of these costs since they have an apparent effect on returns. (Stoll & Whaley, 1983) However, while studying NYSE and

AMEX, Schultz (1983) found that small firms can generate excess returns for a short period, even after accounting for transaction costs.

## 2.5 Valuation ratios

Studies have examined value anomaly using different valuation ratios and showed both clear and mixed evidence regarding value anomaly depending on the ratio. Next, we will discuss the ratios utilized in this study and review their performance in previous studies.

### 2.5.1 E/P ratio

One of the most common valuation ratios used in financial literature is the E/P ratio. It compares the firm's recently reported yearly earnings and its current market value. A more common formula is to divide a firm's earnings per share with its current stock price, which returns the same answer. Therefore, E/P can be formulated with the following equation:

$$\frac{E}{P} = \frac{\text{Earnings per share}}{\text{Stock price}} \quad (3)$$

E/P is often referred to as earnings yield, interpreted as the rate of return that a firm generates to its shareholders. E/P is often presented as its inverse – P/E. An intuitive interpretation for P/E is to consider how many years it takes to generate earnings to the shareholders equal to the current market price. However, both approaches assume the earnings to stay constant, which usually does not match with reality.

One of the first reported occurrences of E/P anomaly was conducted by Nicholson (1960) and McWilliams (1966). Both used raw-returns in evaluation examining the U.S. markets, although omitting risk-adjusted performance metrics. Basu (1977) took risk-adjusted returns into the evaluation and showed that E/P anomaly can still be detected. Banz (1981) and Reinganum (1981) argued that E/P anomaly is another form of size anomaly. However, Basu (1983) responded that this is not the case since the E/P anomaly can be captured even after controlling for size. He adds that the size effect disappears after controlling returns for differences in risk and E/P ratios. On

the contrary, Banz and Breen (1986) agreed with Banz (1981) and Reinganum (1981), showing that E/P anomaly can only be detected when accompanied by size. Similar results were also found in Japanese markets by Chan, Hamao, and Lakonshok (1991). The authors reported that E/P anomaly was not found individually, but rather when controlling for size. Moreover, they suggested stronger evidence of individual B/P and CF/P anomaly. Again, the same conclusion is presented from the U.S. market by Fama and French (1992), reporting the relationship between E/P and size. On the contrary, Roll (1995) exhibited that in fact, E/P anomaly can be detected individually without size proxy. On a global scale, Fama and French (1998) found that the highest value premium with E/P was found in only 2/13 markets. However, in terms of overall returns, E/P returns were not the highest of any of the markets. Advocating views on the E/P anomaly have been presented in more recent studies by Athanassakos (2011) and Israel and Moskowitz (2013), who both suggested that the E/P anomaly can be detected in the U.S. market. The former used a sample period of 1985-2006, and the latter used a remarkably long period of 1926-2011.

E/P anomaly has also been detected in the Finnish stock markets. Leivo and Pätäri (2009) showed that E/P could overperform comparable growth portfolios almost without exception and mostly with statistical significance. Moreover, E/P value portfolios could exceed the market with statistical significance when the holding period is five years. Even though the Finnish evidence of E/P anomaly is apparent, some comparably more robust evidence is shown with D/P and EBITDA/EV. (Leivo & Pätäri, 2009)

### **2.5.2 B/P ratio**

B/P ratio is presumably the most examined valuation ratio within the anomaly literature. It measures a firm's book value of equity in relation to its current market value. Another formulation uses the book value of equity per share and divides it by the current stock price. Therefore, B/P can be formulated in the following way:

$$\frac{B}{P} = \frac{\text{Book value of equity per share}}{\text{Stock price}} \quad (4)$$

B/P ratio can be considered a relatively stable measure since companies' book values are not prone to drastic changes (Lindström, 2007, 201). B/P is sometimes considered a margin of safety measure giving an intuitive approach to a firm's valuation. However, a firm's book value does not always directly reflect a firm's realization value. (Bodie, Kane & Marcus, 2001, 618) This stems from differences between industries: Firms that operate in industries with a substantial amount of intangible assets may lead to relatively lower B/P ratios, and firms with high tangible assets may lead to higher B/P ratios. (Jones, 2007, 281) Thus, the interpretation is not always unanimous in different industries.

Presumably, one of the first to find the functionality of B/P was Stattman (1980), exhibiting significant B/P anomaly in the U.S. markets. However, it was later suggested that his results were biased due to survivorship and look-ahead bias caused by employed sample selection criteria (Pätäri & Leivo 2017). Despite Stattman's erroneous data processing, Rosenberg, Reid, and Lanstein (1985) found evidence in favor of B/P anomaly. They reported that high B/P stocks performed significantly better than low B/P stocks in the U.S. markets during 1973-1984. Later, similar results were reported in the Japanese market by Chan et al. (1991). The authors used E/P, B/P, CF/P, and size as portfolio selection criteria and found the most considerable difference in returns between B/P value and growth portfolios. Again, complementary evidence was found by Fama and French (1992), who showed that the most significant results were obtained with B/P in the U.S. markets during 1963-1990. On a global scale, Fama and French (1998) found that B/P had the largest premium in 6/13 markets compared to the other valuation ratios (E/P, CF/P & D/P).

Using size as a factor along with B/P has shown that the premium varies depending on firms' market values. Fama and French (2012), along with Israel and Moskowitz (2013), showed that there is a negative relationship between firm size and B/P premium. Thus they suggested that small-cap value firms tend to generate more excess returns than large-cap value firms.

In contrast to the evidence from the U.S. markets and some other global markets, B/P premium has shown no strong evidence in the Finnish markets. Pätäri and Leivo (2009) used tertile portfolios and found that B/P did not significantly overperform the growth portfolio or market. Moreover, it underperformed in relation to all other individual ratios (E/P, EBITDA/EV, CF/P &

D/P) during 1993-2008. Using extended holding periods with the same sample period, B/P underperformed the growth portfolio in some cases and remained in the lowest of the rankings in comparison with other individual ratios (Leivo & Pätäri, 2009). On the contrary, Leivo et al. (2009) used quintile portfolios and a period of 1991-2006. They found that B/P was among the best performing value portfolios in terms of annual returns. However, taking risk-adjustments into account in terms of Sharpe and Sortino ratio, the performance of B/P declined significantly due to relatively larger volatility.

### 2.5.3 D/P ratio

The D/P ratio can be translated as dividend yield. It displays the ratio between dividends per share and the current stock price:

$$\frac{D}{P} = \frac{\text{Dividends per share}}{\text{Stock price}} \quad (5)$$

A theoretical framework has been constructed around the dividend yield and its ability to predict returns. There are currently at least three proposed hypotheses: tax-effect hypothesis, dividend neutrality hypothesis, and signaling hypothesis. *The tax-effect hypothesis* states that investors receive higher risk-adjusted before-tax returns on stocks with higher expected dividend yields to be compensated for previously higher taxation of dividend income relative to capital gain income (Brennan, 1970). *The dividend-neutrality hypothesis* suggests that investors' higher return requirements for holding higher dividend-yield stocks would lead value-maximizing firms to adjust their dividend policy by decreasing paid dividends, lowering the cost of capital, and therefore increasing their stock price (Black & Scholes, 1974). *The signaling hypothesis* states that dividend yields and their changes reflect managements' beliefs concerning the future prospects of the firm, thus higher tendency for dividend payments can be interpreted as a signal of the managements' trust being able to pay future dividends (Dielman & Oppenheimer, 1984; Denis, Denis & Sarin, 1994; Sant & Cowan, 1994).

The prediction power of dividend yield on stock returns has gained mixed results among academics. Black and Scholes (1974) suggested no evidence on greater dividend yields generating significant returns. However, Litzenberger and Ramaswamy (1979) examined NYSE stocks and

questioned the methods of Black and Scholes by showing a strong positive relationship between D/P and expected returns. Reinforcing results advocating higher dividend yields were also reported by Elton, Gruber, and Rentzel (1983), Litzenberger and Ramaswamy (1982), Rozeff (1984) as well as Fama and French (1988). Disparate results were presented by Blume (1980) and Keim (1985), who found a U-shaped relationship between dividend yield and returns. They stated that stocks that had either high or zero dividend yield outperformed low dividend yield stocks. Similar results were reported by Naranjo, Nimalendran, and Ryngaert (1998) over a decade later, showing overperformance of zero-yield stocks over low dividend yield stocks.

Despite the existence of D/P premium by several studies, the performance of D/P has not been extremely sufficient in relation to other individual ratios: Fama and French's (1998) international study showed that D/P was able to exceed three other ratios in only one out of 13 major global markets. Naranjo et al. (1998) reported that the value premium exists between the highest and lowest decile portfolios. However, the third decile generated the highest returns. The lowest value premium of D/P in relation to other ratios in the U.S. markets was reported by Loughran and Wellman (2011), Israel and Moskowitz (2013), and Hou, Xue, and Zhang (2015). Studies outside the U.S. have exhibited more mixed results. Levis (1989) reported that D/P had the highest value premium in the U.K., whereas Miles and Timmermann (1996) reported a negative value premium in the same market. Bird and Whitaker (2003) combined seven different European markets (U.K., France, Germany, Italy, Switzerland, Netherlands, and Spain) and reported that a negative value premium is evident on a continental level.

In the Finnish markets, the dividend yield has been reported among the most robust ratios. Leivo and Pätäri (2009) exhibited statistical significance with D/P premium with all holding periods from one to five years and also the highest annualized returns among other individual valuation ratios.

#### **2.5.4 FCF/P ratio**

Reported earnings figures have caused skepticism due to differences in calculation practices of discretionary accruals, such as depreciations and amortizations, and changing calculation principles of earnings figures due to changing accounting standards (Chan, Chan, Jegadeesh &

Lakonishok, 2006; Callao & Jarne, 2010). This has sparked academics' motivation to examine the effect of generated cash flows on stock returns (Bernard & Stober, 1989; Chan et al., 1991). CF/P anomaly has shown competent results: Lakonishok et al. (1994) reported a significant value premium of 9,9 % and higher returns with CF/P than B/P or E/P in the U.S. markets. Davis (1994), as well as Desai, Rajgopal, and Venkatachalam (2004) exhibited the second-highest value premium with CF/P, also in the U.S. markets. Outside of U.S., Fama and French (1998) reported a significant and highest value premium in 4/12 countries with CF/P. Chan et al. (1991) reported the second-highest value premium and second-highest returns in the Japanese market.

CF/P has gained promising results also in the Finnish market. Leivo and Pätäri (2011) reported the highest value premium and value sextile portfolio returns among six individual valuation ratios. However, their previous study from 2009 showed a slightly weaker performance among individual ratios when tertile portfolios were used, showing underperformance compared to EBITDA/EV, E/P, and D/P in risk-adjusted terms.

A notable observation is that these studies have used different measures as the cash flow component. For example, Bernard & Stober (1989) uses only current accruals as the cash flow measure. Chan et al. (1991) determine cash flows as earnings plus depreciations. Fama and French (1998) do not specify their composition of cash flows, whereas Leivo and Pätäri (2009) determine CF/P ratio as the sum of fully diluted earnings per share excluding extraordinary items and depreciations and amortizations per share.

Moreover, the previously mentioned studies exclude capital expenditures and discretionary cash out-flows, which indicate the excess cash left after all required cash-out flows used to fund investments (Lehn & Poulsen, 1989). This form of cash flows can be translated to free cash flows (FCF). Hackel, Livnat, and Rai (2000) and a response study by Jokipii and Vähämaa (2006) took the capital expenditures and discretionary cash out-flows by management into account, showing promising results: Hackel et al. (2000) suggested superior returns with FCF/P in relation to returns of S&P 500 index, similar beta and size portfolios as well as similar book-to-market portfolios. Jokipii and Vähämaa (2006) also suggested FCF/P anomaly from the Finnish market showing

overperformance against the market, giving motivation to this study to see if FCF/P would generate similar results.

Therefore, the used cash flow-based valuation ratio is free cash flow per share divided by the current stock price:

$$\frac{FCF}{P} = \frac{\text{Free cash flow per share}}{\text{Stock price}} \quad (6)$$

where the free cash flow component consists of funds from operations deducted by capital expenditures and cash dividends paid.

### **2.5.5 EBITDA/EV ratio**

Enterprise value (EV) based valuation ratios, such as EBITDA/EV, EBIT/EV, have been relatively uncommon in anomaly literature until recently. They have had increasing popularity since the enterprise value component accounts for leverage effects, including the net value of a company's debt (Pätäri & Leivo, 2017). Enterprise value can be formulated as follows:

$$\text{Enterprise value (EV)} = \text{Firm market value} + \text{Net debt} - \text{Cash} \quad (7)$$

In this study, EBITDA/EV ratio is used:

$$\frac{EBITDA}{EV} = \frac{\text{Earnings before interest, taxes, depreciation, and amortization}}{\text{Enterprise value}} \quad (8)$$

A reason for the use of EBITDA instead of EBIT in the equation is that depreciation methods might differ between firms. Since EBITDA is unaffected by these differences, the comparison between firms can be more balanced (Pätäri & Leivo, 2017). It is, however, argued that EBIT could be a better measure: Chan and Lui (2011) suggest that EBIT gives better guidance on profit and sustainability for the future, and thus, EBIT/EV gives a more distinct view on a company's true profitability than EBITDA/EV. On the other hand, empirical evidence shows that EBITDA/EV would be a more competent measure. Gray and Vogel (2012) found that EBITDA/EV was the best performing portfolio among a total of 25 quintile portfolios. Reinforcing results have also been reported by Leivo et al. (2009) from the Finnish markets. They found that EBITDA/EV was the best

performing individual valuation ratio from a total of 20 quintile portfolios in terms of risk-adjusted metrics.

### ***2.5.6 Composite value measures***

The idea of composite value criteria is to improve portfolio performance by combining two or more fundamental valuation metrics. It is suggested that indicators with relatively low correlation can generate added value when combined together (Pätäri & Leivo, 2017). Assumably one of the first to utilize combinatory metrics in an investment strategy was Graham (1949) in his book *The Intelligent Investor*, which is also considered a benchmark in value investing along with *Security Analysis*. He introduced the "Graham's number", which combines P/E and P/B ratio. As a rule of thumb, he suggests firms with P/E lower than 15 and P/B lower than 1,5, or their multiplication lower than 22,5. Alongside the moderate ratios, he suggests other measures, including moderate enterprise size, sufficiently strong financial condition, earnings stability, dividend record, and earnings growth. (Graham, 2003, 348-349)

In the academic literature, researchers have found more or less evidence on performance improvement using composite value measures. Dhatt, Kim, and Mukherji (1999) were presumably one of the first to report enhanced composite portfolio returns. Their sample consisted of stocks from the U.S. small-cap index Russell 2000. The authors showed that the combinatory value portfolio of B/P, E/P, and D/P had a slight performance improvement of 1,56 % per annum over the best individual ratio E/P. Their later work compared 11 different combinatory portfolios of U.S. stocks, where a combination of S/P and E/P generated the highest value premium and highest value portfolio returns (Dhatt, Kim & Mukherji, 2004). Piotroski (2000) combined B/P with F-score. The F-score utilizes nine different profitability, liquidity, leverage, and operating efficiency metrics as binary scoring measures. He showed that, on average, portfolios with high B/P combined with high F-score outperformed individual high B/P portfolios by 7,5 % per year. However, the combinatory portfolio was substantially smaller on average, making the portfolio comparison unequal. More robust evidence was suggested by Chan and Lakonishok (2004), who reported the highest value premium of 21,1 % between extreme decile portfolios when

combining small-cap firms with high B/P, CF/P, E/P, and S/P ratios together. The difference between similar portfolios with large-cap firms showed an 11,9 % value premium. However, no comparison with different combinatory measures or individual ratios were reported.

Outside the U.S., van der Hart, deZwart, and van Dijk (2003) examined emerging markets with 684 stocks on average. Their best performing combinatory value portfolio of B/P and E/P showed merely a 0,2 % monthly value premium, and only a very slight improvement of 0,1 % per month over the best individual ratio E/P during 1985-1999. Chan and Lakonishok (2004) reported a relatively high value premium of 16,8 % between large stocks from the MSCI EAFE index of developed non-U.S. countries with B/P, CF/P, E/P, and S/P combination during 1989-2001, although without any comparison with different ratios. Bird and Casavecchia (2007) used a relatively large sample of 8000 stocks from 15 European countries total utilizing a combination of B/P-, E/P-, and S/P-quintiles and holding periods of 3, 6, 12, 24, and 36 months during 1989-2004. They concluded that no improvements could be made with the combinatory portfolios compared to individual ratio portfolios.

Certain composite value combinations have shown promising results from the Finnish stock market. Leivo et al. (2009) compared a total of seven different individual and composite value measures. They reported that the highest yearly returns of 32,25 % were generated using a 2-composite value measure based on B/P and EBITDA/EV during 1991-2006. During the same year, Leivo and Pätäri (2009) compared six individual ratios and three combination ratios with 1-5 year holding periods and a sample period of 1993-2008. A combination of D/P and EBITDA/EV performed the best during one-, two-, and four-year holding periods in annual return terms. A combination of B/P, D/P, and EBITDA/EV generated the highest returns during a four-year holding period.

### **3. Data and Methodology**

In this chapter, the data procedures and methodologies used for portfolio evaluation are addressed. First, the data is reviewed. Then, the portfolio formation procedures are inspected in detail. Lastly, the metrics for portfolio evaluation are introduced: Sharpe ratio, SKASR, FF3, and statistical significance.

#### **3.1 Data**

All data is gathered from Thomson Reuters Datastream. All the data and relevant procedures are processed in Microsoft Excel. The data consists of stocks from the Helsinki stock exchange, including all stocks from the main list and the Helsinki First North list. Monthly data and a sample period of 1996-2020 is used. The total sample length is 288 observations, of which 206 belong to the bull market periods, and 82 belong to the bear market periods. Bull and bear markets cycles are determined as a continuous performance of the market index of at least +20 % and -20 %, respectively. Since the performance is measured with 1-, 2-, 3-, and 4-year holding periods, the 24-year sample period divides the holding periods evenly. Thus, the total number of portfolio reformations for each holding period is 24, 12, 8, and 6, respectively. In order to avoid survivorship bias, all stocks that were delisted during the sample period are also included. In order to measure returns correctly, adjustments for dividends and splits are made appropriately by using the stocks' total return index, which takes dividends and stock splits into account.

The risk-free rate used is one-month Helibor-rate during 1.5.1996-1.12.1998 and switched to one-month Euribor rate from 1.1.1999 until the end of the sample period. The monthly rates are converted to correspond to a monthly yield. The Helibor-rates are from the Bank of Finland database, while the Euribor-rates are downloaded from Thomson-Reuters datastream.

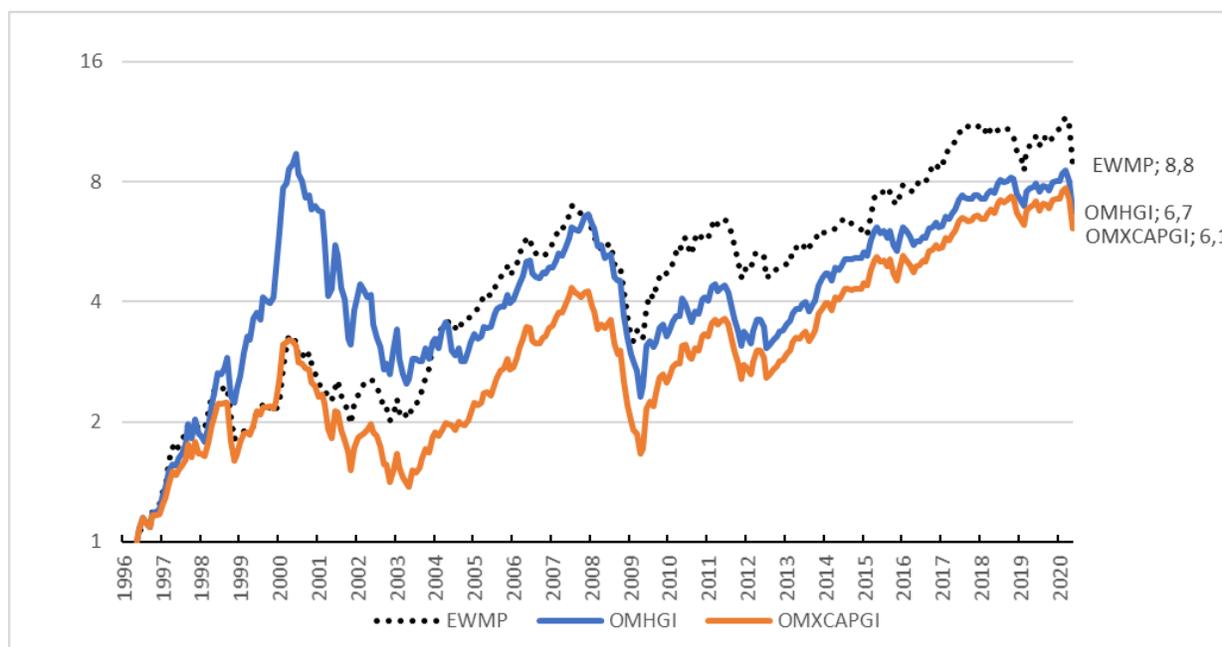
#### **3.2 Portfolios**

The portfolios are always constructed on the first trading day of May. This is based on delays related to publications of financial statements, and thus look-ahead bias is avoided. The portfolios

are divided into quartiles to allow sufficient diversification. The stocks in quartile 1 (Q1) consist of stocks with the highest valuation ratios, and thus, Q1 portfolios are considered as value portfolios, whereas Q4 portfolios are considered as growth portfolios, including the stocks with the lowest ratios. Equal weights are used for all stocks during portfolio formation. If any relevant financial data is absent during the portfolio formation moment, the stocks in question are forced to be omitted from the portfolios during the portfolio formation moment to avoid the stocks being chosen for a wrong portfolio. The smallest total amount of stocks during a portfolio formation was 62 with EBITDA/EV in 1996, and the largest was 153 with B/P in 2019 (with a 1-year holding period). Possible deviations from the number of yearly stocks were ratio-specific due to some incomplete financial data of some firms downloaded from Datastream. If the number of stocks cannot be divided equally to all quartiles during portfolio formation, it is made sure that portfolios Q1 and Q4 always have an equal amount of stocks. Also, if a stock was delisted due to privatization, takeover or bankruptcy (or any other reason), the stock is "sold" at the current market price, and the leftover cash is reinvested to the remaining stocks in the same portfolio according to the stocks' current relative weights during the first trading day of the first following month in order to maximize profits. However, all transaction costs, such as commissions, as well as bid-ask-spreads, are excluded.

The market portfolio is constructed using all the stocks from the Helsinki stock exchange main list and First North list. Zivney and Thompson (1989) suggest that the most reliable results regarding abnormal portfolio returns are achieved using a market portfolio that is constructed the same way as the portfolios. Thus, the market portfolio is constructed by weighting all the stocks equally every year during the first trading day of May. If a stock is delisted, reinvestments to remaining stocks in the market portfolio are applied the same way as in the quartile portfolios. In the case of this study, the equally weighted market portfolio is also more competitive than market indices from the Helsinki stock exchange that are weighted in terms of firms' market capitalization. This is illustrated in figure 1: The equally weighted market portfolio (EWMP) shows clear outperformance over both OMX Helsinki gross index (OMXHGI) and OMXH weight limited gross index (OMXHCAPGI) in terms of average annual excess returns (9,51 %, 8,28 % & 7,84 % respectively) and annual volatility (18,2 %, 25,1 % and 19,7 % respectively). The overperformance

can be considered as compensation for the transaction costs that are omitted from the analysis. Thus more reliable results are obtained when comparing the performance between the quartile portfolios and the equally weighted market portfolio.



**Figure 1. Comparison of compounding excess returns of different market indices.** *The EWMP stands for Equally Weighted Market Portfolio, the OMXHGI is the OMX Helsinki gross index, and the OMXHCAPGI is the OMX Helsinki weight limited gross index. The compounding returns are presented on a logarithmic scale.*

The composite value measure portfolios are constructed by ranking individual ratios and summing the ranks of the chosen ratios together accordingly. Thus, the firm with the highest valuation ratio obtains a rank one, and the lowest ratio firm obtains the lowest rank. The combined rank values are then summed together, and the firms with the lowest summed ranks belong to the value portfolio quartile (Q1), and the firms with the highest summed ranks belong to the growth portfolio quartile (Q4). In this study, a total of 10 different composite value measures are used as portfolio formation criteria, of which five are combinations of two value measures, and the remaining five are combinations of three value measures. The combinatory selections are based on either their common use, versatility, significant previous performance, or low correlation between individual ratios. The first composite value measure is named 2A and is a combination of E/P and B/P, which is similar to the Graham's Number introduced in chapter

2.5.6. However, Graham multiplies the two ratios (P/E & P/B), while this study sums the ranks together. The second two-ratio combination is named 2B and combines B/P and EBITDA/EV. This combination has the lowest correlation (0,769) among all other ratios used in this study (Figure 2), which justifies the selection. The third two-ratio measure is D/P combined with FCF/P and notated as 2C. While D/P measures the expected dividend yield, FCF/P accounts for recently paid dividends along with capital expenditures. It also has the fourth-lowest correlation (0,821) among other ratios (Figure 2), giving more justification for the selection. For comparison, another dividend yield based ratio is selected, which is a combination of D/P and EBITDA/EV and notated as 2D. Pätäri and Leivo (2009) showed the combination to generate the highest annual returns on a full sample period and a larger alpha spread than using the two ratios individually. EBITDA/EV also takes leverage differences into account better than E/P, thus providing an additional dimension to relative valuation. The last two-ratio measure is 2E with a combination of FCF/P and EBITDA/EV, which also has the second-lowest correlation (Figure 2). Moreover, the two ratios alone take profitability, leverage, investments (capital expenditures), and paid dividends into account, providing even more versatility to the relative valuation.

**Table 1. Correlation between the returns of value portfolios based on individual valuation ratios.**

	E/P	B/P	D/P	FCF/P	EBITDA/EV
E/P	1,000				
B/P	0,841	1,000			
D/P	0,917	0,817	1,000		
FCF/P	0,841	0,928	0,821	1,000	
EBITDA/EV	0,864	0,769	0,905	0,771	1,000

The first three-measure combination is named 3A and combines E/P, B/P, and D/P. The second is 3B, which replaces E/P with FCF/P. Thus, FCF/P, B/P, and D/P are combined. The third, named 3C, is fairly similar to the two previous, combining EBITDA/EV, B/P, and D/P. These three combinations provide great comparability to each other since they all include B/P and D/P. In the study published by Leivo and Pätäri (2009), the previously mentioned combination (3C) also

generated the largest alpha among all portfolios, thus similar results could be expected in this study. The last two portfolios take the versatility of FCF/P and EBITDA/EV into account and are combined with relatively low correlation ratios. Thus, the fourth selection is a combination of B/P, FCF/P, and EBITDA/EV, named 3D, and the last combines D/P, FCF/P, and EBITDA/EV, notated as 3E.

### 3.3 Sharpe ratio

One of the most well-known portfolio performance metrics is the Sharpe ratio. The ratio is a risk-adjusted measure, which uses the asset's standard deviation as the risk measure. The Sharpe ratio is formulated in the following way:

$$\text{Sharpe ratio} = \frac{r_i - r_f}{\sigma_i} \quad (9)$$

Where  $r_i$  is the average return of asset  $i$ ,  $r_f$  is the risk-free rate and  $\sigma_i$  is the standard deviation of asset  $i$ . (Sharpe, 1966)

Its popularity stems from its simplicity, giving an intuitive approach to any investment strategy's essential aspects – risk and return (Sharpe, 1994). It is considered one of the best performance measures due to its simplicity and solid theoretical framework (Eling, 2008).

Regardless of the common use and popularity, the ratio has a weakness in its interpretation when two or more negative Sharpe ratios are compared. For example, if two portfolios with equally negative returns have different standard deviations, the one with higher volatility has a higher Sharpe ratio. To counteract this weakness, Israelsen (2005) developed a modified Sharpe ratio, which makes negative Sharpe ratios comparable together:

$$\text{Modified Sharpe ratio} = \frac{r_i - r_f}{\frac{ER}{\sigma_i |ER|}} \quad (10)$$

Where  $ER$  is the average excess return of portfolio  $i$ . In this study, the modified version is beneficial when examining bear market periods since all portfolios have negative returns during those periods.

The Sharpe ratio assumes that the distribution is mean-centered, and therefore, it works best in situations where the portfolio returns are normally distributed (Lo, 2002; Eling & Schumacher, 2007). Thus, the Sharpe ratio has gained critique due to oversimplifying risk since it does not account for distribution skewness or kurtosis, potentially leading to biases. For example, a right-skewed return distribution would be desirable for a portfolio, however, the Sharpe ratio would penalize for the upside potential. (Pätäri, 2011)

### 3.4 SKASR (Skewness and Kurtosis Adjusted Sharpe Ratio)

The asymmetries in deviation can be adjusted by replacing the standard deviation with skewness and kurtosis adjusted deviation (SKAD). Introduced by Pätäri (2011), this ratio is known as SKASR (Skewness and Kurtosis Adjusted Sharpe Ratio) or adjusted Sharpe ratio, and is used along with the original Sharpe ratio in this study. To obtain SKAD, Cornish-Fisher (1937) expansion of Z value – adjusted Z value ( $Z_{CF}$ ) – is calculated first:

$$Z_{CF} = Z_C + \frac{1}{6}(Z_C^2 - 1)S + \frac{1}{24}(Z_C^3 - 3Z_C)K - \frac{1}{36}(2Z_C^3 - 5Z_C)S^2 \quad (11)$$

where  $Z_C$  is the critical value for normal distribution,  $S$  is the skewness, and  $K$  is the kurtosis of the given return distribution. Respectively, formulas for  $S$  and  $K$  are given as follows:

$$S = \frac{1}{N} \sum_{t=1}^N \left( \frac{r_{it} - \bar{r}_i}{\sigma} \right)^3 \quad (12)$$

$$K = \frac{1}{N} \sum_{t=1}^N \left( \frac{r_{it} - \bar{r}_i}{\sigma} \right)^4 - 3 \quad (13)$$

where  $N$  is the number of observations,  $\bar{r}_i$  denotes the average return, and  $r_{it}$  is the portfolio return at time  $t$ . Then, the SKAD is calculated as follows:

$$SKAD = \sigma * \frac{Z_{CF}}{Z_C} \quad (14)$$

Lastly, the skewness and kurtosis adjusted Sharpe ratio can be obtained. In order to avoid the biased estimates stemming from the negative returns during the bear markets, the Israelsen (2005) extension is also applied. Therefore, SKASR can be formulated with the following equation:

$$SKASR = \frac{r_i - r_f}{\frac{ER}{SKAD_i^{|ER|}}} \quad (15)$$

where  $SKAD_i$  is the skewness- and kurtosis-adjusted monthly excess returns of portfolio  $i$ , and  $ER$  denotes the portfolio's average excess return.

### 3.5 Three-factor model

Introduced in chapter 2.2, the Fama-French (1992) three-factor model (FF3) was developed to improve the shortcomings of the original capital asset pricing model (CAPM) by Sharpe (1964) and Lintner (1965). Fama and French (1992) found that the excess market return as an explanatory variable was not sufficient enough to explain assets' price movements. Thus they introduced the SMB-factor (small-minus-big) and HML-factor (high-minus low). The FF3 equation is formulated as follows:

$$(r_i - r_f) = \alpha_i + \beta_{i1}(r_m - r_f) + \beta_{i2}SMB + \beta_{i3}HML + e_i \quad (16)$$

where

$r_i - r_f$	=	portfolio excess return over risk-free rate,
$\alpha$	=	overall excess return determined by the model,
$r_m - r_f$	=	market excess return over risk-free rate
$\beta_{it}$	=	beta of factor $t$ ,
$SMB$	=	small-minus-big factor (excess return of small cap over large cap),
$HML$	=	value-minus-growth factor (excess return of high B/P firms over low B/P firms), and
$e_i$	=	error term.

In this study, the HML and SMB factors are created from market indices provided by Morgan Stanley Capital International (MSCI) and downloaded from Thomson-Reuters Datastream. For the HML factor, we use Finnish value and growth stock indices. The value index monthly returns are subtracted by the growth index monthly returns to obtain the factor values, thus reflecting the excess returns of value stocks over growth stocks. As for the SMB, we use Finnish small-cap and large-cap indices. The small-cap index monthly returns are subtracted by the large-cap index monthly returns, thus reflecting the excess returns of small-cap firms over large-cap firms.

This study utilizes the FF3 by interpreting the estimated annual alphas ( $\alpha$ ) obtained from the model. The larger the alpha is, the better is the portfolio performance. We also examine the alpha-spread between each value portfolio with a comparable growth portfolio. This spread is interpreted as the value premium, and thus, the larger the spread is, the larger is the effect of value anomaly with the concerning value measure.

Lastly, to examine the size effect in value portfolios, we also conduct a two-factor model and compare the obtained alphas with the three-factor model alphas. The two-factor equation is equivalent to the three-factor equation, except the SMB factor is omitted. If the differences between the two models' alphas are statistically significant, we conclude that the concerning portfolio was heavily affected by firm size. More specifically, if a statistically significant positive spread is detected, a portfolio's performance was highly affected by small firm size. On the contrary, a large negative spread would indicate a considerable influence of large-cap stocks in the portfolio returns.

### **3.6 Statistical testing**

Statistical testing is applied to improve the robustness of the overall results. In this thesis, the purpose is to test the statistical significance of the differences between the performance of different portfolios. Thus, the null hypothesis of these tests is that there is no difference in given performance measures. Rejecting the null hypothesis would mean that the differences are statistically significant, and therefore, we can be reasonably confident that the result is not obtained by chance. The critical value for statistical significance used for all the tests is 5 % (0,05).

### 3.6.1 Jobson-Korkie-Memmel Z-test

Jobson and Korkie (1981) utilized a two-sided Z-test for measuring the differences between Sharpe and Treynor ratios between two portfolios. This thesis uses the Memmel (2003) version of the test, which simplifies the Jobson-Korkie test statistic without losing statistical properties. The Jobson-Korkie-Memmel Z-test is utilized for both Sharpe and SKASR measures. In both cases, the value portfolios are tested against comparable growth portfolios, and all quartile portfolios are tested against the market portfolio. The Z-statistic for both the Sharpe ratio and SKASR differences is calculated as follows:

$$Z = \frac{\widehat{Sk}_n - \widehat{Sk}_i}{\sqrt{\widehat{\theta}}} \quad (20)$$

where  $\widehat{Sk}$  is the Sharpe ratio or SKASR of portfolios  $i$  and  $n$ .  $\widehat{\theta}$  is the asymptotic variance between the two ratios, which is calculated with the following equation:

$$\widehat{\theta} = \frac{1}{T} \left[ 2 - 2\rho_{in} + \frac{1}{2} (Sk_i^2 + Sk_n^2 - 2Sk_iSk_n\rho_{in}^2) \right] \quad (21)$$

where  $\rho_{in}$  is the correlation between portfolios  $i$  and  $n$ .

After obtaining the Z-statistic, a critical Z-value is applied to measure the boundary of a statistically significant Z-value. The corresponding 2,5 % critical Z-value for the positive and negative sides is 1,96 and -1,96 respectively, reflecting the overall 5 % (0,05) risk level. If the obtained Z-value passes underneath -1,96 or exceeds 1,96, the Sharpe or SKASR differences are considered statistically significant, and thus, there is a fairly low probability that the result is obtained by chance.

### 3.6.2 Welch's t-test

The Welch's (1947) t-test is designed for data samples with unequal variances. The Welch's t-test statistic is calculated with the following formula:

$$\text{Welch's } t - \text{statistic} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}}} \quad (22)$$

where  $\bar{X}_1$  is the alpha of a value portfolio,  $\bar{X}_2$  is the alpha of a comparable growth portfolio or the market portfolio,  $s_1^2$  is the standard error of the value portfolio,  $s_2^2$  is the standard error of the comparable growth portfolio,  $N_1$  is the sample size of the value portfolio, and  $N_2$  is the sample size of the comparable growth portfolio.

To find the critical t-value, degrees of freedom  $\nu$  needs to be determined. The Welch's t-test degrees of freedom are determined as follows:

$$\nu = \frac{\left(\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}\right)^2}{\frac{s_1^4}{N_1^2 \nu_1} + \frac{s_2^4}{N_2^2 \nu_2}} \quad (23)$$

Where  $\nu_1 = N_1 - 1$  and  $\nu_2 = N_2 - 1$ .

This study tests a two-sided test for statistical significance of alpha spreads between value and growth portfolios conducted from the Fama-French three-factor model. For the size effect in value and growth portfolios, we test the statistical significance of the spread between the three-factor and two-factor alphas.

## 4. Results

In this section, the overall results are examined extensively. First, we will view the performance of the value strategy against the market and growth strategy during the full sample period and the bull and bear markets with one-year holding period portfolios. Concurrently, we will see if the composite value measures can enhance the performance of individual value measure portfolios (chapter 4.1). We will continue by examining the extended holding period portfolios' performance for the full sample period and see if they can enhance the value strategy performance compared to the one-year holding period portfolios (chapter 4.2). Lastly, we will find the effect of firm size in value and growth portfolio performance by comparing Fama-French 3-factor alphas with 2-factor alphas and seek for statistical significance in their differences (chapter 4.3).

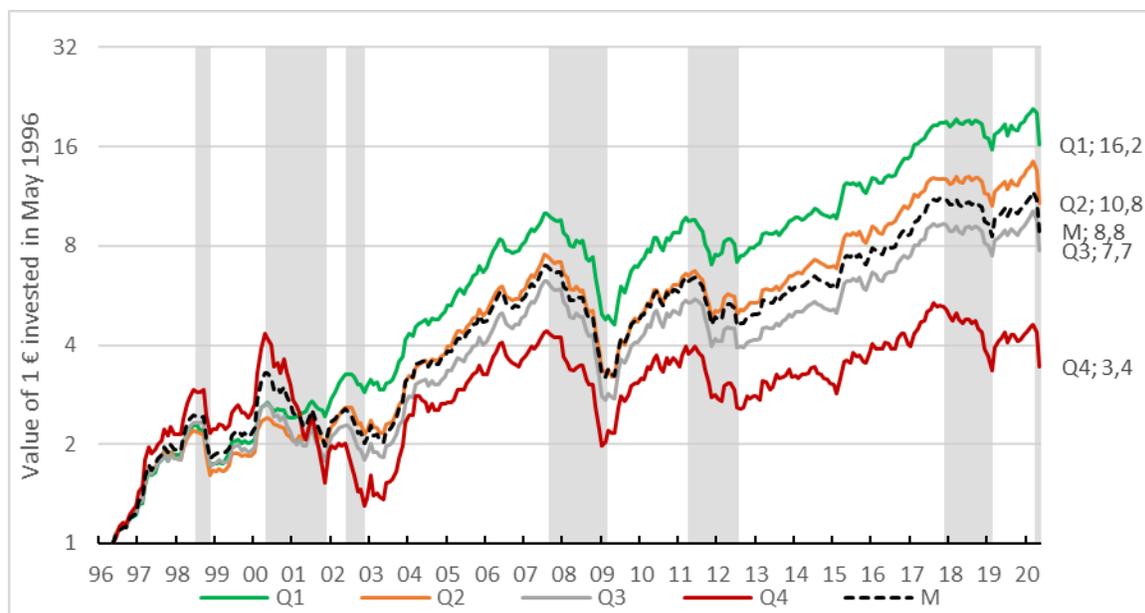
### 4.1 Performance of Individual and composite value measure portfolios

We proceed by examining the performance of the value strategy against the growth strategy and the market with one-year holding period portfolios and compare the performance between individual and composite value measure portfolios. We begin by examining the full sample period and continue with the bull and bear market inspection. The full sample period consists of 288 observations, while the total length of all the bull and bear market periods are 206 and 82, respectively.

#### 4.1.1 Full sample period

The overall performance of the value strategy shows enhanced returns in relation to the market and growth portfolios. Figure 2 shows the average compounding performance of all individual and composite value measure portfolios combined for each quartile (Q1 – Q4) on a logarithmic scale with a one-year holding period. The sample period begins with a strong upswing. At the outset, the growth strategy seems to exceed the value strategy and the market until the “dotcom” crash of the early 2000s, where the growth portfolios do not seem to retain their value as well as the value portfolios and the overall market. Through the course of the seven bull and seven bear

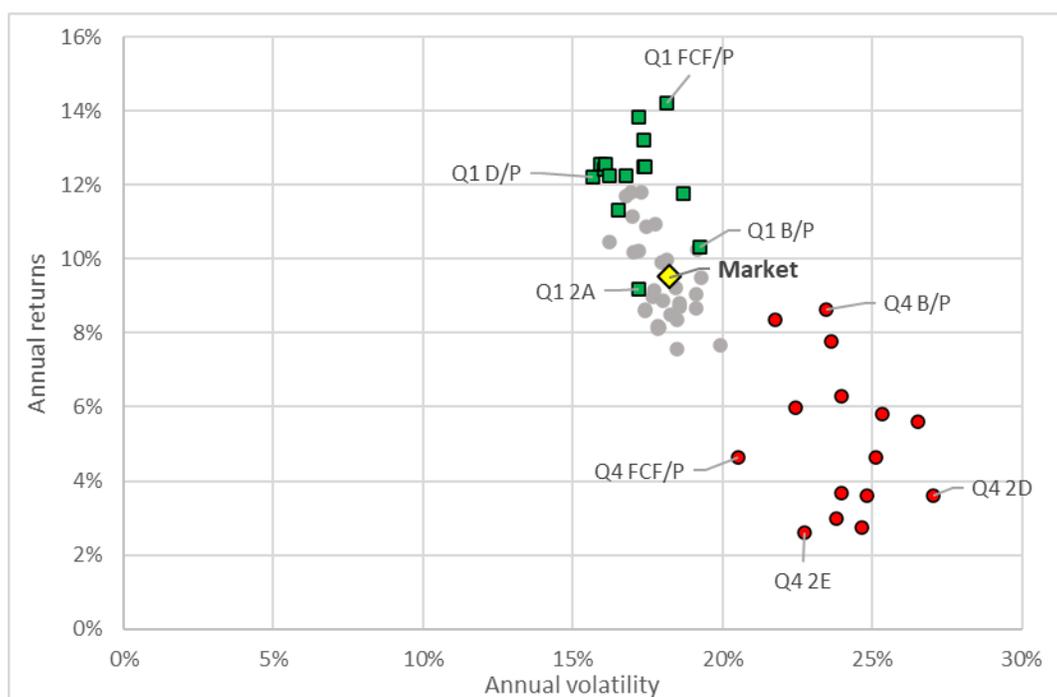
market periods total, it is evident that the difference in returns between value and growth portfolios begins to increase substantially, with value strategy showing clear over-performance over the growth portfolios and the market in the long-term. Hence, the average value strategy's overall performance shows the highest total return of over 16 times (16,2) the initial investment, while the market has generated slightly under nine times (8,8) its initial value. The average growth strategy has merely tripled (3,4) its value, having the lowest overall performance.



**Figure 2. Compounding returns of average quartile portfolios of one-year holding period during the full sample period (1996-2020).** The graph presents the average compounding returns of all individual value measure and composite value measure portfolios combined. The gray areas show the bear market periods, during which the market has declined at least -20 % from a previous peak of at least +20 %. A total of 7 bull market and 7 bear market periods are identified. The compounding returns are presented in logarithmic format.

In terms of the traditional mean-variance framework, the value portfolios have generated greater returns with less risk than the growth portfolios and the market. Figure 3 presents the relationship between portfolio volatility and returns during the full sample period. The plot shows that the highest average annual returns of 14,20 % were generated by stocks with high free cash flows per share relative to stock price (Q1 FCF/P), while the lowest risk was achieved with high dividend yield stocks (Q1 D/P) with an annual volatility of 15,66 %. Moreover, it is shown that the value portfolio with the lowest average annual returns of 9,20 % (Q1 2A) was yet able to exceed the

growth portfolio with the highest returns of 8,62 % (Q4 B/P). Furthermore, the value portfolio with the highest volatility of 19,26 % (Q1 B/P) shows less risk than the growth portfolio with the lowest volatility of 20,51 % (Q4 FCF/P). Thus, all the value portfolios seem to cluster more or less to the low-risk and high-returns side, whereas the growth portfolios tend to cluster to the high-risk and low-returns side. All the value portfolios, except portfolio 2A, seem to generate higher annual returns than the market, while none of the growth portfolios exceed the market. These implications seem to counteract with the traditional financial theories, the CAPM by Sharpe (1964) and Lintner (1965), and the modern portfolio theory by Markowitz (1952), which both assume a positive relationship between risk and return, thus higher risk expectations should result in higher returns and vice versa.



**Figure 3. Relationship between risk and return during the full sample period (1996-2020).** Value portfolios are presented in green squares, growth portfolios are shown in red circles, the yellow diamond represents the market, and gray circles represent all other portfolios (Q2 and Q3 portfolios). The horizontal axis indicates volatility, and the vertical axis indicates average annual returns.

Table 2 exhibits the full sample period results of all the value and growth portfolios for a one-year holding period, including all risk-adjusted performance metrics with statistical significances against the market (sign. Qi vs. market) and the value portfolios against growth portfolios (sign.

Q1 vs. Q4). All the results are presented in annual terms. The table shows that 14/15 value portfolios have exceeded the market's 9,51 % average annual return, and all the value portfolios exceed the growth portfolios. As for the risk-adjusted metrics, the Sharpe ratios show that all 15 value portfolios overperform the market, of which 10 with statistical significance under the 5 % risk level. All the value portfolios also overperform their comparable growth portfolios and 13/15 with statistical significance. Taking the deviations in return distributions into account, the skewness and kurtosis adjusted Sharpe ratios (SKASR) show that only the performance of the 2A portfolio (the combination of E/P and B/P) falls below the market, and thus 14/15 value portfolios exceed the market with the SKASR measure, although merely 4/12 of them with statistical significance. Lastly, the three-factor alphas ( $\alpha$ ) show that all the value portfolios generated a positive alpha, of which 13/15 are statistically significant.

The first five portfolio measures in table 2 exhibit the individual value measure portfolios. In terms of annual returns, the free cash flow based FCF/P has generated the highest returns of 14,20 %. Moreover, it dominates all other individual ratios with every risk-adjusted performance metric with a Sharpe ratio of 0,783, SKASR of 0,714, and three-factor alpha ( $\alpha$ ) of 5,17 %, all statistically significant under 5 % risk level against both the market and the comparable FCF/P growth portfolio. This result complies with Hackel et al. (2000) and Jokipii and Vähämaa (2006), as the authors exhibited outperformance of high FCF/P stocks over the market and all other portfolio selection criteria. However, the highest value premium (Q1 vs. Q4) was obtained with EBITDA/EV among the individual ratios with an alpha spread of 10,27 %. The lowest-performing value portfolio is the B/P measure, having the lowest performance with all the valuation metrics and the lowest value premium among all the individual ratios.

The following five value measures in table 2 exhibit the performance of 2-composite value measure portfolios (2A – 2E). In terms of annual returns, none of the portfolios could exceed the FCF/P value portfolio's 14,20 % annual returns. Taking the risk adjustments into account, combining FCF/P with EBITDA/EV (2E) generated a Sharpe ratio of 0,803, overtaking the individual FCF/P portfolio. A similar result was obtained by combining FCF/P with D/P (2C), generating a marginally enhanced Sharpe ratio of 0,788. Both results were statistically significant against the market and their comparable growth portfolios.



However, taking the skewness and kurtosis of their return distributions into account, both of the SKASRs of value portfolios 2E (0,713) and 2C (0,616) fall below the SKASR of the FCF/P value portfolio (0,714). However, the SKASRs between 2E and FCF/P differ only by 0,001. Compared to their Sharpe ratios, the lower SKASRs of 2E and 2C indicate that their return distributions are more negatively skewed than the FCF/P portfolio, thus having more months with decreasing value or months with stronger decreasing values. Looking at the three-factor alphas, the highest two-value measure alpha of 5,10 % was obtained with the combination of the lowest correlating value measures B/P and EBITDA/EV (portfolio 2B), thus falling slightly short from the 5,17 % alpha of the FCF/P. Hence, none of the 2-composite value measures could exceed the alpha of the FCF/P, even though 4/5 of them are statistically significant, and all are showing overperformance over their comparable growth portfolios with positive alpha spreads. However, enhanced value premium was best obtained by combining EBITDA/EV with FCF/P (2E) with a value premium of 12,03 %, overtaking the individual EBITDA/EV measure. Larger value premiums were also obtained with D/P based combinations: EBITDA/EV with D/P (2D, 11,75 %) and FCF/P with D/P (2C, 11,40 %), all statistically significant under 5 %. The least favorable two-measure value portfolio is the 2A, a combination of E/P and B/P. Although it shows overperformance over most growth portfolios, it has a deteriorating effect on individual E/P and B/P value portfolios. Table 2 shows that the 2A's annual return of 9,20 % and all its risk-adjusted performance metrics are the lowest concerning all other value portfolios.

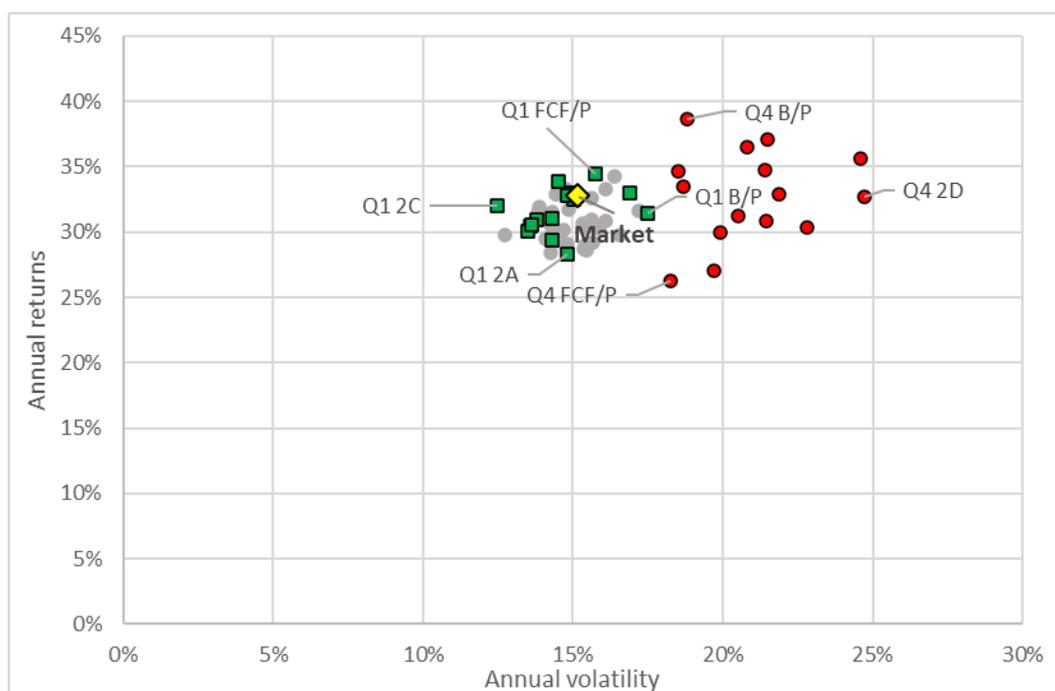
Lastly, we compare the three-value measure portfolios (3A – 3E) with the individual and two-composite value measure portfolios (table 2). Again, looking at the annual returns first, none of the value portfolios 3A – 3E could exceed the FCF/P value portfolio. Thus the concerning individual measure generated the highest returns among all the 15 value measures. As for the Sharpe ratios, no further enhancements were obtained over the highest value of 0,803 of the two-value measure 2E (FCF/P & EBITDA/EV). Moreover, taking the deviations from distribution normality into account, the highest three-composite measure SKASR of 0,652 obtained by value portfolio 3B (B/P, FCF/P & EBITDA/EV) was not able to surpass the overall highest SKASR of the individual FCF/P value portfolio (0,714). The same holds with the three-factor alphas, and thus the alpha of the FCF/P measure (5,17 %) further dominates the highest three-composite measure alpha of

4,68 % obtained with the combination of B/P, D/P, and FCF/P (3B). The highest alpha spread among the 3A–3E portfolios was generated with 3E (D/P, FCF/P, and EBITDA/EV) with a value premium of 11,57 %. However, it does not exceed the overall highest value premium of the value portfolio 2E (12,03 %), though surpassing the individual EBITDA/EV value premium (10,27 %).

Complying with Leivo et al. (2009) and Pätäri and Leivo (2009), value anomaly is evident in the Finnish stock markets, and thus, the value portfolios exceed the market and the growth portfolios almost without exception and with statistical significance in most cases. Similar to this study, the authors suggest that portfolio performance can be somewhat enhanced using combinations with the EBITDA/EV measure. However, this thesis implies that even superior results could be obtained using the free cash flow based FCF/P measure in terms of both annual returns and risk-adjusted metrics. Thus, the best results are obtained by selecting stocks with a high free cash flow rate per share. The individual value measure mostly dominates all other value portfolios with all performance metrics, except only surpassed by the Sharpe ratios of its combinatory counterparts 2E (FCF/P & EBITDA/EV) and 2C (D/P & FCF/P). In addition, many of the combination portfolios are able to amplify the value premiums of the individual ratios, and the highest value premium of 12,03 % was obtained with the two-composite measure value portfolio FCF/P and EBITDA/EV (2E).

#### **4.1.2 Bull market periods**

Next, the value strategy's performance during the bull market periods is examined. Figure 4 shows the relationship between risk and return for every portfolio during all bull market periods on average. Again, it is evident that the value portfolios tend to have relatively low risk compared to the growth portfolios. Similar to the full sample period (see figure 3 and figure 4), the most volatile value portfolio (Q1 B/P) is still less-riskier than the least volatile growth portfolio (Q4 FCF/P). However, the annual returns of the value portfolios are exceeded by many of the growth portfolios: The highest value portfolio returns are obtained with FCF/P, however, being surpassed by six growth portfolios, with Q4 B/P generating the highest average returns (38,65 %) among all the portfolios.



**Figure 4. Relationship between risk and return during the bull market periods.** Value portfolios are presented in green squares, growth portfolios are shown in red circles, the yellow diamond represents the market, and gray circles represent all other portfolios (Q2 and Q3 portfolios). The horizontal axis indicates volatility, and the vertical axis indicates average annual returns.

Table 3 shows the overall results of the value and growth portfolios during the bull market periods. Looking at the overall performance of the value portfolios first, we can see that only 5/15 exceed the market, and the same amount exceed their comparable growth portfolios in terms of annual returns. As for the Sharpe ratios, 11/15 of the value portfolios exceed the market due to their low-risk nature. However, only the overperformance of portfolio 2C (D/P & FCF/P) is statistically significant. On the contrary, the Sharpe ratio of the B/P value portfolio (1,796) falls below the market (2,155) with statistical significance. The performance against the growth portfolios is more superior with the Sharpe ratios: 14/15 value portfolios exceed their Q4 counterparts, and 10/15 with statistical significance. Adjusting for distribution skewness and kurtosis, only 2/15 of SKASRs of the value portfolios exceed the market, neither of them with statistical significance. In addition, the underperformance of portfolios D/P, 2A, and 3A become statistically significant.



Although the positively skewed return distributions increase the Sharpe ratios during the bull markets for all portfolios due to lower adjusted risk (SKAD), the same holds for the market. The market's positive skewness seems to be even stronger than that of the value portfolios, thus surpassing many of the value portfolios' Sharpe ratios after deviation adjustments. Compared with the growth portfolios, 12/15 SKASRs of the value portfolios are higher than that of their comparable growth portfolios, and half of them (6) surpass them with statistical significance. As for the three-factor alphas, overperformance is identified with 14/15 value portfolios, with 10/15 cases statistically significant. Moreover, the alpha spreads are positive in 14/15 instances, and 12/15 of them are statistically significant. Since the three-factor model measures risk with portfolio betas, the positive alpha spreads can be explained by the relatively low betas of the value portfolios determined by the three-factor model versus relatively high betas of the growth portfolios.

Looking at the portfolios' performances in more detail, the highest annual returns of 38,65 % were generated with the B/P growth portfolio. The best performing value portfolio FCF/P generated the seventh-highest annual returns of 34,44 % among all the portfolios. The lowest annual returns were generated with low free cash flows per share to price ratio, and thus the annual returns of the FCF/P growth portfolio were merely 26,25 %.

The performance of the value portfolios becomes slightly enhanced when considering risk-adjustments. Among all of the portfolios, the highest Sharpe ratio (2,569) is obtained with the two-value measure portfolio Q1 2C (D/P with FCF/P), enhancing the results of both individual measures D/P (2,241) and FCF/P (2,186). The high ratio is explained by low volatility (12,47 %), which is the lowest of all the portfolios. The ratio is also highly statistically significant against the market and its comparable Q4 growth portfolio. Enhanced Sharpe ratios were also obtained with all other two-composite measure value portfolios (2B, 2D & 2E) except the portfolio 2A. However, none of them are statistically significant against the market, but 2D and 2E overperform their comparable growth portfolios with statistical significance. As for the three-value measure portfolios, the Q1 of 3B, 3C, 3D, and 3E overperform their Q4 counterparts with statistical

significance. Moreover, a slightly improved Sharpe ratio of 2,210 was obtained with the 3D value portfolio compared to the individual ratios it combines (B/P 1,796, FCF/P 2,186 and EBITDA/EV 2,163). The lowest value portfolio performance is shown with the B/P ratio, being outperformed by the market and its Q4 counterpart and having the lowest Sharpe ratio among all the value portfolios.

As for the deviation adjusted Sharpe ratios, surprisingly, the highest value is obtained by the 2A's third quartile (Q3) with a SKASR of 3,199. Its Sharpe ratio (1,912) was highly penalized by strong positive skewness in returns. Thus, taking the skewness and kurtosis into account, its volatility falls from 17,21 % to merely 9,89 %. Hence, regardless of a moderate return of 31,65 %, the portfolio obtains the highest SKASR (Appendix 6, Panel B). The highest value portfolio SKASR was obtained with the individual measure FCF/P, having the fourth largest SKASR of 3,024 among all the portfolios, falling below Q3 2A (3,199), Q2 FCF/P (3,199), and Q3 2C (3,025). On the contrary, the B/P value portfolio again shows an apparent underperformance against the market and the comparable growth portfolio after deviation adjustments.

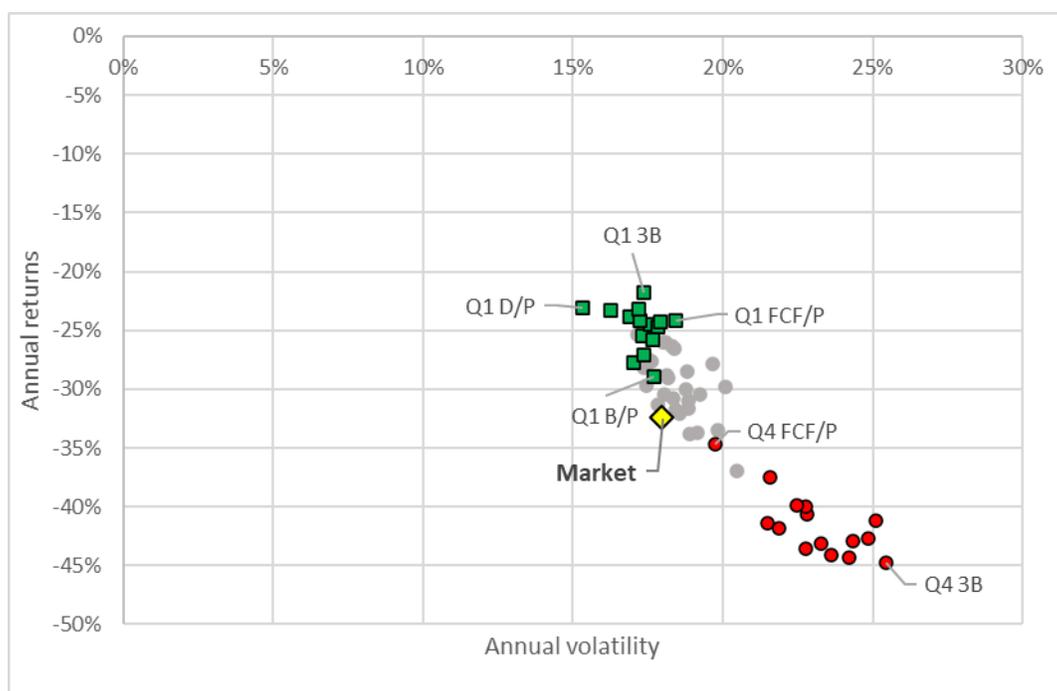
The overall highest bull period alpha (7,43 %) was obtained with the 2C value portfolio. The result is mostly explained by its low risk, having the lowest beta ( $\beta$ ) of 0,727 among all the portfolios. However, the highest alpha spread of 15,98 % was obtained by combining EBITDA/EV and D/P. On the contrary, the B/P value portfolio exhibits the lowest alpha spread of -2,66 %. Hence the growth strategy would be more advisable with B/P growth stocks than B/P value stocks during bull markets.

We can conclude that the long-term overperformance of the value strategy is not emphasized during economically flourishing times. However, due to the low-risk nature of the value portfolios, many of them seem to become more favorable after risk-adjustments. Regardless, most of the performances are not statistically significant. Enhanced risk-adjusted results could be best obtained with the two-value measure combination of D/P and FCF/P, with the overall highest Sharpe ratio and three-factor alpha 2,569 and 7,43 % respectively, both statistically significant against the market and its comparable growth portfolio. Although most value portfolios are capable of greater risk-adjusted performance in comparison with their comparable growth

portfolios, they do not exceed the market as successfully. A risk-averse investor would still prefer the value strategy due to lower risk, although with the cost of lower returns. The deviation between the high and low-performing growth portfolios is higher than with the value portfolios. Thus a growth-investor should be more careful with their stock selections than a value-investor during periods of market upswings.

#### 4.1.3 Bear market periods

During times of economic distress, the overperformance of the value strategy is the most distinguishable. Figure 5 shows the relationship between portfolio returns and volatility during the bear market periods. The graph shows that the average annual returns of all the portfolios are negative. However, the value portfolios seem to retain their value more successfully than all the growth portfolios and the market.



**Figure 5. Relationship between risk and return during bear market periods.** Value portfolios are presented in green squares, growth portfolios are shown in red circles, the yellow diamond represents the market, and gray circles represent all other portfolios (Q2 and Q3 portfolios). The horizontal axis indicates volatility, and the vertical axis indicates average annual returns.

Similar to the full sample period and bull market periods, the riskiest value portfolio (Q1 FCF/P) is less risky than the growth portfolio with the least risk (Q4 FCF/P). In addition, all the value portfolios have retained more value than the market, of which the three-composite value portfolio 3B (B/P, D/P & FCF/P) has retained the most value. The difference between the lowest-performing value portfolio (Q1 B/P) and the market is 3,55 %, while the difference with the highest value retaining growth portfolio (Q4 FCF/P) is 5,80 % in annual terms (table 4).

Table 4 exhibits the results for the value portfolios, growth portfolios, and the market during the bear market periods. Although the value portfolios with individual value measures have performed relatively well in terms of annual returns, the portfolios' value could be maintained more successfully with many of the composite value measures. The best result would be obtained by combining B/P, D/P, and FCF/P (3B), thus losing the least amount of value on average (-21,81 %). The value of the individual E/P (-27,75 %) and B/P (-28,91 %) portfolios can be slightly protected by combining them together (2A, -27,14 %).

Continuing the risk-adjusted metrics, all the Sharpe ratios and adjusted Sharpe ratios (SKASR) are negative due to all portfolios having negative returns. As discussed in chapter 3.3, the interpretation of negative Sharpe ratios can be problematic if portfolios with the same returns have different volatilities: The portfolio with higher volatility obtains a higher Sharpe ratio, and thus the interpretation would become counterintuitive. After applying the Israelsen's (2005) extension, the rankings of both Sharpe and adjusted Sharpe ratios become correctly organized. However, statistical testing for both performance measures no longer function properly, and therefore, the two measures are examined only by their absolute values. Nevertheless, the evidence of all the value portfolios outperforming the market and all the growth portfolios with both Sharpe ratio and SKASR measures is clear. This is due to their relatively low deviations in returns combined with the ability to retain more value. The dominating value measure with the highest values on both performance metrics is the D/P value portfolio, having the second-highest returns of -23,10 % and the lowest risk in terms of annual volatility (15,32 %) and adjusted volatility (18,53 %) among all the portfolios. Almost without exception, the adjusted volatility



Looking at the three-factor alphas in the table 4, similar results are detected. All 15 value portfolio alphas are positive, and all the growth portfolio alphas are negative, and thus the overperformance of the value strategy is evident. However, only 4/15 value portfolio alphas and 10/15 alpha spreads are statistically significant. Although the alphas and the alpha spreads are higher on average compared to the full sample period and bull markets, statistically significant alphas and alpha spreads can be obtained only with extremely positive (or negative) values. This requisite could be explained with higher standard errors caused by higher volatilities during economic distress and the substantially smaller sample size of the bear market periods. The highest alpha of 9,42 % and the highest value premium of 17,63 % are obtained with the value portfolio 3B, which combines B/P, D/P, and FCF/P. Both of the results are statistically significant.

In conclusion, the long-term overperformance of the value strategy can largely be explained by the value portfolios' ability to retain their value more successfully during the bear market periods than the average market and the growth portfolios. All the value portfolios outperformed the market and every growth portfolio with all of the performance metrics. Indeed, an even more superior option would be to invest in risk-free government bonds, which would have yielded +2,54 % annually on average during an average bear market period (table 4). Alternatively, an investor could sell all stocks before a market crash and retain all the remaining wealth, excluding possible transaction costs and taxes. However, the timing of market movements can be challenging. Thus, an investor could easily make inconsiderate investing decisions during turbulent times, causing more problems than benefits in the long-term – for instance, liquidating all positions and buying the same securities later with an even higher price, hence losing potential gains. Ignoring the risk-free options, the value was best retained with the three-composite value portfolio 3B (B/P, D/P & FCF/P combine) with -21,81 % annual returns. The same portfolio also generated the highest alpha of 9,42 % and the highest value premium of 17,63 %, both statistically significant under the 5 % risk level. The D/P value portfolio obtained the highest Sharpe ratio and SKASR. In addition, the same portfolio had the lowest risk with all of the risk metrics (volatility 15,32 %, SKAD 18,53 %, and beta 0,779) among all of the portfolios.

## 4.2 Performance of the value strategy with extended holding periods

The analysis continues by comparing the one-year holding period portfolios and extended holding periods of two, three, and four years. The aim is to find if the extended holding periods can enhance the performance of the value portfolios.

The average value portfolio performance is examined first. Table 5 exhibits all individual and combinatory value portfolios' combined average performance against the market and an average growth portfolio with one-, two-, three- and four-year holding periods. All the differences are positive, implicating that the average value portfolio has exceeded the market and the average growth portfolio with all holding periods and every performance metric. However, enhanced performance with extended holding periods is not entirely evident.

**Table 5. Performance differences of an average value portfolio against the market and an average growth portfolio with one-, two-, three- and four-year holding periods during the full sample period (1996-2020).** The table presents the differences of all the applied performance metrics (average annual returns, Sharpe ratio, SKASR, and three-factor alpha) for the average value portfolio against the market (Q1 vs. market) and the average growth portfolio (Q1 vs. Q4). The performance differences are presented in annual terms.

Holding period length	Return differences		Sharpe ratio differences		SKASR differences		3-factor alphas & alpha spreads	
	Q1 vs. market	Q1 vs. Q4	Q1 vs. market	Q1 vs. Q4	Q1 vs. market	Q1 vs. Q4	$\alpha$	Q1 vs. Q4
<b>1-year</b>	2,70 %	7,08 %	0,197	0,502	0,154	0,414	3,99 %	9,02 %
<b>2-years</b>	2,54 %	7,41 %	0,181	0,502	0,147	0,428	3,85 %	9,11 %
<b>3-years</b>	1,72 %	3,70 %	0,133	0,321	0,093	0,232	3,00 %	5,54 %
<b>4-years</b>	1,18 %	4,91 %	0,101	0,367	0,061	0,287	2,68 %	6,71 %

Looking at the differences in annual returns of table 5 first, the highest average returns over the market are obtained with the one-year holding period value portfolios (2,70 %) on average. The overperformance against the market gradually decreases with each extended year, and thus the lowest average returns are obtained with the four-year holding period value portfolios (1,18 %). However, the highest return differences against the growth portfolios are best obtained with the two-year holding period (7,41 %) and the narrowest difference with the three-year holding period (3,70 %). As for the Sharpe and adjusted Sharpe ratio (SKASR) differences, both of the results against the market are in identical order with the returns: The performance gradually decreases

when extending the holding period, and therefore, the best risk-adjusted average performance is obtained with the one-year holding period and the lowest with the four-year holding period. Against the growth portfolios, the best performance with the Sharpe ratios is identical with the one- and two-year holding periods (0,502), whereas after deviation adjustments, the two-year holding period SKASR (0,428) surpasses the one-year holding period (0,414). Regarding the alphas, the average alpha also decreases gradually from 3,99 % with the one-year holding period to 2,68 % with the four-year holding period. The highest average value premium is obtained with the two-year holding period (9,11 %), and the lowest with the three-year holding period (5,54 %).

Overall, by examining the value portfolios as a collective, no added value could be obtained by holding value stocks longer than one year, at least in respect of the market. In comparison with the growth portfolios (Q1 vs. Q4), slight enhancements could be obtained in terms of overall returns and risk-adjusted metrics when extending the holding period to two years. Moreover, the value premium seems to expand slightly with the two-year holding period. However, it is important to emphasize that transaction costs and taxation procedures were omitted from this study. Thus, higher portfolio reformation frequency would deteriorate the overall returns more than that of the more extended holding period portfolios in a real-life scenario.

Looking at the performance of the value portfolios in more detail, some enhanced results could be obtained with specific portfolios. Table 6 exhibits the results for all the value portfolios of the one-year holding period, as well as the extended holding periods of two-, three- and four years during the full sample period. Examining the average annual returns of the individual ratios first (Table 6, Panel A), extending the holding period length to two years has enhanced the returns of B/P (12,33 %), D/P (12,62 %), and FCF/P (14,67 %). Moreover, the holding period of four years also enhanced the one-year B/P value portfolio returns (11,66 %), although falling below the two-year returns. Moving to the Sharpe ratios in Table 6 Panel A, enhanced risk-adjusted returns are also obtained with the two-year holding period of B/P (0,626) and D/P (0,786), of which D/P is statistically significant against the market and both against their comparable growth portfolios under the 5 % risk level. Similarly, holding high B/P stocks for four years, the Sharpe ratio slightly surpasses (0,614) the one-year holding period Sharpe ratio. Adjusting for skewness and kurtosis in the Sharpe ratios, parallel results can be detected. Thus, enhanced SKASRs over the one-year

holding period are obtained with two-year holding periods of B/P, D/P, and FCF/P value portfolios, of which the two last-mentioned are statistically significant against the market and their comparative growth portfolios. Lastly, the alphas show similar results: the two-year holding period alphas of B/P (3,38 %), D/P (4,73 %), and FCF/P (5,47 %) exceed their corresponding one-year holding period value portfolios, of which D/P and FCF/P are statistically significant. The four-year alpha of B/P (3,21 %) also exceeds its one-year equivalent, although with no statistical significance. Wider alpha spreads could be obtained only with a two-year holding period of B/P and a four-year holding period of both B/P and E/P. Overall, the highest returns in terms of absolute returns and risk-adjusted returns can be obtained with FCF/P extending its holding period to two-years, while zero added value could be obtained with EBITDA/EV.

Table 6 panel B shows the performance metrics of two-composite value portfolios for each holding period length. Looking at the returns first, larger returns with the portfolio 2A can be obtained with all of the extended holding periods, of which two year holding period has the highest yield (10,59 %). The enhancements, however, do not exceed many of the other two-composite value portfolios. Additionally, only the three-year holding period of the value portfolio 2D could outperform its one-year holding period equivalent with a marginal difference. Moving to the risk-adjusted metrics, only the Sharpe ratios of two-year and four-year portfolios of the E/P and B/P combination (2A) slightly exceeded their one-year equivalents. Taking skewness and kurtosis into account, the SKASRs of two- and three-year holding period 2A exceeded the one-year equivalent. Moreover, holding the combination of high D/P and EBITDA/EV (2D) stocks for three years provides only an extremely marginal and insignificant advantage over the one-year holding period. As for the alphas, only the 2A value portfolio alphas are enhanced with all extended holding periods, of which the two-years is the highest. However, the overall performance of 2A falls below other value measures almost without exception. Thus, more preferable ratio selections are easily found. The value premiums are slightly amplified with two-year holding periods of portfolios 2A and 2B, and four years with 2A, although being surmounted by many of the other value measures. Even though most of the two-composite measure value premiums are not enhanced with extended holding periods, all of them are positive with a multitude of statistically significant cases.







Table 6 panel C exhibits the performance metrics of three-composite value portfolios for each one- to four-year holding period. When holding a combination of high E/P, B/P, and D/P (3A) stocks for three or four years, the annual returns become slightly higher. Adjusting for risk, the Sharpe ratios for the same portfolio, and same holding periods show slightly better performance. When combining B/P, D/P, and FCF/P (3B), the risk can be slightly reduced with a two-year holding period, and thus the Sharpe of 3B is marginally higher than holding the corresponding value stocks for one year. After adjusting for skewness and kurtosis (SKASR), the enhancements are parallel to the Sharpe ratios, although the portfolio 3A of the three-year holding period loses its statistical significance against the market and comparable growth portfolio. The alphas show that enhancements can be obtained with the same portfolios and holding periods as the Sharpe ratio and SKASR. The same is true for the value premium, although also the alpha spread of the 3C value portfolio is slightly larger with the two-year holding period than the one-year holding period.

In conclusion, the results show that extended holding periods have a somewhat gradually deteriorating effect on long-term portfolio performance up to the maximum used holding period of four years when considering the overall average performance. By examining the value portfolio performance individually, relatively slight enhancements can be obtained with certain portfolios. The best overall performance in relation to all value portfolios and holding periods is obtained with a two-year holding period of high free-cash-flow to price (FCF/P) stocks, which overperformed all the other portfolios in terms of annual returns (14,67%), adjusted Sharpe ratio (0,740) and three-factor alpha (5,47%), all with statistical significance against the market and its comparable growth portfolio. Moreover, the B/P and D/P value portfolios' performance can be slightly enhanced with the two-year holding period. Regarding the composite value measure portfolios, only marginal enhancements can be detected with portfolios 2A, 2D, 3A, and 3B. In comparison, with Leivo and Pätäri (2009), they found that while two- and three-year holding periods had no enhancing effect, a further extension to four years recovered the performance to the levels of one- and two-year holding periods in terms of average performance. Interestingly, their maximum holding period length of five years had even further improvements to the four-year results. However, the five-year holding period was left out of this study. Since the sample

period in this study is different, and different value measures are used as portfolio selection criteria, they are a probable explanation for the deviating results.

Since the omitted transaction costs and taxes would have a more deteriorating effect on portfolios with a higher reformation frequency, the performance of the longer holding period portfolios would be amplified in comparison. Thus, taking these factors into account, the alternate results would potentially seem more favorable for the longer holding periods. The effect would be the strongest in the portfolios, which already had some enhancing effects over the one-year strategy. Moreover, portfolios with only a subtle underperformance would potentially change their outcome to be more profitable than the one-year strategy.

### **4.3 Size effect**

Lastly, the size effect in the value and growth portfolios is examined during the full sample period, bull markets, and bear markets. As Banz (1981) and Reinganum (1981) originally argued, E/P anomaly is another form of size anomaly. The aim is to find if similar implications are found in the Finnish stocks and multiple other value measures. As described in chapter 3.5, a two-factor model is constructed to compare the alphas with the three-factor model alphas. In the two-factor model, the Small-Minus-Big (SMB) factor is omitted from the parameter estimation, and thus, any significant differences between the three-factor and two-factor alphas would indicate that firm size is a significant factor in explaining the portfolio returns: A positive difference is an indication of the small firm effect, and vice versa, a negative difference indicates large firm effect.

Table 7 shows the alpha differences for all value and growth portfolios during the full period, bull markets, and bear markets. All of the full sample period differences of the value portfolios are slightly positive, indicating that the effect of small firms could have some explanatory power in explaining the value anomaly. On the contrary, the majority of the growth portfolios show negative differences, suggesting an inverse of the small firm effect.

**Table 7. Three-factor versus two-factor alphas for measuring size effect in value & growth portfolios.** The table presents the results for the full sample period (1996-2020), bull markets and bear markets. The 3-factor model utilizes the market-factor, HML-factor, and SMB-factor in parameter estimation, while the SMB-factor is omitted from the 2-factor alphas. A positive difference indicates small firm effect, whereas a negative difference indicates large firm effect.

3-factor vs. 2-factor alpha spreads						
Variable	Full sample period		Bull market periods		Bear market periods	
	$\alpha$ -spread	(sign.)	$\alpha$ -spread	(sign.)	$\alpha$ -spread	(sign.)
<b>E/P</b>						
Q1	0,10 %	(0,962)	-0,08 %	(0,979)	0,24 %	(0,955)
Q4	-0,08 %	(0,972)	0,05 %	(0,986)	-0,15 %	(0,978)
<b>B/P</b>						
Q1	0,03 %	(0,988)	0,00 %	(0,999)	0,83 %	(0,868)
Q4	-0,17 %	(0,954)	0,11 %	(0,973)	-0,67 %	(0,936)
<b>D/P</b>						
Q1	0,01 %	(0,995)	-0,02 %	(0,994)	-0,31 %	(0,943)
Q4	-0,04 %	(0,990)	0,01 %	(0,997)	-0,22 %	(0,969)
<b>FCF/P</b>						
Q1	0,08 %	(0,969)	-0,01 %	(0,997)	1,50 %	(0,755)
Q4	0,05 %	(0,981)	0,00 %	(0,999)	1,19 %	(0,811)
<b>EBITDA/EV</b>						
Q1	0,12 %	(0,957)	-0,06 %	(0,985)	1,37 %	(0,794)
Q4	-0,10 %	(0,975)	0,05 %	(0,989)	-0,32 %	(0,959)
<b>2A (E/P B/P)</b>						
Q1	0,06 %	(0,977)	-0,01 %	(0,997)	1,12 %	(0,812)
Q4	-0,12 %	(0,963)	0,12 %	(0,972)	0,85 %	(0,871)
<b>2B (B/P EBITDA/EV)</b>						
Q1	0,09 %	(0,969)	-0,03 %	(0,992)	1,32 %	(0,815)
Q4	-0,09 %	(0,972)	0,07 %	(0,982)	0,04 %	(0,995)
<b>2C (D/P FCF/P)</b>						
Q1	0,09 %	(0,963)	-0,05 %	(0,984)	0,41 %	(0,929)
Q4	0,00 %	(0,999)	0,01 %	(0,997)	0,63 %	(0,909)
<b>2D (D/P EBITDA/EV)</b>						
Q1	0,08 %	(0,969)	-0,05 %	(0,985)	0,42 %	(0,931)
Q4	-0,11 %	(0,973)	0,07 %	(0,986)	-0,08 %	(0,991)
<b>2E (FCF/P EBITDA/EV)</b>						
Q1	0,08 %	(0,970)	-0,02 %	(0,995)	1,30 %	(0,796)
Q4	0,01 %	(0,998)	-0,02 %	(0,994)	-0,69 %	(0,909)
<b>3A (E/P B/P D/P)</b>						
Q1	0,07 %	(0,971)	-0,02 %	(0,993)	0,94 %	(0,855)
Q4	-0,14 %	(0,957)	0,16 %	(0,962)	1,66 %	(0,758)
<b>3B (B/P D/P FCF/P)</b>						
Q1	0,06 %	(0,976)	-0,04 %	(0,987)	0,10 %	(0,986)
Q4	-0,13 %	(0,963)	0,12 %	(0,973)	0,34 %	(0,963)
<b>3C (B/P D/P EBITDA/EV)</b>						
Q1	0,06 %	(0,978)	-0,03 %	(0,993)	0,56 %	(0,914)
Q4	-0,08 %	(0,978)	0,07 %	(0,984)	0,39 %	(0,953)
<b>3D (B/P FCF/P EBITDA/EV)</b>						
Q1	0,09 %	(0,965)	-0,03 %	(0,993)	1,50 %	(0,768)
Q4	-0,07 %	(0,979)	0,05 %	(0,990)	-0,58 %	(0,933)
<b>3E (D/P FCF/P EBITDA/EV)</b>						
Q1	0,14 %	(0,950)	-0,09 %	(0,975)	0,49 %	(0,924)
Q4	-0,03 %	(0,991)	0,01 %	(0,998)	-0,39 %	(0,950)

However, none of the differences are close to statistical significance since all the p-values are over 0,9. Concurrently, the widest positive alpha spread is only 0,14 %, and the widest negative spread is -0,17 %. The SMB betas of the value portfolios (appendix 1) are all positive, suggesting a positive relationship between the excess returns of small firms and portfolio returns. However, most SMB betas of the value and growth portfolios show no statistical significance either, apart from the value portfolio 3E. Thus, the excess returns of small firms over large firms are not a quite strong factor in explaining the full sample period returns.

The differences during the bull market periods are mostly reversed since most of the alpha differences are negative for the value portfolios and positive for the growth portfolios. Parallel to the full sample period, all of the differences are close to zero, and the p-values are over 0,9, indicating that the size effect does not explain the returns of the value and growth portfolios during periods of market upswings either.

During the bear market periods, the alpha spreads are slightly larger, of which 1,66 % is the widest, and 0,04 % is the narrowest. In the case of value portfolios, 14/15 instances are positive, while 8/15 of the growth portfolios are positive, and the rest are negative. Although the implications of the size anomaly affecting the performance of value portfolios are stronger during times of economic distress, the results show no statistical significance in any of the portfolios, the smallest p-value being only 0,755. Thus, it is evident that the returns of the value portfolios are not determined by firm size during any specific time period.

In conclusion, although the positive alpha spreads and positive SMB betas of the value portfolios would imply the existence of size effect, the results side with Pätäri and Leivo (2009): The conclusion is that size effect does not explain the returns of value portfolios in the Finnish stock markets, contradicting with the original evidence from the U.S. markets suggested by Banz (1981) and Reinganum (1981). While Pätäri and Leivo (2009) also found larger positive alpha differences in value portfolios during the bear markets, the results were not statistically significant in either case.

## 5. Conclusions

This thesis examines the impact of holding period length, composite value measures, and firm size on value strategy performance in the Finnish stock markets during 1996-2020. The valuation ratios utilized in portfolio composition are E/P, B/P, D/P, FCF/P, and EBITDA/EV. In addition, a total of 10 composite value measure portfolios are formulated, of which five are two-ratio combinations, and the remaining five are three-ratio combinations. The portfolios' performance is evaluated with average annual returns and three risk-adjusted metrics - Sharpe ratios, Adjusted Sharpe ratios (SKASR), and three-factor alphas. Statistical testing is applied for the risk-adjusted metrics to enhance the robustness of the results.

While the risk-based theories presented in chapter 2.3.2 would suggest that value premium is a cause of the higher risk of value stocks, the results exhibit the complete opposite: The overperformance of the value portfolios can be – in most cases – explained by the substantially lower risk of value stocks compared to growth stocks and the market.

### 5.1 Summary

The first objective of this study was: *Do value portfolios outperform the market and the growth portfolios consistently?* Since the first studies published by Nicholson (1960) and Basu (1977), the existence of value anomaly has been detected globally by numerous academics ever since (e.g., Fama & French, 1992; 1998; 2012; Lakonishok et al., 1994; Pätäri & Leivo, 2009). Contributing to the existing literature, this study reinforces the previous studies with the evidence of the value anomaly in the Finnish markets during the sample period of 5/1996 – 5/2020. Almost without exception, the value portfolios outperform the market and comparable growth portfolios with all of the performance metrics and with numerous statistically significant cases. Considering the overall average performance, the value strategy's total cumulative returns yielded over 16 times (16,2) the original investment. In comparison, the market yielded nearly nine times (8,8) returns, and the growth strategy merely tripled (3,4) the initial investment. Complying with Hackel et al. (2000) and Jokipii and Vähämaa (2006), superior results are obtained with stocks having high free cash flow to price ratio (FCF/P). The FCF/P value portfolio exhibited the best results in terms of

annual returns and most of the risk-adjusted metrics. On the contrary, among the lowest-performing value portfolios is the B/P portfolio, parallel to the evidence of Pätäri and Leivo (2009). This result contradicts with Chan et al. (1991) and Fama and French (1992), who suggested the overperformance of the B/P value portfolio over other measures in the U.S. markets.

However, the value strategy's overperformance is not quite distinguishable during bull market periods since many of the growth portfolios and the market outperform the value portfolios in terms of annual returns. While the overall highest bull market returns are generated with the B/P growth portfolio, the highest value portfolio returns of FCF/P placed seventh among all the portfolios, surpassed by six growth portfolios. However, after taking risk-adjustments into account, most value portfolios seem more favorable than their comparable growth portfolios due to their significantly lower risk profiles. Regardless, most of the results were not statistically significant, and preferable options could be found from some growth portfolios.

The long-term overperformance of the value strategy is emphasized during bear market periods since all the value portfolios outperform the market and every growth portfolio with all of the performance metrics due to losing less value and having substantially lower risk. Although the best yield would certainly be obtained with risk-free investments during market descents, timing the market can be a troublesome endeavor, potentially causing more losses than merely following the same strategy in the long-term.

The second objective of the study was: *Do composite value measures enhance the performance of the value portfolios?* Parallel to Leivo and Pätäri (2009), Dhatt et al. (1999), and Dhatt et al. (2004), some enhancements can be obtained using combinations of two or more valuation ratios together. While finding the single best composite value measure is not fairly obvious, employing the FCF/P and EBITDA/EV measures could provide improvements most consistently. While the individual FCF/P value portfolio dominates other portfolios in the long-term, some of its combinatory counterparts could surpass its risk-adjusted performance during bull and bear markets. Generally, most of the composite value measures successfully amplified the value premiums of individual measures regardless of the time period, and therefore, the

overperformance value portfolios against growth portfolios are most emphasized with ratio combinations.

The third objective of the study was: *Do extended holding periods enhance the performance of the value portfolios?* The average statistics implied that the performance against the market decreased gradually with every extended year up to the maximum of four years used in this study. Compared to the growth portfolios, the performance was best enhanced with a two-year holding period. The overall best-performing value portfolio's (FCF/P) performance could be further enhanced with a holding period of two-years, overperforming all other portfolios with different holding periods in terms of annual returns, SKASR, and alpha statistics. Only marginal enhancements could be obtained by extending the holding periods of portfolios 2A, 2D, 3A, and 3B, although being surpassed by many of the other one-year holding period value portfolios. While the enhancements with extended holding periods mostly ratio-specific and relatively insignificant, it has to be emphasized that since transaction costs and taxes were omitted from this study, the costs would deteriorate the returns of the portfolios with higher portfolio formation frequency more. Thus, if the costs were accounted for, the portfolios with longer holding periods would potentially become more favorable than the portfolios with shorter holding periods.

The last objective of this thesis was: *Does size anomaly explain the returns of value portfolios?* While Banz (1981) and Reinganum (1981) reported evidence on the small firm effect in U.S. stocks, later studies have not supported the anomaly's existence (Dimson & Marsh, 1999; Chan et al., 2000). This study shows that after adjusting for firm size, the performance of the value portfolios did not change drastically during the full sample period, the bull, or the bear markets. Although the indications were the strongest during the bear market periods, showing some small firm effect, all the portfolios' high – thus insignificant – p-values showed no robust evidence of size anomaly within neither the value nor growth portfolios. This result is parallel to Pätäri and Leivo (2009), who also found slightly stronger indications of small firm effect in the value portfolios during bear market periods in Finnish stocks, though without any statistically significant instances. Thus, we can state with fair confidence that small firms are not a decisive factor in explaining the overperformance of the value portfolios in the Finnish stock markets.

## 5.2 Concluding remarks

Potential real-life application of the strategies presented in this study should be implemented with caution. While the historical evidence has been quite robust in favor of the value strategy, history is not a guarantee of future revenue. Nevertheless, the relatively long sample period of 24 years utilized in this study provides reliability to the results due to the multiple different economic cycles that fit in the data sample. The statistical testing applied for the risk-adjusted performance metrics strengthen the robustness of the results. Moreover, the equally weighted market index constructed from the data sample appeared as a more competent index for the quartile portfolios than conventional market indices due to having higher average annual returns and lower volatility. Thus, the market's comparably strong performance could be considered a substitute for the transaction costs that were otherwise omitted from the quartile portfolios. However, when comparing the one-year holding period portfolios with the extended holding periods, the portfolios with higher portfolio formation frequency would have been exposed to more transaction costs. Therefore, for the purpose of further studies that compare different holding period lengths, transaction costs could be included to provide results that would correspond to a real-life scenario.

Further additional applications would include more sophisticated pricing models: Since Asness et al. (2018) were able to detect consistent size anomaly after accounting for the Quality-Minus-Junk factor (QMJ) in their pricing model, further studies could adopt the concerning factor in the attempt to find similar results in the Finnish markets. Moreover, while this study used a simple ranking system in the composite value measure portfolio formation, different multiple criteria decision-making methods could be applied, such as TOPSIS (Technique for Order Preference by Similarity to an Ideal Solution) or AHP (Analytic Hierarchy Process). These methods have shown to be capable of separating future outperforming stocks relatively well in the U.S. stock markets by Pätäri, Karell, Luukka, and Yeomans (2018). Thus, implementing those in the Finnish stock markets could also be considered.

## References

- Acharya, V. & Pedersen, L. (2005) Asset pricing with liquidity risk. *Journal of Financial Economics*. 77, 2, 375-410.
- Asness, C., Frazzini, A., Israel, R., Moskowitz, T. & Pedersen, L. (2018) Size matters, if you control your junk. *Journal of Financial Economics*. 129, 3, 479-509.
- Athanassakos, G. (2011) The Performance, Pervasiveness and Determinants of Value Premium in Different us Exchanges: 1985–2006. *Journal of Investment Management* 9, 33–73.
- Ball, R. (1992) The earnings-price anomaly. *Journal of Accounting and Economics*. 15, 2-3, 319-345.
- Banz, R. (1981) The relationship between return and market value of common stocks. *Journal of Financial Economics*. 9, 1, 3-18.
- Banz, R. & Breen, W. (1986) Sample-dependent results using accounting and market data: some evidence. *Journal of Finance*. 41, 779–793.
- Barberis, N., Shleifer, A. & Vishny, R. (2005) A model of investor sentiment. *Advances in behavioral finance*. 2, 423-459.
- Bartov, E. & Kim, M. (2004) Risk mispricing and value investing. *Review of Quantitative Finance and Accounting* 23: 353–376.
- Basu, S. (1977) Investment performance of common stocks in relation to their price-earnings ratios. *Journal of Finance*. 32, 663-682.
- Basu, S. (1983) The relationship between earnings' yield, market value and return for NYSE common stocks: Further evidence. *Journal of financial economics*. 12, 129-156.
- Berk, J. (1995) A critique of size-related anomalies. *The Review of financial studies*. 8, 275-286.
- Bernard, V. & Stober, T. (1989). The nature and amount of information in cash flows and accruals. *Accounting Review*. 64, 624–652.

Bird, R. & Whitaker, J. (2003) The performance of value and momentum investment portfolios: recent experience in the major European markets. *Journal of Asset Management*. 4, 221–246.

Black, F. (1993) Beta and return. *Journal of Portfolio Management* 20: 8–18.

Black, F., Jensen M. & Scholes M. (1972) The capital asset pricing model: some empirical tests. *Studies in the Theory of Capital Markets*. 79-121.

Black, F. & Scholes, M. (1974) The effects of dividend yield and dividend policy on common stock prices and returns. *Journal of Financial Economics*. 1, 1–22.

Blume, M. (1980) Stock returns and dividend yields: Some more evidence. *Review of Economics and Statistics*. 62, 567–577.

Bodie, Z., Kane A. & Marcus, A., (2001) *Investments* 5th Edition. McGraw-Hill.

Brennan, M. (1970) Taxes, market valuation and corporate financial policy. *National Tax Journal*. 23, 417–427.

Callao, S. & Jarne, J. (2010) Have IFRS affected earnings management in the European Union? *Accounting in Europe*. 7, 159–189.

Campbell, J., Hilscher, J. & Szilagyi, J. (2008) In search of distress risk. *Journal of Finance*. 63, 2899–2939.

Bird, R. & Casavecchia, L. (2007) Value enhancement using momentum indicators: The European experience. *International Journal of Managerial Finance*. 3, 229–262.

Chan, K., Chan, L., Jegadeesh, N. & Lakonishok, J. (2006) Earnings quality and stock returns. *Journal of Business*. 79, 1041–1082.

Chan, L., Hamao, Y. & Lakonishok, J. (1991) Fundamentals and stock returns in Japan. *The journal of finance*. 46, 1739-1764.

Chan, L. K.C., Karceski, J., and Lakonishok, J. (2000). New paradigm or same old hype in equity investing? *Financial Analysts Journal* 56, 23–36.

Chan, L. & Lakonishok, J. (2004) Value and growth investing: review and update. *Financial Analysts Journal*. 60, 71–86.

Chan, R. & Lui, B. (2011) EV/EBIT ratio: the best of both worlds. *Better Investing* 60: 27–28.

Connor, G. & Sehgal, S. (2001). Tests of the Fama and French model in India. Discussion Paper 37. Financial Markets Group, London School of Economics.

Conrad, J. (1995) The New Finance: The Case against Efficient Markets. *The Journal of Finance*. 50, 1348-1352.

Conrad, J., Cooper, M. & Kaul, G. (2003) Value versus glamour. *Journal of Finance*. 58, 1969–1995.

Cooper, M., Dimitrov, O., Rau, P. (2001) A Rose.com by Any Other Name. *The Journal of Finance*. 56, 2371-2388.

Cornish, E. & Fisher R. (1937) Moments and cumulants in the specification of distributions. *Review of the International Statistical Institute*. 307-320.

Daniel, K., Hirshleifer, D. and Subrahmanyam, A. (2001) Overconfidence, arbitrage, and equilibrium asset pricing. *Journal of Finance*. 56, 921–965.

Davis, J. (1994) The cross-section of realized stock returns: the pre-compustat evidence. *Journal of Finance*. 49, 1579–1593.

De Bondt, W. & Thaler, R. (1985) Does the stock market overreact? *Journal of Finance*. 40, 793–805.

Denis, D.J., Denis, D.K. & Sarin, A. (1994) The information content of dividend changes: cash flow signaling, overinvestment and dividend clienteles. *Journal of Financial and Quantitative Analysis*. 29, 567–587.

Desai, H., Rajgopal, S. & Venkatachalam, M. (2004) Value-glamour and accruals mispricing: one anomaly or two? *Accounting Review*. 79, 355–385

Dhatt, M., Kim, Y. & Mukherji, S. (1999) The value premium for small-capitalization stocks. *Financial Analysts Journal*. 55, 60–68.

Dhatt, M., Kim, Y. & Mukherji, S. (2004) Can composite value measures enhance portfolio performance? *Journal of Investing*. 13, 42–48.

Dielman, T. & Oppenheimer, H. (1984) An examination of investor behavior during periods of large dividend changes. *Journal of Financial and Quantitative Analysis*. 19, 197–216.

Dimson, E. & Marsh, P. (1986) Event Study Methodologies and the size effect. *Journal of Financial Economics*. 17, 113-142

Dimson, E. & Marsh, P. (1999) Murphy's Law and Market Anomalies. *Journal of Portfolio Management*. 25, 53-69.

Doukas, J., Kim, C. & Pantzalis, C. (2004) Divergence of opinions and the performance of value stocks. *Financial Analysts Journal*. 60, 55–64.

Eling, M. (2008) Does the Measure Matter in the Mutual Fund Industry? *Financial Analyst Journal*. 64, 54–66.

Eling, M. & Schuhmacher, F. (2007) Does the choice of performance measure influence the evaluation of hedge funds? *Journal of Banking & Finance*. 31, 2632-2647.

Elton, E., Gruber, M. & Rentzler, J. (1983) A simple examination of the empirical relationship between dividend yields and deviations from the CAPM. *Journal of Banking and Finance*. 7, 135–146.

Fama, E. (1970) Efficient Capital Markets - A Review of Theory and Empirical Work. *The Journal of Finance* 25, 383–417.

Fama, E. (1991) Efficient Capital Markets II. *The Journal of Finance* 46, 1575-1617.

Fama, E. & French, K. (1988) Dividend yields and expected stock returns. *Journal of Financial Economics* 22, 3–25.

Fama, E. & French, K. (1992) The Cross-section of Expected Stock Returns. *The Journal of Finance*. 47, 2, 427-465.

Fama, E. & French, K. (1993). Common Risk Factors in the Returns on Stocks and Bonds, *Journal of Financial Economics* 33, 3–56.

Fama, E. & French, K. (1995) Size and Book-to-Market Factors in Earnings and Returns. *The Journal of Finance*. 50, 131-155.

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

Fama, E. & French, K. (1998) Value versus Growth: The International Evidence. *The Journal of Finance*. 53, 1975-1999.

Fama, E. & French, K. (2004). The Capital Asset Pricing Model: Theory and Evidence. *Journal of Economic Perspectives*. 18, 25-46.

Fama, E. & French, K. (2012) Size, value, and momentum in international stock returns. *Journal of Financial Economics*. 105, 457-472.

Findlay, M. & Williams, E. (2000) A Fresh Look at the Efficient Market Hypothesis: How the Intellectual History of Finance Encouraged a Real "Fraud-on-the-Market". *Journal of Post Keynesian Economics*. 23, 181-199.

Garleanu, N., Panageas, S., Yu, J. (2012) Technological Growth and Asset Pricing. *The Journal of Finance*. 67, 1265-1292.

Graham, B. (1949) *The Intelligent Investor*. New York, NY: Harper & Row Publishers.

Graham, B. (2003) *The Intelligent Investor – Revised Edition*. New York, NY: Harper & Row Publishers.

Gray, W. & Vogel, J. (2012). Analyzing valuation measures: a performance horserace over the past 40 years. *Journal of Portfolio Management*. 39, 112-121.

Guerard, J. Jr., Xu, G. & Gultekin, M. (2012) Investing with momentum: the past, present, and future. *Journal of Investing*. 21, 68–80.

Hackel, K., Livnat, J. & Rai, A. (2000) A Free Cash Flow Investment Anomaly. *Journal of Accounting, Auditing and Finance*. 15, 1-24.

van der Hart, J., deZwart, G. & van Dijk, D. (2003) Stock selection strategies in emerging markets. *Journal of Empirical Finance*. 10, 105–132.

Harvey, C. & Siddique, A. (2000). Conditional skewness in Asset Pricing Tests. *Journal of Finance*. 55, 1263-1296.

Hou, K. & Moskowitz, T. (2005) Market Frictions, Price Pelay, and the Cross-Section of Expected Returns. *The Review of Financial Studies*. 18, 981-1020.

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

Israel, R. & Moskowitz, T.J. (2013) The Role of Shorting, Firm Size, and Time on Market Anomalies. *Journal of Financial Economics*. 108, 275–301.

Israelsen, C. (2005) A Refinement to the Sharpe Ratio and Information Ratio. *Journal of Asset Management* 5, 423–427.

Jegadeesh, N. & Titman, S. (1993) Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *The Journal of Finance*. 48, 65–91

Jobson, J. & Korkie, B. (1981). Performance Hypothesis Testing with the Sharpe and Treynor Measures. *The Journal of Finance*. 36, 889-908.

Jokipii, A. & Vähämaa, S. (2006) The Free Cash Flow Anomaly Revisited: Finnish Evidence. *Journal of Business & Accounting*. 33, 961-978.

Jones, C.P. (2007) *Investments* 10th edition. John Wiley & Sons, Inc.

Keim, D. (1985) Dividend yields and stock returns: Implications of abnormal January returns. *Journal of Financial Economics*. 14, 473–489.

Lakonishok, J., Shleifer, A., Vishny, R. (1994) Contrarian Investment, Extrapolation, and Risk. *The Journal of Finance*. 49, 5, 1541-1578.

Leivo, T. & Pätäri, E. (2009) The impact of holding period length on value portfolio performance in the Finnish stock markets. 8, 71-86.

Leivo, T., Pätäri, E. & Kilpiä, I. (2009) Value Enhancement Using Composite Measures: The Finnish Evidence. *International Research Journal of Finance and Economics*. 33, 7-30.

Leivo, T. & Pätäri, E. (2011) Enhancement of value portfolio performance using momentum and the long-short strategy: the Finnish evidence. *Journal of Asset Management*. 11, 401–416.

Lehn, K. & Poulsen, A. (1989) Free Cash Flow and Stockholder Gains in Going Private Transactions. *The Journal of Finance*. 44, 771-787.

Levis, M. (1989) Stock market anomalies: a re-assessment based on UK evidence. *Journal of Banking and Finance*. 13, 675–696.

Lindström, K. (2007) *Vaurastu arvo-osakkeilla*. Helsinki, Talentum.

Lintner, J. (1965) The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics*. 47, 13-37.

Litzenberger, R. & Ramaswamy, K. (1979) The effect of personal taxes and dividends on capital asset prices: Theory and empirical evidence. *Journal of Financial Economics*. 7, 163–195.

Litzenberger, R. & Ramaswamy, K. (1982) The effects of dividends on common stock prices. *Journal of Finance*. 37, 429–443.

Lo, A. (2002) The statistics of Sharpe ratios. *Financial Analysts Journal*. 58, 36-52.

Loughran, T. & Wellman, J. (2011) New evidence on the relation between the enterprise multiple and average stock returns. *Journal of Financial and Quantitative Analysis*. 46, 1629–1650.

Malkiel, B. (2003) The Efficient Market Hypothesis and Its Critics. *Journal of Economic Perspectives*. 17, 59-82.

Markowitz, H. & Xu, G. (1994) Data mining corrections. *Journal of Portfolio Management*. 21, 60–69.

McWilliams, J. (1966) Price, earnings and P-E ratios. *Financial Analyst Journal*. 22, 137–142.

- Memmel, C. (2003) Performance Hypothesis Testing with the Sharpe Ratio. *Finance Letters* 1, 21-23.
- Merton, R. (1987) A Simple Model of Capital Market Equilibrium with Incomplete Information. *The Journal of Finance*. 42, 483-510.
- Miles, D. & Timmermann, A. (1996) Variations in expected stock returns: evidence on the pricing of equities from a cross-section of UK companies. *Economica*. 63, 369–382.
- Naranjo, A., Nimalendran, M. & Ryngaert, M. (1998) Stock returns, dividend yields, and taxes. *Journal of Finance*. 53, 2029–2057.
- Nicholson, S. (1960) Price-earnings Ratios, *Financial Analysts Journal*. 16, 43–45.
- Petkova, R. & Zhang, L. (2005) Is value riskier than growth? *Journal of Financial Economics*. 78, 187–202.
- Piotroski, J. (2000) Value investing: the use of historical financial statement information to separate winners from losers. *Journal of Accounting Research*. 38, 1–41.
- Pätäri, E. (2011) Does the risk-adjustment method matter at all in hedge fund rankings?. *International Research Journal of Finance and Economics*. 75, 69–99.
- Pätäri, E. & Leivo, T. (2009) Performance of the Value Strategies in the Finnish Stock Markets. *Journal of Money, Investment and Banking*. 8, 5-24.
- Pätäri, E. & Leivo, T. (2017) A closer look at value premium: Literature review and synthesis. *Journal of Economic Surveys*. 31, 79–168.
- Pätäri, E., Karell, V., Luukka, P. & Yeomans, J. (2018) Comparison of the multicriteria decision-making methods for equity portfolio selection: The U.S. evidence. *European Journal of Operational Research*. 265, 655-672.
- Reinganum, M. (1981). The misspecification of capital asset pricing empirical anomalies based on earnings yield and market values. *Journal Financial Economics* 9, 19–46.

Roll, R. (1977) A Critique of the Asset Pricing Theory's Tests. Part I: On Past and Potential Testability of the Theory. *Journal of Financial Economics*. 4, 129-176.

Roll, R. (1995) Style return differentials: illusions, risk premiums, or investment opportunities. *The Handbook of Equity Style Investment*. New Hope, USA: Frank J.

Rosenberg, B., Reid, K. & Lanstein, R. (1985) Persuasive Evidence of Market Inefficiency. *Journal of Portfolio Management*. 11, 9–16.

Rozeff, M. (1984) Dividend yields are equity risk premiums. *Journal of Portfolio Management*. 11, 68–75.

Sant, R. & Cowan, A. R. (1994) Do dividends signal earnings? The case of omitted dividends. *Journal of Banking and Finance*. 18, 1113–1133.

Schultz, P. (1983) Transaction costs and the small firm effect: A comment. *Journal of Financial Economics*. 12, 81-88.

Schwert, G. (2002) Anomalies and Market Efficiency. NBER Working Paper Series. 9277, 1–54.

Sharpe, W. (1964) Capital asset prices: a theory of market equilibrium under conditions of risk. *Journal of Finance*. 19, 425-442.

Sharpe, W. (1966) Mutual fund performance. *The Journal of Business*. 39, 119-138.

Sharpe, W. (1994) The Sharpe ratio. *Journal of Portfolio Management*. 21, 49-59.

Shleifer, A. & Vishny, R. (1997) The Limits of Arbitrage. *The Journal of Finance*. 52, 1, 35-55.

Stattman, D. (1980) Book values and stock returns. *A Journal of Selected Papers*. 4, 25-45.

Stoll, H. & Whaley, R. (1983) Transaction costs and the small firm effect. *Journal of Financial Economics*. 12, 57-79.

Van Dijk, M. (2011) Is size dead? A review of the size effect in equity returns. *Journal of Banking & Finance*. 35, 3263-3274.

Vassalou, M. & Xing, Y. (2004) Default Risk in Equity Returns. *The Journal of Finance*. 59, 831,868

Zhang, L. (2005). The value premium *Journal of Finance*. 60, 67-103

Zivney, T. & Thompson, D. (1989). THE EFFECT OF MARKET PROXY REBALANCING POLICIES ON DETECTING ABNORMAL PERFORMANCE. *Journal of Financial Research*. 12, 293-299.

## APPENDICES

**Appendix 1. Results of the Fama-French 3-factor model.** *The model results for all the value and growth portfolios are presented.  $\alpha$  shows the alphas determined by the model,  $\beta$ -MR shows the portfolios' risk in relation to the market,  $\beta$ -HML shows the betas of the High-Minus-Low value factor, and the  $\beta$ -SMB shows the betas of the Small-Minus-Big size factor. The goodness of fit statistics Adjusted  $R^2$  are shown in the last column. Statistical significances are presented in parentheses—bolded significances with a star (\*) notation show statistical significance under 5 % (0,05) risk level.*

Variable	3-factor alpha statistics		3-factor beta statistics						Adj. $R^2$
	$\alpha$	(sign.)	$\beta$ -MR	(sign.)	$\beta$ -HML	(sign.)	$\beta$ -SMB	(sign.)	
<b>E/P</b>									
Q1	2,74 %	(0,064)	0,944	<b>(0,000)*</b>	-0,005	(0,779)	0,029	(0,100)	0,862
Q4	-4,48 %	<b>(0,010)*</b>	1,239	<b>(0,000)*</b>	-0,015	(0,512)	-0,026	(0,230)	0,872
<b>B/P</b>									
Q1	1,61 %	(0,330)	0,958	<b>(0,000)*</b>	0,027	(0,214)	0,010	(0,614)	0,835
Q4	-1,50 %	(0,468)	1,177	<b>(0,000)*</b>	0,014	(0,615)	-0,051	<b>(0,046)*</b>	0,821
<b>D/P</b>									
Q1	4,43 %	<b>(0,001)*</b>	0,791	<b>(0,000)*</b>	0,007	(0,656)	0,003	(0,845)	0,852
Q4	-5,66 %	<b>(0,009)*</b>	1,341	<b>(0,000)*</b>	-0,023	(0,426)	-0,011	(0,674)	0,836
<b>FCF/P</b>									
Q1	5,17 %	<b>(0,000)*</b>	0,919	<b>(0,000)*</b>	-0,004	(0,813)	0,023	(0,174)	0,866
Q4	-4,28 %	<b>(0,008)*</b>	1,032	<b>(0,000)*</b>	0,014	(0,526)	0,016	(0,420)	0,854
<b>EBITDA/EV</b>									
Q1	4,05 %	<b>(0,012)*</b>	0,850	<b>(0,000)*</b>	-0,021	(0,307)	0,036	(0,065)	0,813
Q4	-6,23 %	<b>(0,005)*</b>	1,231	<b>(0,000)*</b>	0,035	(0,243)	-0,030	(0,276)	0,813
<b>2A (E/P B/P)</b>									
Q1	1,39 %	(0,353)	0,849	<b>(0,000)*</b>	0,024	(0,212)	0,018	(0,310)	0,832
Q4	-1,55 %	(0,399)	1,102	<b>(0,000)*</b>	-0,008	(0,752)	-0,036	(0,109)	0,835
<b>2B (B/P EBITDA/EV)</b>									
Q1	5,10 %	<b>(0,002)*</b>	0,847	<b>(0,000)*</b>	0,012	(0,562)	0,026	(0,183)	0,810
Q4	-3,87 %	<b>(0,031)*</b>	1,142	<b>(0,000)*</b>	0,004	(0,870)	-0,027	(0,223)	0,852
<b>2C (D/P FCF/P)</b>									
Q1	4,71 %	<b>(0,000)*</b>	0,800	<b>(0,000)*</b>	0,005	(0,753)	0,024	(0,114)	0,861
Q4	-6,69 %	<b>(0,000)*</b>	1,231	<b>(0,000)*</b>	-0,010	(0,656)	-0,001	(0,951)	0,874
<b>2D (D/P EBITDA/EV)</b>									
Q1	4,81 %	<b>(0,001)*</b>	0,788	<b>(0,000)*</b>	0,014	(0,457)	0,024	(0,177)	0,821
Q4	-6,95 %	<b>(0,004)*</b>	1,341	<b>(0,000)*</b>	0,023	(0,473)	-0,035	(0,251)	0,810
<b>2E (FCF/P EBITDA/EV)</b>									
Q1	4,92 %	<b>(0,001)*</b>	0,866	<b>(0,000)*</b>	-0,039	<b>(0,035)*</b>	0,022	(0,192)	0,849
Q4	-7,11 %	<b>(0,000)*</b>	1,156	<b>(0,000)*</b>	-0,004	(0,877)	0,002	(0,922)	0,861
<b>3A (E/P B/P D/P)</b>									
Q1	3,50 %	<b>(0,014)*</b>	0,821	<b>(0,000)*</b>	0,016	(0,371)	0,021	(0,208)	0,838
Q4	-2,73 %	(0,132)	1,214	<b>(0,000)*</b>	0,005	(0,834)	-0,042	(0,063)	0,853
<b>3B (B/P D/P FCF/P)</b>									
Q1	4,68 %	<b>(0,001)*</b>	0,805	<b>(0,000)*</b>	0,004	(0,822)	0,017	(0,297)	0,841
Q4	-4,85 %	<b>(0,017)*</b>	1,289	<b>(0,000)*</b>	0,013	(0,634)	-0,040	(0,114)	0,839
<b>3C (B/P D/P EBITDA/EV)</b>									
Q1	4,45 %	<b>(0,004)*</b>	0,821	<b>(0,000)*</b>	0,018	(0,357)	0,018	(0,342)	0,805
Q4	-5,74 %	<b>(0,006)*</b>	1,266	<b>(0,000)*</b>	0,013	(0,654)	-0,024	(0,347)	0,835
<b>3D (B/P FCF/P EBITDA/EV)</b>									
Q1	3,91 %	<b>(0,008)*</b>	0,869	<b>(0,000)*</b>	-0,014	(0,468)	0,027	(0,127)	0,842
Q4	-6,62 %	<b>(0,001)*</b>	1,186	<b>(0,000)*</b>	0,020	(0,475)	-0,024	(0,362)	0,819
<b>3E (D/P FCF/P EBITDA/EV)</b>									
Q1	4,39 %	<b>(0,004)*</b>	0,790	<b>(0,000)*</b>	-0,017	(0,397)	0,039	<b>(0,032)*</b>	0,804
Q4	-7,18 %	<b>(0,000)*</b>	1,244	<b>(0,000)*</b>	0,029	(0,284)	-0,010	(0,689)	0,838































**Appendix 7. All one-year holding period quartile portfolios during bear market periods – Return, risk, and performance metrics.**

Panels A, B & C show the results for individual value measure, 2-composite measure, and 3-composite measure quartile portfolios, respectively. Statistical significances are presented in parentheses—bolded significances with a star (\*) notation show statistical significance under 5 % (0,05) risk level.

Panel A: Individual value measure portfolios

Variable	Average annual return	Annual volatility	Sharpe ratio statistics			SKAD	SKASR statistics			$\beta$	3-factor alpha statistics			
			Sharpe Ratio	Qi vs. market difference	Q1 vs. Q4 difference		SKASR	Qi vs. market difference	Q1 vs. Q4 difference		$\alpha$	(sign.)	Q1 vs. Q4 $\alpha$ -spread	(sign.)
<b>E/P</b>														
Q1	-27,75 %	17,01 %	-0,047	0,011	0,042	20,40 %	-0,057	0,010	0,035	0,869	1,82 %	(0,551)	11,07 %	<b>(0,029)*</b>
Q2	-26,32 %	18,31 %	-0,048	0,010		21,70 %	-0,057	0,009		0,928	5,30 %	(0,130)		
Q3	-28,51 %	18,79 %	-0,054	0,005		21,29 %	-0,061	0,006		0,973	4,01 %	(0,225)		
Q4	-41,40 %	21,51 %	-0,089	-0,031		22,07 %	-0,091	-0,025		1,092	-9,25 %	<b>(0,025)*</b>		
<b>B/P</b>														
Q1	-28,91 %	17,69 %	-0,051	0,007	0,052	21,02 %	-0,061	0,005	0,047	0,872	0,16 %	(0,963)	3,97 %	(0,573)
Q2	-28,22 %	17,37 %	-0,049	0,009		20,97 %	-0,059	0,007		0,882	1,11 %	(0,718)		
Q3	-31,05 %	18,87 %	-0,059	0,000		22,53 %	-0,070	-0,004		1,016	2,91 %	(0,267)		
Q4	-41,16 %	25,11 %	-0,103	-0,045		26,23 %	-0,108	-0,042		1,203	-3,80 %	(0,534)		
<b>D/P</b>														
Q1	-23,10 %	15,32 %	-0,035	0,023	0,064	18,53 %	-0,043	0,023	0,054	0,779	4,39 %	(0,153)	14,35 %	<b>(0,007)*</b>
Q2	-27,85 %	19,67 %	-0,055	0,004		23,53 %	-0,066	0,001		1,004	6,76 %	(0,082)		
Q3	-33,54 %	19,83 %	-0,067	-0,008		20,15 %	-0,068	-0,001		1,027	-0,41 %	(0,899)		
Q4	-43,57 %	22,75 %	-0,099	-0,041		22,23 %	-0,097	-0,031		1,150	-9,96 %	<b>(0,024)*</b>		
<b>FCF/P</b>														
Q1	-24,19 %	18,42 %	-0,045	0,014	0,024	22,07 %	-0,053	0,013	0,020	0,927	8,52 %	<b>(0,011)*</b>	12,43 %	<b>(0,013)*</b>
Q2	-33,74 %	19,17 %	-0,065	-0,006		22,15 %	-0,075	-0,009		1,033	0,53 %	(0,845)		
Q3	-36,99 %	20,48 %	-0,076	-0,017		20,78 %	-0,077	-0,011		1,041	-4,39 %	(0,264)		
Q4	-34,71 %	19,75 %	-0,069	-0,010		21,05 %	-0,073	-0,007		0,989	-3,91 %	(0,281)		
<b>EBITDA/EV</b>														
Q1	-25,45 %	17,32 %	-0,044	0,014	0,047	20,36 %	-0,052	0,014	0,039	0,840	2,74 %	(0,461)	12,74 %	<b>(0,034)*</b>
Q2	-27,61 %	17,64 %	-0,049	0,010		19,40 %	-0,054	0,013		0,899	2,69 %	(0,423)		
Q3	-29,75 %	20,09 %	-0,060	-0,001		24,01 %	-0,071	-0,005		1,066	7,21 %	<b>(0,013)*</b>		
Q4	-41,80 %	21,89 %	-0,091	-0,033		21,63 %	-0,090	-0,024		1,072	-10,00 %	<b>(0,037)*</b>		
<b>Market</b>	-32,46 %	17,98 %	-0,058			20,40 %	-0,066							
<b>Rf</b>	2,54 %	0,57 %												

Panel B: 2-composite value measure portfolios

Variable	Average annual return	Annual volatility	Sharpe ratio statistics			SKAD	SKASR statistics			$\beta$	3-factor alpha statistics			
			Sharpe Ratio	Qi vs. market difference	Q1 vs. Q4 difference		SKASR	Qi vs. market difference	Q1 vs. Q4 difference		$\alpha$	(sign.)	Q1 vs. Q4 $\alpha$ -spread	(sign.)
<b>2A (E/P B/P)</b>														
Q1	-27,14 %	17,35 %	-0,047	0,011	0,034	20,65 %	-0,056	0,010	0,026	0,857	2,10 %	(0,530)	5,70 %	(0,261)
Q2	-28,98 %	17,74 %	-0,051	0,007		22,03 %	-0,064	0,002		0,918	1,75 %	(0,606)		
Q3	-29,03 %	18,19 %	-0,053	0,006		20,59 %	-0,060	0,006		0,959	2,60 %	(0,354)		
Q4	-37,46 %	21,57 %	-0,081	-0,022		22,02 %	-0,082	-0,016		1,095	-3,60 %	(0,348)		
<b>2B (B/P EBITDA/EV)</b>														
Q1	-24,50 %	17,59 %	-0,043	0,015	0,050	20,40 %	-0,050	0,016	0,041	0,843	4,62 %	(0,243)	10,40 %	(0,083)
Q2	-29,73 %	17,45 %	-0,052	0,006		20,87 %	-0,062	0,004		0,874	-1,17 %	(0,721)		
Q3	-30,46 %	19,21 %	-0,059	0,000		22,86 %	-0,070	-0,003		1,027	4,28 %	(0,163)		
Q4	-40,62 %	22,80 %	-0,093	-0,034		22,46 %	-0,091	-0,025		1,145	-5,78 %	(0,201)		
<b>2C (D/P FCF/P)</b>														
Q1	-24,74 %	17,83 %	-0,044	0,014	0,056	22,65 %	-0,056	0,010	0,041	0,910	7,76 %	<b>(0,018)*</b>	16,45 %	<b>(0,002)*</b>
Q2	-27,35 %	17,46 %	-0,048	0,011		19,58 %	-0,054	0,013		0,896	2,87 %	(0,381)		
Q3	-33,84 %	18,88 %	-0,064	-0,006		20,88 %	-0,071	-0,004		0,996	-1,41 %	(0,593)		
Q4	-43,09 %	23,30 %	-0,100	-0,042		22,51 %	-0,097	-0,031		1,183	-8,69 %	<b>(0,038)*</b>		
<b>2D (D/P EBITDA/EV)</b>														
Q1	-23,30 %	16,25 %	-0,038	0,020	0,070	20,08 %	-0,047	0,019	0,057	0,803	4,91 %	(0,160)	14,07 %	<b>(0,023)*</b>
Q2	-26,60 %	18,40 %	-0,049	0,009		21,14 %	-0,056	0,010		0,929	5,39 %	(0,114)		
Q3	-30,00 %	18,76 %	-0,056	0,002		20,63 %	-0,062	0,004		0,967	1,48 %	(0,673)		
Q4	-44,37 %	24,20 %	-0,107	-0,049		23,42 %	-0,104	-0,038		1,187	-9,16 %	(0,073)		
<b>2E (FCF/P EBITDA/EV)</b>														
Q1	-24,28 %	17,92 %	-0,044	0,015	0,047	21,11 %	-0,051	0,015	0,036	0,896	6,48 %	(0,065)	10,62 %	(0,063)
Q2	-28,88 %	18,14 %	-0,052	0,006		21,75 %	-0,063	0,003		0,925	2,59 %	(0,448)		
Q3	-32,51 %	17,99 %	-0,058	0,000		20,61 %	-0,067	-0,001		0,933	-2,36 %	(0,364)		
Q4	-39,97 %	22,75 %	-0,091	-0,033		21,95 %	-0,088	-0,022		1,162	-4,14 %	(0,351)		
<b>Market</b>	-32,46 %	17,98 %	-0,058			20,40 %	-0,066							
<b>Rf</b>	2,54 %	0,57 %												

Panel C: 3-composite value measure portfolios

Variable	Average annual return	Annual volatility	Sharpe ratio statistics			SKAD	SKASR statistics			$\beta$	3-factor alpha statistics			
			Sharpe Ratio	Qi vs. market difference	Q1 vs. Q4 difference		SKASR	Qi vs. market difference	Q1 vs. Q4 difference		$\alpha$	(sign.)	Q1 vs. Q4 $\alpha$ -spread	(sign.)
<b>3A (E/P B/P D/P)</b>														
Q1	-23,83 %	16,88 %	-0,040	0,018	0,049	20,62 %	-0,049	0,017	0,041	0,822	5,20 %	(0,151)	11,52 %	<b>(0,032)*</b>
Q2	-27,59 %	17,57 %	-0,048	0,010		21,36 %	-0,059	0,007		0,907	3,10 %	(0,347)		
Q3	-30,73 %	18,35 %	-0,056	0,002		20,79 %	-0,064	0,002		0,977	1,19 %	(0,693)		
Q4	-39,90 %	22,48 %	-0,090	-0,031		22,48 %	-0,090	-0,024		1,123	-6,32 %	(0,112)		
<b>3B (B/P D/P FCF/P)</b>														
Q1	-21,81 %	17,38 %	-0,038	0,020	0,076	22,04 %	-0,048	0,018	0,060	0,859	9,42 %	<b>(0,018)*</b>	17,63 %	<b>(0,010)*</b>
Q2	-30,43 %	18,04 %	-0,055	0,003		22,03 %	-0,067	-0,001		0,901	-1,01 %	(0,777)		
Q3	-31,69 %	18,85 %	-0,060	-0,001		20,09 %	-0,064	0,003		0,987	0,90 %	(0,764)		
Q4	-44,77 %	25,46 %	-0,114	-0,056		24,14 %	-0,108	-0,042		1,245	-8,20 %	(0,132)		
<b>3C (B/P D/P EBITDA/EV)</b>														
Q1	-24,20 %	17,24 %	-0,042	0,017	0,063	21,30 %	-0,052	0,015	0,050	0,855	5,70 %	(0,119)	12,61 %	<b>(0,042)*</b>
Q2	-25,32 %	17,15 %	-0,043	0,015		20,09 %	-0,051	0,015		0,842	3,25 %	(0,351)		
Q3	-32,06 %	18,56 %	-0,060	-0,001		21,17 %	-0,068	-0,002		0,986	0,35 %	(0,907)		
Q4	-42,92 %	24,33 %	-0,104	-0,046		23,75 %	-0,102	-0,036		1,204	-6,91 %	(0,164)		
<b>3D (B/P FCF/P EBITDA/EV)</b>														
Q1	-25,79 %	17,67 %	-0,046	0,013	0,061	20,66 %	-0,053	0,013	0,049	0,867	3,50 %	(0,326)	8,05 %	(0,201)
Q2	-26,05 %	18,08 %	-0,047	0,011		23,04 %	-0,060	0,006		0,884	4,46 %	(0,272)		
Q3	-31,37 %	17,82 %	-0,056	0,002		20,36 %	-0,064	0,002		0,927	-1,20 %	(0,687)		
Q4	-42,68 %	24,86 %	-0,106	-0,048		23,98 %	-0,102	-0,036		1,243	-4,54 %	(0,378)		
<b>3E (D/P FCF/P EBITDA/EV)</b>														
Q1	-23,14 %	17,19 %	-0,040	0,019	0,064	21,49 %	-0,050	0,016	0,049	0,857	7,18 %	<b>(0,048)*</b>	16,30 %	<b>(0,007)*</b>
Q2	-26,00 %	18,01 %	-0,047	0,012		20,34 %	-0,053	0,013		0,904	5,36 %	(0,127)		
Q3	-31,79 %	18,43 %	-0,059	0,000		20,86 %	-0,066	0,000		0,973	-0,62 %	(0,813)		
Q4	-44,05 %	23,62 %	-0,104	-0,046		22,51 %	-0,099	-0,033		1,182	-9,13 %	(0,055)		
Market	-32,46 %	17,98 %	-0,058			20,40 %	-0,066							
Rf	2,54 %	0,57 %												