Alternative Value Premia

Master Thesis
Economics and Business Administration
Finance and Investments

Copenhagen Business School

Emil Maltha Rasmussen - 92268
Simon Bo Grønning - 115858

Supervisor: Rasmus Tangsgaard Varneskov - Department of Finance

Number of characters: 214,120 - Number of pages: 107

15-05-2019

Abstract

The asset pricing literature has historically focused on the book-to-market (HML) factor as a driver of the value premium, but several scholars have noted that alternative fundamental-to-price ratios capture value premia comparable to the HML factor. Alternative value premia have in the literature received little attention; hence this study seeks to fill this gap in the literature by analyzing the historical performance of alternative value factors starting from 1963 through 2018, using CRSP and Compustat datasets.

We test whether returns of long-short factors based on alternative value metrics are explained by the five-, q- and six-factor models and investigate how characteristics of alternative value factors deviate from the traditional HML factor using multidimensional sorts of firms based on their fundamentals which allow for the detection of alpha generating subsections. Fama-MacBeth regressions are then used to investigate if past changes in size drive value premia on a value-sorted portfolio level. Lastly, time-dependent factors are analyzed to examine if the value premia are driven by subperiods where the difference between cheap and expensive stocks are particularly large.

Alternative value factors generally deliver higher returns as standalone strategies and these higher returns are often attributed to exposure to profitability risk factors. The cash-flow-from-operations-to-enterprise-value factor generated alpha unaccounted for by the five- and q-factor models which were, as for multiple other value strategies primarily driven by the short-leg and showed a general tendency to intensify when controlling for gross profits to assets. Results are ambiguous and lack enough statistical significance to conclude to what extent, past changes are a driver of alternative value metrics. Timing value strategies based on the relative value spread did not prove to be an effective strategy, suggesting value strategies also generates positive returns in the subperiods where the value spread is small. While factor subsections drive many factor premia, we call for future research to examine how these subsections interact as they hold the potential to improve the existing asset pricing models.

Table of content

1 Introduction

The CAPM which built on the work of Markowitz was introduced in the 1960s and were based on the idea that rational equity market investors only require compensation for undiversifiable risk. In the late 1970s, researchers started discovering return anomalies based on empirical data which could not be explained by the CAPM. These anomalies were believed, by rational market advocates, to arise as investors required compensation for various underlying risks other than market risk. One of these anomalies discovered was the famous value premium, which was based on the discovery by (Stattman, 1980; Rosenberg, Reid, & Lanstein, 1985) that the book-to-market ratio could explain a significant proportion of the variance in the cross-section of stock returns.

In 1993, (Fama & French, 1993) proposed their famous three-factor model that captured most of the anomalies discovered up until then, in which they included the book-to-market value factor. Since then the hunt for new risk factors by academics has been substantial, and a multitude of new factors and anomalies have been discovered leading to (Fama & French, 2015), in 2015 augmenting their three-factor model with two new factors, so the model would again be able to capture most of the cross-sectional variation of returns. While (Fama & French, 2006; Fama & French, 2011) hold that alternatives value metrics also proxy for the underlying value risk premium and that they can essentially substitute one for another, we find the body of literature lacks a comprehensive comparison and description of the similarities and differences between book-to-market and other value premium proxies, and we, therefore, seek to fill this gap.

In recent years, behavioral explanations for why anomalies exist in the cross-section of stock returns have increasingly received attention, and risk-based rational explanations for the value premium among other anomalies have been challenged. (Novy-Marx, 2013)'s recent study challenges the explanation by showing the entire book-to-market premium can be explained by a subsection of firms with high-profit margins. Another recent study by (Gerakos & Linnainmaa, 2018) shows that the entire value premium can be explained by changes in market cap over the prior five years. With limited success, we tested the two propositions jointly to shed light on which subsection may be the driver of the value premium.

Additionally, we also test these propositions for alternative profitable value metrics. Besides investigating various value metrics individually, we also combine these into a compound value metric to see whether we can improve the performance of value premium strategies. (Asness, Liew, Pedersen, & Thapar, 2017) found that the value spread had predictable power of subsequent returns for equities based on book-to-market and alternative value strategies for other asset classes. Based on these findings we developed a simple, implementable strategy to investigate if it is possible to time the value premium for various value metrics in practice using a 10-year rolling historical estimation window.

2 Delimitation

The scope of the study is limited to the investigation of the improvement of value strategies using what we classify as "profitable value" factors as well as the traditional HML factor, as well as "profitable value" factors in combination with (Novy-Marx, 2013) profitability factor.

Our equity universe was limited to US equities on the NYSE, NASDAQ and AMEX exchanges and their return performance based on various fundamental data for the period 1963 to 2018, excluding microcap stocks and financials.

It was also within the scope of the study to discover new "profitable value" factors and combinations of value and quality factors that potentially could deliver risk-adjusted returns in excess of traditional asset pricing models. This also entails sorting firms on multiple dimensions to test whether returns have been driven by a subsection of firms with specific characteristics.

Because we sought to find implementable investment strategies that can outperform these asset pricing models we used a factor construction methodology different from the methodology in (Fama & French, 2015). In our methodology, we formed equal weighted portfolios with annual rebalancing as opposed to the standard methodology used by (Fama & French, 1993) who use value-weighted portfolios which are rebalanced monthly. Value-weighted returns tend to skew portfolios towards larger companies and away from companies with more extreme scores on the formation metrics. Furthermore, we used 20th and 80th percentile NYSE breakpoints for our construction of profitable value factors, (Novy-Marx, 2013)'s profitability factor and the book-to-market factor when combined with other factors.

The reason for using these wider breakpoints was to analyze the performance of a concentrated but still practically implementable factor strategy. We did not use even wider breakpoints, other than for robustness checks to keep a reasonable number of firms in each portfolio formed to avoid a high impact from the idiosyncratic firm risk.

It was also within the scope to investigate whether these factors could reasonably be timed as an implementable strategy based on the relative spread of the high and low portfolios average scores, like the way done in (Asness et al., 2017). It was beyond our scope to draw wider implications about market timing in relation to market efficiency.

We did not account for trading cost since it is not realistic to assume all investors face the same trading costs. Excluding trading, costs can significantly bias the attractiveness of a strategy. All the strategies we analyzed used yearly rebalancing and the portfolio turnover is therefore somewhat limited. We also excluded microcap stocks to avoid holding the most illiquid stocks. We, therefore, consider the market-impact and other transaction costs to only have had a marginal impact on the results presented.

Many of the factor strategies have a short-side, and this makes the factor strategies impossible to implement for many investors, as it is common to be subject to various shorting constraints and shorting cost. Nevertheless, it is a practical approach to use factors for the discovery of potential anomalies and investors can still implement the long-side of the factor strategies. Therefore, we decided to analyze the strategies based on the assumption that investors were under no shorting constraints and faced no shorting costs but occasionally tested the long-side in isolation to see whether a strategy mainly was driven by the long- or short-side of the factor. It was beyond the scope to analyze the limits of investors shorting constraints.

3 Literature Review

The literature review is divided into three sections. In the first section, we review various explanations for the value premium. The second section contains a more in-depth review of the most relevant papers for the thesis that formed the bedrock for our methodology and hypotheses. In the third section, we formulate our hypotheses and explain the logic for each of them based on the reviewed literature.

3.1 General Literature Review

(Linnainmaa & Roberts, 2016) states that two fundamental questions permeate the empirical asset pricing literature, namely identifying a simple empirical asset pricing model that provides a good description of the cross-section of returns as well as a precise description between the boundaries of risk-based and behavioral explanations for the many anomalies. Several decades ago the CAPM model was built, and various academics found that the market beta could explain a significant portion of the cross-sectional expected returns (Lintner, 1965; Sharpe, 1964).

Over the years it has become clear that traditional asset pricing models such as the CAPM and Fama French three-factor, Fama French Carhart four-factor and Fama French five-factor models only insufficiently explained the cross-section of expected returns. According to (Cochrane, 2011): “We thought 100% of the cross-sectional variation in expected returns came from the CAPM, now we think that is about zero, and a zoo of new factors describes the cross-section”.

After the initial discovery of the market beta, many academics have been searching for patterns and firm fundamental characteristics that could reasonably provide explanatory power for the cross-section of expected returns, which has led to hundreds of published papers on this topic revealing factors outside the classical asset pricing models (Harvey, et al., 2016). (Harvey et al., 2016) documents 316 different anomalies that have been published in top academic journals. Another example of the multitude of factors discovered by academics is (Hou, Xue, & Zhang, 2017) recent study where they try to replicate 447 anomalies. Anomalies usually fall under one of the three overarching explanations:

(1) unmodelled risk
(2) mispricings
(3) data-snooping

In the following section, we first define and introduce the value factor followed by a discussion on the literature related to the three overarching explanations for the value premium.

3.2 The Value Anomaly

The value anomaly is one of the most researched anomalies in the empirical asset pricing literature and various researchers have found that value stocks have a strong tendency to outperform growth stocks, on average, while not being accounted for by (Sharpe, 1964)'s CAPM (Rosenberg, Reid, & Lanstein, 1985; Fama & French, 1992, 1993). In 1993 (Fama & French, 1993) published their new asset pricing model, the three-factor model that could explain much of the variation in the cross-section of stock returns based on differences in covariance with the market, firm size and book-to-market ratio. The three-factor model essentially consolidated the CAPM and many of the previously discovered fundamental anomalies discovered by (Banz, 1981; Bhandari, 1988; Rosenberg, Reid, & Lanstein, 1985; Sharpe, 1964; Stattman, 1980), that historically had predicted the cross-sectional returns. The HML factor in the Fama French three-factor model is based on the book-to-market ratio which is traditionally known as the value factor.

3.3 Possible Explanations for the Existence of the Value Premium

Researchers have called in question the validity of the value premium as a viable trading strategy (Loughran, 1997). (Loughran, 1997) found the HML value premium to be statistically insignificant among the largest quantile of firms and mostly driven by the short-side of smaller growth firms as well as a January seasonal effect for the period 1963 to 1997. However, (Fama & French, 2006) found the value premium to be statistically significant when they moved out-of-sample into the pre-1963 era, as well as for both large and small stocks in 14 major markets outside the US in the period 1975 to 2004. Additionally, they also found it to be statistically significant from 1963 to 2004, for all sizes of stocks, using earnings-to-price as an alternative proxy for the value premium. Similar findings have been published by (Linnainmaa & Roberts, 2016) which also found the value premium to be significant in the out of sample period from 1926 to 1963.

3.3.1.1 Unmodelled Risk

(Fama & French, 1993) consistent with the efficient market hypothesis, argues that the value premium is compensation investors require for undiversifiable risk exposure they bear by holding value firms, but have struggled to identify the source of risk. This overarching idea is known as the risk hypothesis and is congruent to the idea that the value premium must stem from the unmodelled risk that some rational investors are aware of and require compensation for being exposed to. Based on this fundamental assumption, the time-varying risk hypothesis has been proposed by (Fama & French, 1993), that argues the HML factor strategies are exceptionally volatile in bad times, where the premium required by investors for risk itself, is higher compared to periods with financial stability.

(Carlson, Fisher, & Giammarino, 2004) have proposed an operating leverage hypothesis for the value premium. They suggest that operating leverages magnifies firm exposure to economic risk. They, therefore, argue that firms that are more profitable in relation to their price, i.e., value firms, are just firms that have more operating leverage thus a higher risk exposure. (Favilukis & Lin, 2016) more specifically suggest the value premium stems from higher operational leverage caused by wage rigidity but find that their model cannot fully explain this effect since the higher market betas explain much of the volatility of these value firms with high wage rigidity. (Petkova & Zhang, 2005) suggests a similar explanation for the value premium, that it is driven by costly reversibility and countercyclical price of risk but struggles to explain firms with a low market beta that earn the value premium.

(Lettau & Wachter, 2007) have proposed a duration-based explanation of the value premium which suggests value firms are firms with short-duration assets that are riskier and demand higher compensation. (Novy-Marx, 2013) finds that both the risk, operating leverage, and the duration-based hypotheses are strictly in contrast to the average returns observed empirically when the value premium is analyzed jointly with his profitability measure, a phenomenon we will review later.

(Kogan & Papanikolaou, 2013) finds five-factors that tend to comove including an alternative value premium proxy (price-to-earnings) and propose the comovement of these can be explained by investment-specific technology (IST) shocks. Firms with high growth opportunities tend to benefit more from IST shocks as the productive capital they employ in the future will embody these recent technological changes. (Kogan & Papanikolaou, 2013) therefore suggest the value premium is a risk premium driven by exposure to IST shocks.

3.3.1.2 Mispricing

Several studies argue that the risk hypothesis must be incorrect (Chopra, Lakonishok, & Ritter, 1992; De Bondt & Thaler, 1987; Lakonishok, Shleifer, & Vishny, 1994; Novy-Marx, 2013). Many empirical results challenge the risk explanations, for instance (Lakonishok et al., 1994), who notes value betas tends to be higher than growth betas in good times and lower in bad times, giving value strategies very desirable characteristics that are hard to account for with risk-based models.

Several scholars believe a systematic and predictable behavioral expectation errors drive the value premium, rather than unmodelled risk alone (Asness et al., 2017; Piotroski & So, 2012)**.** (Piotroski & So, 2012) argues that the HML-factor and alternative value proxies capture systematically pessimistic and optimistic performance expectations for glamour[1] and value firms. The value premium may arise because investors have a strong tendency to overweight firms' historical fundamentals and thereby underweight new financial data that contradict the past performance which leads to the investor neglecting the tendency for these firms to mean-revert (Lakonishok et al., 1994). A value investor can then capture a premium when corrections arise from the reversal of these expectation errors (Piotroski & So, 2012).

(Piotroski & So, 2012) tried to identify firms where the expectations based on their value or glamour classification were incongruent with the actual strength of their fundamentals by sorting each firm into value or glamour portfolios based on their FSCORE. The FSCORE is a sum of nine binary signals related to the improvement or deterioration of a firms' fundamentals, first used by (Piotroski, 2000). (Piotroski & So, 2012) find that the average return generated by the subsection of value (glamour) firms which have a high (low) FSCORE is significantly higher (lower).

(Piotroski & So, 2012) further, supplement their analysis by looking at inter-temporal variation in investor sentiment and find that, periods of high investor sentiment produced market prices where performance expectations deviated further and more frequently from fundamentals. Interestingly, returns of the congruent value (glamour) strategy displayed no relation with the investor sentiment, while the incongruent value (glamour) portfolios returns increased (decreased). They conclude the value premium is likely to be driven by value/glamour firms incongruent with their fundamentals, thus by mispricings, while value/glamour firms congruent to their fundamental do not generate a value premium. They

[1]Here (Piotroski & So, 2012) use the term glamour rather than growth, because the term growth may resonate more with an unmodelled risk explanation, while the term glamour refers to the overly optimistic expectations towards firms with poor scores on value metrics such as the book-to-market ratios.

argue their findings is consistent with expectation biases stemming from many factors such as optimism, anchoring, representativeness and confirmation biases, and add to the doubt of risk-based explanations.

3.3.1.3 Data-Snooping

Due to the high and historically increasing discovery rate of new factors, academics in the field have increasingly become concerned about data-snooping (Harvey, Liu, & Zhu, 2016; Linnainmaa & Roberts, 2016; Yan & Zheng, 2017). In 2015 (Fama & French, 2015) proposed their new five-factor model, that empirically could explain away many of the newly discovered anomalies and thus provided a “shared story” (Karolyi, 2015). However, when (Linnainmaa & Roberts, 2016) in their study tested the profitability and investment anomalies proposed by (Fama & French, 2015, 2016) out-of-sample in the pre-1963 period, they found no statistically reliable premia from the two newly added factors.

It is difficult to gauge the magnitude of data-snooping as it is unknown to what extent the data has been scrutinized by academics previously. This is a problem as “the more scrutiny a collection of data is subjected to, the more likely it is that interesting (spurious) patterns emerge” (Lo & MacKinlay, 1990) and thus many variables will look like they generate abnormal returns by pure chance (Harvey et al., 2016; Lo & MacKinlay, 1990; Yan & Zheng, 2017).

A significant contributor to the data-snooping risks is an inherent publication bias (Harvey et al., 2016). It is difficult to get an insignificant result published, and an army of academics, therefore, test hundreds or thousands of different variables. It is highly likely they stumble upon significant results by chance and get it published (Harvey et al., 2016). An additional and more subtle publication bias arises because it is particularly challenging to publish a replication study in economics and finance, academics therefore often attempt to publish new factors rather than rigorously verifying previously discovered factors (Harvey et al., 2016).

(Linnainmaa & Roberts, 2016) in their study analyze 36 well-known accounting-based anomalies by testing how they fare out-of-sample in their pre and post-discovery periods. They find that most of these anomalies lose their power when they move out-of-sample in either direction while volatility and correlation with other anomalies increase. Only 16 of the 36 factors tested have significantly positive alphas when tested out-of-sample using the CAPM and the three-factor model. Their results, therefore, suggest that data-snooping have had a significant impact on the discovery of new factors. Due to these widely acknowledged data-snooping concerns, (Harvey et al., 2016) have called for a higher threshold for passing a factor as valid and suggest they need to be statistically different from zero with a minimum t-statistic of 3.

3.4 Alternative Value Metrics

As previously touched upon, several scholars have noted that the value premium can also be captured with alternative value metrics (Fama & French, 2006, 2011; Gray & Vogel, 2012; Lettau & Wachter, 2007). We define a value metric to be a ratio between market capitalization (MCAP) or enterprise value (EV) to an income statement, cash flow statement or balance sheet item.

Book-to-market have historically been a crude but effective way to sort value stocks and capture the value premium. (Fama & French, 1992)'s reasoning behind using book-to-price as a value proxy is that it is easily measurable and has historically provided a robust characterization of the cross-section of expected returns but also found that earnings-to-price in the period 1963 to 1990 was largely absorbed by the book-to-market and size factors and therefore decided to focus on book-to-market. (Fama & French, 2011) argues that:

"*We always emphasize that different price ratios are just different ways to scale a stock's price with a fundamental, to extract the information in the cross-section of stock prices about expected returns. One fundamental (book value, earnings, or cash flow) is pretty much as good as another for this job, and the average return spreads produced by different ratios are similar to and, in statistical terms, indistinguishable from one another. We like the book-to-market ratio because the book value in the numerator is more stable over time than earnings or cash flow, which is essential for keeping turnover down in a value portfolio. Nevertheless, there are problems in all accounting variables, and book value is no exception, so supplementing book-to-market ratio with other ratios can in principle improve the information about expected returns. We periodically test this proposition, so far without much success.*"

(Gray & Vogel, 2012) in their study find statistically significant differences in the performance of various alternative value metrics and therefore call in question (Fama & French, 2011)'s statement that one fundamental is pretty much as good as another.

What motivated our investigation of alternative value metrics is the lack of arguments for why market cap scaled by book equity should be a better proxy for value premia than the market cap or enterprise value scaled to some other income statement, balance sheet or cash flow item.

3.4.1 Alternative Value Metrics Included for Further Analysis

We chose to investigate the following value metrics as we found an economic rationale behind including each one of them:

1. Revenue-to-enterprise-value (REVEV)
2. Gross-profit-to-enterprise-value (GPEV)
3. EBITDA-to-enterprise-value (EBITDAEV)
4. EBIT-to-enterprise-value (EBITEV)
5. Cash-flow-from-operations-to-enterprise-value (CFOEV)

EBITEV is commonly used in the finance industry by practitioners, and it was therefore particularly interesting to investigate its return characteristics and compare them to the traditional HML factor. By using enterprise value rather than market cap, we normalize different debt structures to reduce the impact of leverage and taxes (Modigliani & Miller, 1958). Practitioners such as (Greenblatt, 2010) uses EBIT because it is "possible to put companies with different levels of debt and different tax rates on an equal footing when comparing earnings yields" (Greenblatt, 2010).

EBITDAEV is also interesting to compare to the HML factor, as it normalizes the depreciation and amortization. Multiple studies have tested the predictability of EBITDA scaled by various fundamentals (Gray & Vogel, 2012; Tjero-Pech, Welden, & House, 2008). (Tjero-Pech et al., 2008) find EBITDA scaled by assets has significant predictive power over the cross-section of stock returns for agricultural firms, it is therefore interesting to investigate if the same effect can be found for the general market if EBITDA is scaled by enterprise value. The EBITDA-to-Enterprise-Value metric has also been tested by (Gray & Vogel, 2012)**,** that found the portfolio with the most attractive quintile EBITDA-to-enterprise-value ratios generated a significant three-factor alpha.

GPEV was also chosen as an alternative value metric as (Novy-Marx, 2013) findings show gross profits scaled by assets have predictive power. Instead, we scaled by enterprise value as we wanted to account for the price an investor pays for the stock, which according to our definition makes it a value metric. (Gray & Vogel, 2012) in their study tested the same factor we used, gross profits scaled to enterprise value, but with a slightly different construction methodology and found significant alpha unexplained by the three-factor model.

We also decided to include the revenue-based value metric, REVEV, to investigate if firms with higher revenue did better when controlling for the stock price, although this factor may be severely polluted by large differences in business models that dictate the profit margins. It was also a way to rule out that significant GPEV alpha returns were primarily driven by differences in revenue. The lack of papers investigating revenue-based value metrics made the comparison interesting.

Several studies have shown that firms' cash flow has significant predictive power over future stock returns (Lakonishok et al., 1994; Sloan, 1996), we, therefore, include cash-flow-from-operations-to-enterprise-value (CFOEV). (Lakonishok et al., 1994) find that the cash-flow-to-price metric has significant predictive power over future-stock returns. (Sloan, 1996) find that the accrual component of a company's earnings has significant negative predictive power over future stock returns, again indicating that a metrics based on cash flow might be able to better capture the value premium compared to a metric based solely on an income statement item. The similar metric cash-flow-to-enterprise-value was also included as an alternative value factor in (Gray & Vogel, 2012). They found the two highest quintiles of portfolios formed on the metric generated significant three-factor alpha. We decided to focus on the operating part of the cash flow variable as we wanted to normalize the impact of a firm's capital structure, and thereby focused on the operating part of the firm's activities.

3.5 Key Papers

The following section introduces the most central papers that inspired the hypotheses and the methodology used throughout the paper. This includes the introduction of the asset pricing factor models that will be used in the analyses throughout the paper.

3.5.1 Factor Models

3.5.1.1 Fama French Three-factor Model

The key literature often uses the three factors from the three-factor model proposed by (Fama & French, 1993) to test if anomalies are explained by existing factors. The three-factor model predicts that the cross-sectional excess returns can be described as:

$$r - r_f = \alpha + \beta_1(r_m - r_f) + \beta_2(SMB) + \beta_3(HML) + \epsilon, \tag{1}$$

$r - r_f$ denotes the expected excess return on a stock or a portfolio on a specific cross-section and is dependent on r_f which denotes the risk-free rate, α is the intercept, β_1, β_2, and β_3 are the sensitivity coefficients to the three-factors respectively, and the market factor is given by the market excess return $(r_m - r_f)$. The *HML* (high minus low) is based on going long the high book-to-market, while shorting the low book-to-market firms and similarly for *SMB*, short the big firms and buys the small firms. See the methodology section for details on the *HML* and *SMB* factor constructions.

With these factors, it is possible to test whether portfolio returns of a given strategy can be explained by exposure to the three factors by simultaneously OLS regressing the strategy returns on the returns of the three factors in the three-factor model. If the three factors explain a strategy, it will show up as beta coefficients, i.e., loadings on the factors rather than as a positive intercept, i.e., positive alpha. Portfolio returns that have a high loading on the *HML* factor are what is traditionally associated with the value premium in the academic literature.

3.5.1.2 Fama French Five-factor Model

After the formulation of the Fama French three-factor model, academics discovered many new anomalies that the three-factor model was unable to explain (Fama & French, 2015). (Fama & French, 2015) Find that the addition of the two new factors—the investment and profitability factor—can explain away many of these anomalies such as the factors discovered by (Aharoni, Grundy, & Zeng, 2013; Novy-Marx, 2013). The Fama French five-factor model is given by,

$$r - r_f = \alpha + \beta_1(r_m - r_f) + \beta_2(SMB) + \beta_3(HML) + \beta_4(RMW) + \beta_5(CMA) + \epsilon. \quad (2)$$

As the five-factor model is an extension of the three-factor model, the new elements in the formula are: the profitability factor, *RMW*, and investment factor, *CMA*, *and* their respective exposure coefficients β_4 and β_5.

While the three original factors are included because of their ability to explain away much of the variation in the cross-section of expected returns, (Fama & French, 1992) rationalize the addition of the two new factors using (Modigliani & Miller, 1963)'s decomposition of the dividend discount model divided by book equity,

$$\frac{M_t}{B_t} = \frac{\sum_{\tau=1}^{\infty} E(Y_{t+\tau} - dB_{t+\tau})/(1+r)^{\tau}}{B_t}. \quad (3)$$

M_t is the market equity at time *t*, $Y_{t+\tau}$ is the total equity earnings for period $t + \tau$, B_t is the book equity at time *t*, $dB_{t+\tau} = B_{t+\tau} - B_{t+\tau-1}$ is the change in book equity, or according to the clean surplus accounting model, retained earnings, and *r* is the approximate long-term average expected return or the internal rate of return on expected dividends. They then argue that if M_t and B_t is fixed, a higher increase in book equity ($dB_{t+\tau}$ i.e., investments) implies a lower expected return. This is the logic behind the addition of the *CMA* (investment) factor addition.

The rationale for the inclusion of the profitability factor *RMW* follows a similar idea. Suppose M_t and B_t is fixed in eq. (3). Then, higher earnings imply a lower required return. While this may seem intuitive, a major flaw in the model is that of the profitability and investment factors, are that they build on the assumption that they are good proxies for future profitability and investment (Hou, Mo, Xue, & Zhang, 2018), which has to be taken into account. See the methodology section for details on the *CMA* and *RMW* factor construction.

(Hou et al., 2018) argue the q-factor model is superior to the Fama French five-factor model. We still decided to include the classical five-factor model in our analysis to make our constructed factors comparable to other's strategies tested against the five-factor model.

3.5.1.3 Q-Factor model

In 2015, (Hou, Xue, & Zhang, 2015) published a new factor model called the q-factor model that built on investment-based asset pricing anomalies and captured many of the anomalies that proved challenging for the Fama French five-factor model,

$$r - r_f = \alpha + \beta_1(MKT) + \beta_2(ME) + \beta_3(I/A) + \beta_4(ROE) + \epsilon. \tag{4}$$

In the q-factor model $r - r_f$ is the return in excess of the risk-free rate and $\beta_1, \beta_2, \beta_3$ and β_4 are the exposure coefficients to the factor premia *MKT*, *ME*, *I/A* and *ROE*.

The investment and profitability factors in the q-factor model are also based on the principles of the discount dividend model as in the Fama French five-factor model (Hou et al., 2015). (Hou et al., 2015) argues that the q-factor model largely subsumes the five-factor model even after dropping the HML factor. They argue that the value and investment factors in the Fama French five-factor model are closely related because firms with high book-to-market should invest less and hence earn a higher expected return and show that the HML and CMA factor have a correlation of 0.69 (Hou et al., 2015). See the methodology section for details on the factor construction.

3.5.1.4 The Six-factor Model

(Frazzini, Kabiller, & Pedersen, 2013) use a six-factor model to quantify and explain the incredible historical performance and factor exposure of Warren Buffet:

$$r - r_f = \alpha + \beta_1(MKT) + \beta_2(SMB) + \beta_3(HML) + \beta_4(UMD) + \beta_5(BAB) + \beta_6(QMJ) + \epsilon, \tag{5}$$

again, $r - r_f$ is the return in excess of the risk-free rate. The factors in their model are the three-factors from the original three-factor model *MKT* or $(r_m - r_f)$, *SMB*, and *HML* as well as three additional factors *UMD*, *BAB*, and *QMJ*, while $beta_1, beta_2, beta_3, beta_4, beta_5$ and $beta_6$ are the factor exposure coefficients and ϵ is the error term.

UMD is the (up minus down) momentum factor. (Jegadeesh & Titman, 1993) initially discovered the momentum anomaly that since has been further documented by (Asness, 1994; Carhart, 1997). (Asness, Moskowitz, & Pedersen, 2013) show the persistence of momentum returns in markets around the world. A common behavioral explanation for the anomaly in the literature is the tendency of investors to initially underreact to news with a delayed overreaction driving this anomaly.

The intuition behind the betting against beta (*BAB*) factor is that many investors prefer holding high-beta assets as they can achieve a higher return without using leverage, but an unconstrained investor has historically been able to achieve a higher risk-adjusted return by leveraging stocks with low betas.

(Novy-marx, 2018) criticizes the untraditional factor construction used by (Frazzini & Pedersen, 2014), as several non-standard factor construction procedures are used. First, they deviate from the standard methodology by ranking stocks as described above, and secondly, they leverage the portfolio to achieve a beta of 1 rather than buying the market in proportion to the short-side tilt. (Novy-marx, 2018) claims these two procedures in combination essentially functions as a backdoor to equal-weighting the portfolio giving a massive overweight to the firms in the bottom 1% of total market cap. After one account for the transaction costs, the impact of nano and microcap turnover reduces *BAB*'s profitability by almost 60% giving it an insignificant alpha on the five-factor model (Novy-marx, 2018). Despite this reasonable critique, we still decided to test the strategies that pass the five-factor model on the six-factor model to see how they load on these alternative strategies.

The last additional factor is the *QMJ* factor that buys stocks that are high quality and sells stocks that are of lower quality. *QMJ* is a compound of various factors investors should want to pay more for everything else equal and tend to capture companies that are profitable, growing, safe and have a high payout to investors. See the methodology section for detail on the *UMD*, *BAB* and *QMJ* factor construction.

3.5.2 "The other side of value: The gross profitability premium"

Despite (Linnainmaa & Roberts, 2016)'s finding that the RMW profitability premium loses its explanatory power when moving out-of-sample, (Novy-Marx, 2013) argue that substituting earnings by gross profitability and book value of equity with total assets is a better measure for capturing the profitability premium. (Novy-Marx, 2013) states: "value strategies finance the purchase of inexpensive assets through the sale of expensive assets, and profitability strategies achieve the same end by financing the purchase of productive assets through the sale of unproductive assets" and that both strategies yield abnormal returns.

Surprisingly, his profitability factor generates abnormal returns despite the disconnect from a market value component. Because of this disconnect from market price, it makes sense that his profitability measure works exceptionally well together with book-to-market, which does incorporate the price component. Intuitively, HML value portfolios may be polluted with unprofitable stocks, and growth portfolios may be polluted with profitable stocks, thus adding a profitability measure to HML helps avoid buying a stock that is more unprofitable than cheap and selling stocks that are more profitable than expensive (Novy-Marx, 2013). Alternative value metrics may also achieve this goal, which will be considered in the hypotheses section.

(Novy-Marx, 2013) considers gross profitability to be the cleanest accounting measure of true economic profitability, and he also considers it to be a better proxy for future profitability compared to earnings because profitability measures further down the income statement are more polluted.

Profitable firms are quite different from value firms since they resemble classical growth firms as they have similar covariances and generally low book-to-market values (Novy-Marx, 2013). According to (Novy-Marx, 2013) firms with high gross profitability are growth firms, but good growth firms. He believes value investors should pay attention to his gross profitability measure as this increases the returns of a value portfolio based on the traditional book-to-market measure substantially.

3.5.3 "Deep value."

(Asness et al., 2017) investigate if the value premium is time-dependent by testing if periods where the spread between expensive and cheap stocks have predictable power in relation to subsequent returns. They define these periods where the spread is wider than its 80th percentile as "deep value episodes."

They investigate deep value episodes across individual equities, equity index futures, currencies, and global bonds to provide new information about the predictability provided by the value spread. In total, they test 522 value strategies across markets and find about 3000 deep value episodes.

They employ a strategy where they buy value stocks and short growth stocks when the value spread exceeds the 80th percentile measured on the history up until that time and exit the market when the value spread reaches its historical median value. They find that returns to the intra-deep-value strategies, such as strategies based on value spreads within industries, are time-dependent as their returns are particularly high after deep value episodes. They, however, find that deep value strategies are often less significant within specific equity markets and only marginally significant within different types of index futures and currencies.

They also find that the value premium does not appear to be compensation for market risk but may be a compensation for other risks, as market betas for value portfolios are close to zero and even negative during deep value episodes. (Asness et al., 2017) also find the value spread to be at least partly rational, as the earnings of value firms (the long-side of the factor) deteriorate faster than the earnings of growth firms (the short-side of the factor) during these periods and continue to deteriorate for about two years after a deep value episode.

They use a proxy for the tone of news stories and find that the sentiment towards value firms is generally more negative compared to growth firms and that this effect is amplified during deep value episodes. This sentiment trend tends to mean revert in the following year after a deep value event. They also find analyst earnings forecast, tends to be revised negatively for value companies relative to growth companies, before and after a deep value event.

They furthermore find that value stocks face stronger selling pressure right before deep value events and investigate if the effect can be attributed to past returns or past fundamentals and find that the selling pressure is primarily driven by negative returns of the previous months consistent with theories of over-extrapolation of previous returns.

Limits to arbitrage also increase around deep value episodes as both the bid-ask spread and the cost of short selling growth stocks tend to rise (Asness et al., 2017). Furthermore, they document the volatility of the long-short value portfolios tends to rise in deep value periods. They argue that these effects are a result of increased value arbitrage as the demand for shorting growth stocks increase in deep value periods.

3.5.4 "Decomposing value"

A recent study by (Gerakos & Linnainmaa, 2018) suggests there is a significant disconnect between book-to-market ratios and the value premium. (Gerakos & Linnainmaa, 2018) use the notion that a firm is a value firm if:

(1) "It's market-value of equity decreased while its book value of equity remained unchanged;"
(2) "It's book-value of equity increased while its market value stayed the same; or.
(3) "it was already a value stock, and nothing changed."

In their study, they, therefore, decompose book-to-market into a component affected by changes in firm size and a component orthogonal to changes in firm size to find out if one of the three categories drives the value premium. (Gerakos & Linnainmaa, 2018) find that when they control for five years of past annual changes in firm size, book-to-market loses its statistical significance in relation to return prediction. Their decomposition has revealed a significant tendency of firms that have shrunk in size to outperformed those that have grown. Based on their empirical results they conclude "the value premium is therefore about changes in firm size" rather than driven by book-to-market by itself, i.e., firms in category (1), and this means some firms that would be categorized as growth firms based on book-to-market alone also earns the value premium according to their definition (Gerakos & Linnainmaa, 2018). Interestingly, they also find that both size components of book value and the orthogonal component tend to comove. Therefore, the orthogonal component can be used to hedge the value premium (Gerakos & Linnainmaa, 2018).

Additionally, they find that two other well-known anomalies span the value premium: long-term reversals originally documented by (De Bondt & Thaler, 1987) and the net share issues. Intuitively this makes sense, as changes in firm size equal stock returns minus dividends plus net share issues. Consequently, net share issues are part of changes in firm size (Gerakos & Linnainmaa, 2018). According to (Gerakos & Linnainmaa, 2018) tests, the size component of the value premium is not accounted for by current models of risk and therefore could be mispricings, but they do not exclude the possibility of a risk-based explanation.

3.6 Hypotheses

Based on our literature review, we found that there were reasons to investigate alternative value metrics based on enterprise value to some fundamental, as these metrics may be equivalently powerful or better proxies for capturing the value premium.

It is expected most of the return variation based on these alternative metrics can be captured by the HML and RMW factors from the five-factor model. This is because many of the alternative valuation metrics we analyze are closely related to earnings-to-price which can be decomposed into book-to-market and earnings-to-book, i.e., the HML and RMW factors:

$$\frac{B}{MCAP} \cdot \frac{E}{B} = \frac{E}{MCAP}, \qquad (6)$$

B is the book value of equity, *MCAP* is the market cap, and *E* is earnings. We suggest a new terminology for alternative value metrics that can be decomposed into a book-to-market (value) and an earnings-to-book (profitability) like ratio as shown above. Any such metrics that is a ratio between an income-related income statement item or cash flow and market cap or enterprise value can logically be classified as a profitable value metric, as they contain elements that favor these two dimensions.

There are three reasons why the profitable value metrics we analyze may deviate and do better or worse than the HML and RMW.

First, the metrics as factors vary because when decomposed, each component is treated as having its own linear relationship to returns. They may, however, yield more explanatory power when integrated, i.e., the decomposition of *E/MCAP* into *B/MCAP* and *E/B* may yield no incremental information beyond what is already contained in *E/MCAP* or closely related value metrics.

Secondly, we substitute earnings with revenue, gross profits, EBITDA, EBIT, and cash-flow-from-operations. (Campbell & Shiller, 1998) states earnings is a noisy measure as well as (Novy-Marx, 2013) who advocates, the farther one moves down the income statement, the more polluted a measure of profitability may become. Items higher on the income statement are relatively clean accounting measures where the firms have less flexibility to move items back and forth in time. For instance, a value measure based on earnings could miss a firm that has both lower production costs, higher revenue and is spending a substantial amount on investments that would eventually lead to higher profitability (Novy-Marx, 2013). Earnings, therefore, may be a worse predictor of future stock returns compared to gross-profits-to-assets (Novy-Marx, 2013). Thirdly, we choose to use enterprise value instead of market cap.

There are multiple plausible explanations for why these alternative metrics may better capture the value premium. The alternative factors may be exposed to some undiversifiable common risk factors for which investors demand a premium. A premium of these alternative value metrics may also be driven by incongruence between the price and firm fundamentals due to overreactions to negative firm fundamentals, extrapolation of a negative trend or some other investor bias that may or may not be missed by the five-factor model due to the slight variation between the metrics. Note that these explanations are not mutually exclusive.

Based on the ideas from above we derived the following hypotheses:

Hypothesis 1:

Revenue-to-enterprise value (REVEV), Gross-profit-to-enterprise value (GPEV), EBITDA-to-enterprise value (EBITDAEV), EBIT-to- enterprise value (EBITEV), Cash-flow-from-operations-to-enterprise-value (CFOEV) capture alpha relative to existing asset pricing models.

(Huang, Zhang, Zhou, & Zhu, 2019) argue that a single fundamental variable may only sometimes predict returns and may not predict returns in the same way continuously through time due to structural breaks or changes in accounting rules. They, therefore, argue a collection of variables may better capture a firm's economic outlook and therefore increase the predictive power of returns. Because of these potential structural breaks and because value metrics may capture slightly different aspects of firms' fundamentals, we hypothesized that a combination of the metrics might be better able to capture value premia:

Hypothesis 2:
A combination of value metrics, i.e., a compound value metric better predicts a firm's expected return and therefore is able to subsume all its constituent variables with an improved Sharpe ratio and a significant alpha in relation to existing asset pricing models.

(Gerakos & Linnainmaa, 2018) results suggest the book-to-market value premium may be entirely driven by historical changes in size and a trend reversal. (Piotroski & So, 2012) and (Novy-Marx, 2013) results suggest the entire premium could be driven by a subsection of book-to-market sorted firms of relatively good quality measured on assets/gross profits or a high FSCORE.

Most of our value metrics have a natural gravitation towards higher quality firms as shown by the decomposition eq. (1) and our profitable value metrics can individually be decomposed into a profitability measure similar or equivalent to (Novy-Marx, 2013)'s profitability measure and the book-to-market ratio. (Novy-Marx, 2013) show his profitability factor in and of itself generates abnormal returns. We, therefore, tested whether his profitability metric can add anything to our profitable value metrics beyond its already natural gravitation towards more profitable firms.

Hypothesis 3:
(Novy-Marx, 2013) gross-profits-to-assets (GP/A) generally improves the performance of profitable value metrics measured on alpha, Sharpe ratio, and maximum drawdown.

After testing the above hypotheses, we further investigated whether the value premium could be attributed to a subsection of firms that have had negative annual changes in size over the past six years for profitable value metrics rather than just book-to-market as done by (Gerakos & Linnainmaa, 2018).

Hypothesis 4:
The factor premia of value metrics become insignificant when controlling for the annual change in size for the past six years.

Finally, we combined the ideas from (Novy-Marx, 2013) and (Gerakos & Linnainmaa, 2018) and investigated whether value premia are driven by a subsection of firms that are both experiencing trend reversal and have strong fundamentals. Their studies show the HML value premium becomes insignificant when they control for one of the two individually. Therefore, we hypothesized:

Hypothesis 5:
When controlling for both past annual change in size for the past six years and gross-profits-to-assets, value premia become significantly negative.

(Asness et al., 2017) suggest that the value spread based on book-to-price does predict subsequent returns to value strategies. So, we hypothesized that a value strategy could be timed only using historical data with a rolling window that calculates the historical value spread for book-to-market and profitable value factors. While (Asness et al., 2017) uses historical data on the value spread up until that time, we decided to use a historical rolling estimation window to account for potential structural changes in the underlying business environment, which led us to the hypothesis:

Hypothesis 6:
Value strategies generate significantly higher Sharpe ratios when timed based on the value spread using a rolling 10-year prior estimation window.

4 Methodology

4.1 Factor Construction Methodology

4.1.1 The Three-Factor Model Factors

The HML (high minus low) value-weighted portfolio returns are calculated by forming 2 x 3 double-sorted portfolios based on size and book-to-market. The HML factor's returns are calculated by equal weighting each of the four value-weighted portfolios where we long (short) the stocks with a book-to-market ratio above (below) the 70th (30th) book-to-market NYSE breakpoint in both the big and small size categories, that is firms that fall over and under the median of the NYSE market cap. Book-to-market refers to the book value of equity divided by the firms market cap.

Similarly, the small minus big factor (SMB) return in a given cross-section, is calculated by equal weighting the six 2 x 3 double-sorted portfolios, but here one goes long (short) the three portfolios sorted with book-to-market that also fall into the small (big) portfolio. The portfolios used for both HML and SMB are quarterly rebalanced. All firms listed on the NYSE, AMEX, and NASDAQ exchanges are included in the HML and SMB portfolios.

4.1.2 The Five-Factor Model Factors

The CMA and RMW factors are constructed the same way as HML in the three-factor model using the (Fama & French, 1993) standard methodology, except annual percentage change in total assets and operating profitability to book equity is used instead of book-to-market for the factors respectively, with the higher change in total assets and less profitable portfolios being shorted.

The SMB factor in the five-factor model is constructed differently than in the three-factor model. In the three-factor model it is constructed controlling for HML as described above, while in the five-factor model the SMB is constructed by taking the equal-weighted average of the three portfolios constructed in the same way as the original three-factor SMB controlled for HML, but additionally also equivalent portfolios that are instead controlled for CMA and RMW individually. Size is constructed with controls for HML, CMA, and RMW. HML, CMA and RMW are only constructed with control for size, but not each other mutually. That means the returns of these factor portfolios are to some degree a mix of premia from one another (Fama & French, 2015).

4.1.3 The Q-Factor Model Factors

The factor construction in the q-factor model differs slightly from the Fama French five-factor model. The MKT is the market excess return in relation to the risk-free rate as also included in the Fama French model. The I/A (profitability) factor is the annual change in total assets (Compustat annual item AT) divided by 1 year-lagged total assets. The ROE (profitability) factor is income before extraordinary items (Compustat quarterly item IBQ) divided by 1-quarter-lagged book equity. The I/A and ROE factors are constructed with the same NYSE breakpoints as the Fama French five-factor model investment and profitability factors but are based on 2 x 3 x 3 triple sorted portfolios on size, I/A and ROE instead, so they control for size and each other mutually. The ME (size factor) returns each month are given by the simple average return difference between the nine small and nine big portfolios, while I/A (ROE) is the simple average return difference between the six low I/A (high ROE) and six high I/A (low ROE) portfolios. Monthly value-weighted returns are calculated for all portfolios where ROE portfolios are monthly rebalanced at the beginning of each month, while the size and I/A portfolios are rebalanced annually.

4.1.4 The Six-Factor Model Factors

The UMD factor is constructed using a 2 x 3 double sort on size and momentum including stocks on NYSE, AMEX, and NASDAQ. The size portfolio sort is equivalent to SMB, while the momentum sort uses the 30th and 70th NYSE percentile breakpoints based on highest (up) and lowest (down) returns in the past year (excluding last month). The strategy then goes long the up-big and up-small portfolios and short the down-big and down-small portfolios. The positions in the four portfolios are equally weighted and rebalanced monthly.

The BAB factor is constructed by ranking all stocks available in CRSP by their beta at the beginning of each period. The 50% of stocks with a beta above (below) the median are assigned to a high (low) beta portfolio. Within their respective portfolios stock are weighted by their ranked betas, such that lower beta stocks have larger weights in the low-beta portfolio and higher beta stocks have larger weights in the high-beta portfolio. The strategy then goes long the low-beta portfolio and short the high-beta portfolio, and the position in each portfolio is scaled to have a beta of one to make the strategy market neutral. The factor might therefore at a given time hold \$1.4 of low-beta stocks while shorting \$0.7 of high-beta stocks, with an offsetting risk-free position to make the strategy self-financing (Frazzini & Pedersen, 2014).

The QMJ (quality minus junk factor) used in (Asness et al., 2017) is constructed using the same methodology described in (Asness, Frazzini, & Pedersen, 2018), and is defined as stock characteristics that should command a higher stock price. QMJ is constructed by giving firms a score based on their average rank across three variables, namely profitability, growth, and safety. The profitability score is derived from an average score across gross profits, profit margin, earnings, accruals, and cash flows. The growth score is given by the five-year growth in each of the profitability metrics used in the profitability score, and the score is then averaged across those to get the profitability score. The safety score is given by the average score of low volatility of profitability, low leverage and low credit risk, and then averaged across those three.

The QMJ factor is then constructed in the standard method as the HML, CMA, and RMW, by forming six portfolios with the quality score and size. QMJ then equal weight and long (short) the firms with the 30% best (worst) overall quality score in among the 50% largest and smallest firms with monthly rebalancing.

4.1.5 Obtaining Commonly Used Factor Model Data

The calculations of the three factors from the three-factor model in the five-factor model varies slightly from the methodology in the original three-factor model (Fama & French, 1993), as we use the most recent methodology from Ken French's website (French, 2019). The q-factor model returns are calculated as described in the methodology section and equivalent to the calculation in (Hou et al., 2015). The six-factor model is obtained partly by downloading the four factors HML, SMB, MKTRF, and UMD through Ken French's website (French, 2019). The two factors QMJ and BAB is obtained through the AQR database (AQR, 2019). The returns for the UMD and BAB factor in the extended Fama French five-factor model used by (Asness et al., 2017), were also downloaded from (AQR, 2019).

4.2 Methodology for Testing Factor Loadings and Alpha

We tested our factor returns against existing asset pricing models to investigate if the new factor returns were fully explained by known anomalies or if the new factor could generate a significant alpha (intercept) when regressed against the asset pricing factor models. A significant alpha would indicate that the new factors are better at capturing the value-premium compared to the factors included in the asset pricing models. We tested the factors against the Fama French five-factor model, the Q-factor model, and the Six-factor model. We initially started by whether the new factors could generate significant alpha when regressed against the five-factor model. If the factors could generate a significant alpha when regressed against the five-factor model we then regressed the factor returns against the q-factor model and finally the six-factor model if it survived the q-factor model.

4.3 Fama-MacBeth Regressions

The Fama-MacBeth regression as developed by (Fama & Macbeth, 1973) can be used to estimate factor premia with the explanatory power of the expected returns for a given asset. These factors can be anything from macroeconomic factors such as GDP growth and the unemployment rate to firm-specific characteristics such as past annual changes in market equity or gross profits margin. It is also possible to test multiple factors jointly to see if a factor that predicts returns in isolation is subsumed by other factors. An added benefit of the Fama-Macbeth regressions is that it corrects for the cross-sectional correlation of the error terms.

Before we ran the Fama-MacBeth regressions, we winsorized the explanatory variables that were not return based at the 0.5% and 99.5% level, and this modification was essential in this analysis as we otherwise would end up with some very extreme enterprise value-based value metrics. Value metrics with an enterprise value in the denominator are especially prone to extreme outliers as cash is subtracted in the enterprise price value calculation, which in rare cases cause the denominator to be extremely low compared to the numerator which can massively inflate the ratio. These extreme outliers can heavily skew the characteristic of an entire portfolio and have therefore been removed.

It is common for firms to have a negative EBIT, EBITDA, gross profits, cash flows and book-to-price ratios, which is problematic as it is not possible to take the logarithmic values of these variables. We, therefore, remove these firms which have resulted in one-fifth of the observations were dropped for some variables, and this is necessary to adjust for the ratios being skewed. This reduction of observations may have had implications for our estimated factor premia, especially if firms with negative operating earnings behave differently. Our Fama-MacBeth results can therefore only be used to analyze firms with a positive value metric (above zero).

For all our observations we required at least six years of price data as we looked at size changes over the past six years, we also required no missing explanatory variables. Excluding firms with less than six years of data inevitably excludes relatively young and therefore often smaller firms, for which we then had to be cautious about drawing implications about based on our results.

4.3.1.1 Fama-MacBeth Regressions Step 1

If we were to run Fama-MacBeth regressions on an individual firm level basis, we would end up running a time series regression for each firm i with m number of factor exposures, and after that run a cross-sectional regression for T monthly observations of m factors based on monthly returns for each asset i. While we could skip step one and run both time-series and cross-sectional regressions on an individual firm level, we chose to make 20 equal weighted and monthly rebalanced portfolios for which we aggregated firm-specific characteristics across each portfolio by simple averaging to simplify the analysis. This methodology is very similar to the one used by (Fama & Macbeth, 1973) where they allocate firms based on their beta ventiles to get less noisy beta estimates. While we did not allocate firms based on their betas, we chose to follow a very similar methodology where we allocated firms in our universe into the 20 portfolios based on their value metric ventile instead. This portfolio allocation meant, when we analyzed whether changes in size drove the EBITEV premium, we first allocated firms into ventiles based on their EBITEV. Just like for our factor construction we used annual fundamental data and formed the portfolio at the end of June each year.

Since the portfolios were sorted on a value metric, we could estimate the factor premium for that metric well in isolation. The risk arises, when one starts to control for other variables, because when characteristics are averaged for hundreds of firms across portfolios, as portfolio characteristics may start to resemble one another highly in some aspects. This may have reduced the accuracy of our factor premia estimates.

For all our Fama-MacBeth regressions we truncated observations before 31. December 1968 since there were too few firms that met our criteria before that period to divide firms into ventiles.

4.3.1.2 Fama-MacBeth Regressions Step 2

In the second step, factor exposures ($\hat{\beta}_{i,F_{m_{R_i}}}$) were obtained by running a time-series regression for *n* portfolios by regressing returns against the *m* factors, formally:

$$\begin{gathered} R_{1,t} = \alpha_1 + \beta_{1,F_{1_{R_1}}} F_{1_{R_1},t} + \beta_{1,F_{2_{R_1}}} F_{2_{R_1},t} + \dots + \beta_{1,F_{m_{R_1}}} F_{m_{R_1},t} + \epsilon_{1,t} \\ R_{2,t} = \alpha_2 + \beta_{2,F_{1_{R_2}}} F_{1_{R_2},t} + \beta_{2,F_{2_{R_2}}} F_{2_{R_2},t} + \dots + \beta_{2,F_{m_{R_2}}} F_{m_{R_2},t} + \epsilon_{2,t} \\ \vdots \\ R_{n,t} = \alpha_n + \beta_{n,F_{1_{R_n}}} F_{1_{R_n},t} + \beta_{n,F_{2_{R_n}}} F_{2_{R_n},t} + \dots + \beta_{n,F_{m_{R_n}}} F_{m_{R_n},t} + \epsilon_{n,t}, \end{gathered} \tag{7}$$

Where $R_{i,t}$ is the expected return for portfolio *i* (*n* total) at time *t*, while $F_{j_{R_n,t}}$ is the j^{th} factor value (*m* total) for portfolio *i* at time *t* and ϵ the error terms. β_{i,F_m} are the factor exposures that we obtain in step one. Note the factors *F* for which we obtain exposures can be time-dependent, but equal across portfolios or assets for common factors such as macroeconomic variables, but vary for portfolio specific values such as the book-to-market ratio, hence we here use the subscript R_n, to note their asset specific value.

4.3.1.3 Fama-MacBeth Regressions Step 3

In the third step, we then for each portfolio *i* ran *T* cross-sectional regressions where we regressed portfolio returns for each on the m_{R_i} factor exposures ($\hat{\beta}_{i,F_{m_{R_i}}}$), that we obtained in step one:

$$\begin{gathered} R_{i,1} = \gamma_{1,0} + \gamma_{1,1}\hat{\beta}_{i,F_{1_{R_i}}} + \gamma_{1,2}\hat{\beta}_{i,F_{2_{R_i}}} + \dots + \gamma_{1,m}\hat{\beta}_{i,F_{m_{R_i}}} + \epsilon_{i,1} \\ R_{i,2} = \gamma_{2,0} + \gamma_{2,1}\hat{\beta}_{i,F_{1_{R_i}}} + \gamma_{2,2}\hat{\beta}_{i,F_{2_{R_i}}} + \dots + \gamma_{2,m}\hat{\beta}_{i,F_{m_{R_i}}} + \epsilon_{i,2} \\ \vdots \\ R_{i,T} = \gamma_{T,0} + \gamma_{n,1}\hat{\beta}_{i,F_{1_{R_i}}} + \gamma_{n,2}\hat{\beta}_{i,F_{2_{R_i}}} + \dots + \gamma_{n,m}\hat{\beta}_{i,F_{m_{R_i}}} + \epsilon_{i,T}, \end{gathered} \tag{8}$$

$R_{i,t}$, as before, is the expected return for portfolio *i* (*n* total) and $\epsilon_{i,T}$ the error term. We thereby obtain an intercept, and factor risk premia for each portfolio denoted $\gamma_{n,m}$. We assume $\epsilon_{i,T}$ are i.i.d, and are therefore able to obtain the final factor premia γ_m for factor F_m of length *T*, by merely averaging the m^{th} γ over *T*. For the t-statistic we obtain γ_m by averaging the m^{th} γ over *T* and its standard deviation is also calculated over *T* periods. The t-statistic is then:

$$\frac{\gamma_m}{\sigma_{\gamma m}/\sqrt{T}}. \tag{9}$$

4.4 Portfolio Evaluation Metrics

4.4.1 Sharpe Ratio

$$SR = \frac{E(r - r_f)}{\sigma(r - r_f)}. \tag{10}$$

The Sharpe ratio is a measure of the risk-adjusted return. It measures the reward of an investment strategy per unit of risk. The numerator $E(r - r_f)$ is the expected arithmetic risk-free rate subtracted from the historical arithmetic expected return, where r_f is obtained from Ken French’s homepage (French, 2019), and this is measured against the standard deviation of the historical monthly arithmetic returns in excess of the risk-free rate, $\sigma(r - r_f)$. Investors generally prefer higher Sharpe ratios, as they prefer high returns while at the same time having a low-risk exposure. The investor preferences are more complex than what the Sharpe ratio can capture, as investors might avoid investment strategies that have skewed returns and are subject to crash risks (Pedersen, 2015).

4.4.2 Cumulative Value of the Factor

The following formula defines the cumulative value,

$$Cumulative\ value_{t=n} = (1 + retx_{t=1}) \cdot (1 + retx_{t=2}) \cdot \cdots \cdot (1 + retx_{t=n}). \qquad (11)$$

The formula indicates how much money an investor would have at time $t=n$ if the investor invested in the strategy from the initial portfolio formation at time $t=1$. The current cumulative value of the factor is the cumulative value of a factor investment strategy on December 31, 2018.

4.4.3 High Water Mark

A strategy's high water mark is the highest cumulative value that the strategy has had in a given period. We use the formula for the high water mark as given in (Pedersen, 2015):

$$High\ water\ mark = Maximum\ cumulative\ value. \qquad (12)$$

4.4.4 Drawdown, Current Drawdown, and Maximum Drawdown

The drawdown is a way for investors to evaluate the riskiness of an investment strategy. The drawdown is the combined losses a strategy has had since it achieved its high water mark divided by the strategy's high water mark. The formula for the drawdown is given by:

$$Drawdown = (high\ water\ mark\ current\ cumulative\ value) /\ high\ water\ mark\ ). \qquad (13)$$

The current drawdown is defined as a strategy's drawdown at the end of December 2018, which is the last day that returns of the portfolios are estimated. The maximum drawdown is defined as the worst drawdown the strategy has experienced over the period and is an indicator of how bad the strategies have performed in the worst periods since the portfolio formation (Pedersen, 2015).

4.5 HAC test for Sharpe ratio comparison

Testing if the Sharpe ratio of two different investment strategies is the same is an important indicator when analyzing the performance of these strategies. Investors usually turn to the Sharpe ratio as it is a great way of outlining the risk-return relationship of an investment strategy. Multiple techniques have been developed to test if a strategy's Sharpe ratio is significantly better than another strategy's Sharpe ratio. The thesis will rely on the methodology presented in (Ledoit & Wolf, 2008). They argue that traditional testing methods such as the method presented in (Memmel, 2003) are incorrect if returns are correlated over time or if the returns are non-normal. Financial returns are generally non-normal which indicate that an alternative method must be used. They propose multiple solutions on how to adjust for the issues presented, and we decided to focus on the HAC inference as it works well in large samples. Moreover, we use the same notation as presented in (Ledoit & Wolf, 2008) paper.

The difference between the two Sharpe ratio estimates ($\hat{\Delta}$) is calculated as:

$$\hat{\Delta} = \widehat{Sh}_i - \widehat{Sh}_n = \frac{\hat{\mu}_i}{\hat{\sigma}_i} - \frac{\hat{\mu}_n}{\hat{\sigma}_n}. \tag{14}$$

The t-statistic estimator of the difference in Sharpe ratio is calculated as:

$$t_{\hat{\Delta}} = \frac{\hat{\Delta} - \Delta_0}{s(\hat{\Delta})}. \tag{15}$$

The standard error $\hat{\Delta}$ used in the t-statistic above are calculated as:

$$s(\hat{\Delta}) = \sqrt{\frac{\nabla' f(\hat{\upsilon}) \hat{\Psi} \nabla f(\hat{\upsilon})}{T}}, and \tag{16}$$

$$\nabla' f(a,b,c,d) = \left(\frac{c}{c-a^2}, -\frac{d}{(d-b^2)^{1.5}}, -\frac{1}{2}\frac{a}{(c-a^2)^{1.5}}, \frac{1}{2}\frac{b}{(d-b^2)^{1.5}}\right), and$$

where:

$$\hat{\Psi} = \hat{\Psi}_T = \frac{T}{T-4} \sum_{J=-T+1}^{T-1} k\left(\frac{j}{S_T}\right) \hat{\Gamma}_T(j), \tag{17}$$

where:

$$\hat{\Gamma}_T(j) = \begin{cases} \frac{1}{T}\sum_{t=j+1}^{T} \hat{y}_t \hat{y}'_{t-j} & for\ j \geq 0 \\ \frac{1}{T}\sum_{t=-j+1}^{T} \hat{y}_{t+j} \hat{y}'_t & for\ j < 0 \end{cases}, with\ \hat{y}'_t = (t_{ti} - \hat{\mu}_1, r_{tn} - \hat{\mu}_n, r_{ti}^2 - \hat{\gamma}_i, r_{tn}^2 - \hat{\gamma}_n). \tag{18}$$

The two-sided p-value for the null hypothesis that H_0: $\Delta = 0$ is calculated as:

$$\hat{p} = 2\phi\left(-\frac{|\hat{\Delta}|}{s(\hat{\Delta})}\right). \tag{19}$$

It is however known that the HAC inference sometimes have a tendency to reject a true null hypothesis too often in cases where the sample size is small or moderate compared to the estimated significance level, we are therefore aware of this issue when using the HAC test to conclude if a Sharpe ratio is significantly better than another factor's Sharpe ratio (Ledoit & Wolf, 2008).

4.6 Data Sample

4.6.1 General Data Modifications

4.6.1.1 Data Sources and the Stock Universe

Many financial studies of US equities rely on Compustat and CRSP data, due to the relatively high quality and long history of these databases. We therefore also based our analysis on the financial data from these databases. Compustat contains information about firm fundamentals on US equities whereas CRSP contains information about historical stock prices on US securities. Furthermore, we followed the standard procedure of limiting the stock universe to companies traded on AMEX, NYSE, and NASDAQ. Securities trading on other exchanges were excluded by only keeping those with a CRSP exchange code (item EXCHCD) of 1,2 or 3 which corresponds to the NYSE, AMEX, and NASDAQ exchanges.

4.6.1.2 Excluding Foreign Firms

We also followed the common practice of excluding companies incorporated abroad to eliminate the influence of cross-country differences. Additionally, we excluded securities different from common stock such as mutual funds. We achieved this by only retaining firms with a CRSP share code of 10 or 11.

4.6.1.3 Excluding Microcap

For all analyses and factor models, we used the standard method of excluding firms with a market cap below the 20th NYSE percentile. This was done to alleviate the impact of microcap stocks that account for roughly 60% of the number of stocks, but only about 3% of the NYSE, AMEX and NASDAQ market cap (Fama & French, 2008). Many of the anomalies discovered in the microcap universe are likely unexploitable due to the large spread and significant market impact an investor would have because of the low liquidity in these markets (Hou et al., 2015).

4.6.1.4 Excluding Financials

We also adhered to the common practice of excluding financial firms as it is normal for these firm to have high leverage and therefore should not be measured on the same premises as nonfinancial firms (Fama & French, 1992), and this was done by removing all companies with a SIC code starting with "6".

4.6.1.5 Excluding Firms with Multiple Exchange Code Changes

During the analysis, we identified 32 firms that changed the exchange code and later changed back. Some of these exchange code changes looked like punching errors such as exchange codes changing from 12 to 21 in just one observation, we, therefore, excluded these firms from the dataset as well, which did not affect the result substantially.

4.6.1.6 Time Horizon

Our chosen time-period was limited to the period from the end of June 1963 to December 2018. Most academic research using fundamental stock data focuses on the period after 1963 because the Compustat database and its comprehensive data collection were not established before 1962 by Standard and Poor's (Linnainmaa & Roberts, 2016). Standard and Poor's did collect some stock data going back to 1947, but the pre-1963 data is subject to survivorship and backfilling bias (Ball & Watts, 1977; Gerakos & Linnainmaa, 2018). Cashflow variables did however only become available in 1972 and portfolios that are based on the cash flow from operations variable were only formed from June 1972.

4.6.1.7 General Data Preparation Procedures

The Compustat and CRSP databases are two separate databases, and they use different company and security identification numbers. The GVKEY is the specific security identification code that used to identify securities in Compustat, while the CRSP database uses PERMNO. In order to merge the two databases, we used a third dataset that contained both the GVKEY and LPERMNO security identification codes called CCM. The LPERMNO is equivalent to PERMNO in the CRSP dataset. The CCM dataset allowed us to match securities in the two databases with the publishing date for both the company fundamentals and the securities prices. There were however multiple companies that were not traded, and these have been removed from the merged dataset. Moreover, multiple securities were traded without being present in the Compustat database. These securities have also been removed from the merged dataset.

4.6.1.8 Returns

The CRSP database includes both the monthly return variable RET which are the holding period return including reinvested regular stock dividends at the end of each month and RETX which is the holding period return without dividends. Both two return variables have missing values which we replaced with zero. The backtests are therefore not flawless, but still a good indicator of what variables explain the cross-section of returns.

Another issue to take into account is that returns are not adjusted for companies that got delisted over the period. CRSP includes the variable delisting return (item DLRET). Delisting returns were moved to the last day of the month they delist, and then the delisting return was multiplied with the regular dividend-adjusted return (item RET) using the formula:

$$Ret_{adj} = (DLRET + 1) \cdot (RET + 1) - 1. \tag{20}$$

There is however an issue of missing delisting returns. (Shumway & Warther, 1999) find that if missing delisting returns are set to zero percent, a spurious size effect emerges for NASDAQ equities. By using alternative data sources (Shumway & Warther, 1999) estimate that replacing missing values for performance-related delisting returns on NASDAQ by -55% yields a good approximation, while (Shumway, 1997) find -30% for AMEX and NYSE to be the optimal correction for performance-related delisting returns. We implemented both recommendations. The performance-related returns were isolated by replacing missing delisting returns that had a delisting code of 500 or 505–588 with -55% for companies that had an exchange code of 3 and -30% for companies that had an exchange code of 1 or 2. All other delistings with missing delisting returns were set to a delisting return of zero percent.

4.6.2 Obtaining Variables for Value Metrics

4.6.2.1 Obtaining Market Cap

Many variables we used were not directly observable in the databases and we, therefore, combined multiple other variables to obtain them. The total market cap for each security was calculated by multiplying the price variable (item PRC) with shares outstanding (item SHROUT). Companies can, however, have more than one security trading on the exchanges and it was, therefore, necessary to calculate the total market cap for each company by summing the market cap of all its shares based on the unique company identification item (PERMCO). We then replaced the largest of any given firm's shares, measured on market cap, with the summed market cap variable and dropped any other security variables.

4.6.2.2 Obtaining Enterprise Value

To calculate enterprise value (EV), we followed (Novy-Marx, 2013) methodology. Specifically, we used the market cap variable, as calculated above, plus long-term debt (item DLTT), plus debt in current liabilities (item DLC), plus preferred stock, minus cash and short-term investments (item CHE). Preferred stock was calculated using preferred stock redemption value (item PSTKRV), and if missing replaced by preferred stock liquidating value (item PSTKL). If both values were missing, we replaced the value with preferred stock capital total (item PSTK) and finally replaced the value with zero if all preferred stock variables were missing from the data.

4.6.2.3 Obtaining Book Equity

To obtain book equity we followed the procedure in (Fama & French, 1993), using Stockholders' Equity (item SEQ) plus Deferred Taxes and Investment Tax Credit (item TXDITC) minus the value of the outstanding preference shares. We obtained the preference shares the same way as we obtained it for the enterprise value calculation above. For total revenue, gross profits, EBITDA and EBIT we used the available variables in Compustat, (item REVT), (item GP), (item EBITDA) and (item EBIT) respectively.

4.6.2.4 Obtaining the Cash Flow from Operations Variable

Cash flow from operations was not available in the Compustat database for the entire estimation period, so funds from operations (item FOPT) which closely resembles cash flow from operations was used as a substitute until it stopped being available. When FOPT was unavailable the similar variable operating activities cash flow (item OANCF) was used. We used these different variables because FOPT is mostly unavailable after 1990 and the OANCF mostly unavailable before 1987. Neither FOPT nor OANCF were available before 1972, so portfolios for strategies based on cash flow from operations was formed starting from June 1972.

4.6.2.5 Constructing Single Value Metrics

The value metrics we analyzed throughout this study were calculated based on the variables we defined above. Book-to-market was calculated by dividing the book equity variable with market cap, REVEV is total revenue (item REVT) divided by enterprise value, GPEV is gross profits (item GP) divided by enterprise value, EBITDAEV is EBITDA divided by enterprise value, and EBITEV is EBIT divided by enterprise value. To obtain (Novy-Marx, 2013)'s profitability measure, gross profits to assets, we divided the two directly observable variables, gross profits (item GP) by assets total (item AT).

4.6.2.6 Portfolio Formation Methodology Based on Fundamental Data

For the formation of portfolios based on fundamental data, we follow (Fama & French, 1993) approach by lagging company fundamentals until the end of June to make sure most financial statements for the previous year were published at the formation time. The most recently available annual release of the company fundamentals was used, and this implies that companies may be included in the portfolios based on up to one-year-old fundamental data from the formation time.

Other data modifications or procedures will be described when relevant in the results and data analysis section.

5 Results

5.1 Single Metrics

In this section, we first show and discuss the historical performance of the chosen value factors. Next, we present the results for the test of the ability of the Fama French five-factor model, q-factor model and six-factor model to explain the return stemming from these factors. We first investigated if the new factors could generate alpha when controlling with the five-factor model. If this is the case, the factors were then regressed against the q- and six-factor models.

Moreover, the factors were also be decomposed if they generated a significant alpha in relations to the five-factor model. This was done to investigate if the alpha could be captured by only by being short or whether alpha could also be captured with a long-only strategy.

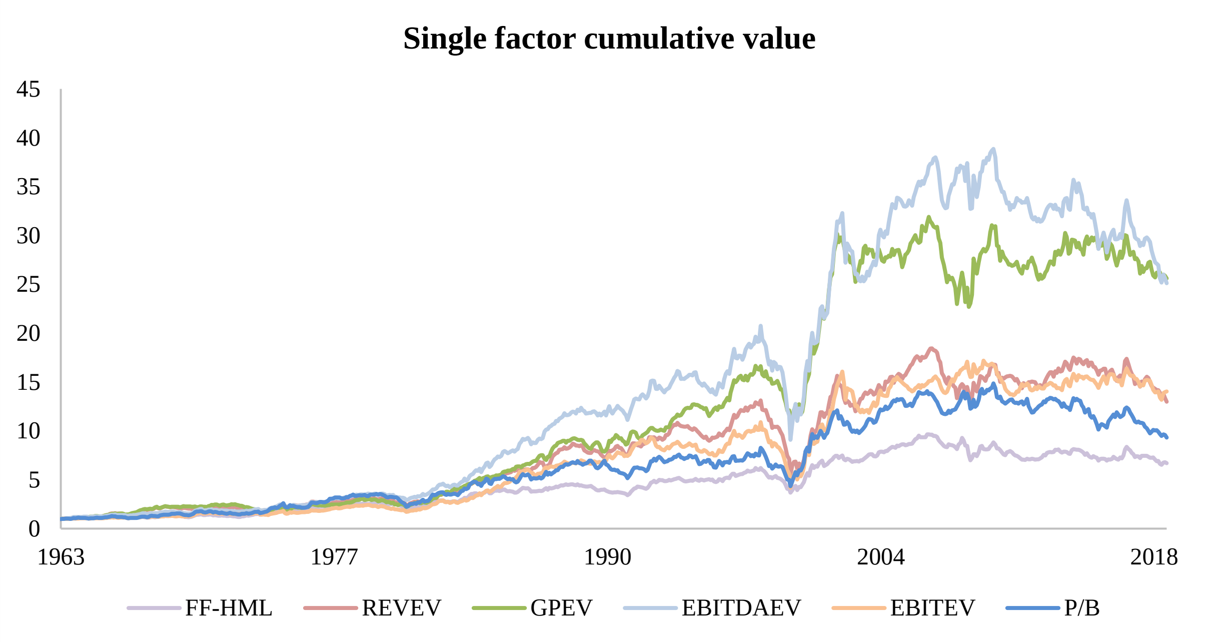

Figure 1. This figure shows the cumulative value of the individual value factors under investigation. The factor strategies included are the original Fama French book-to-market factor with 30th and 70th NYSE value percentile breakpoints, FF-HML; Our replication of the book-to-market factor with 20th and 80th percentile NYSE value breakpoints, P/B; Revenue-to-enterprise-value, REVEV; Gross-profits-to-enterprise-value, GPEV; EBITDA-to-enterprise-value, EBITDAEV; and EBIT-to-enterprise-value, EBITEV, covering 1963 to 2018. Yearly rebalancing is used in portfolio formation.

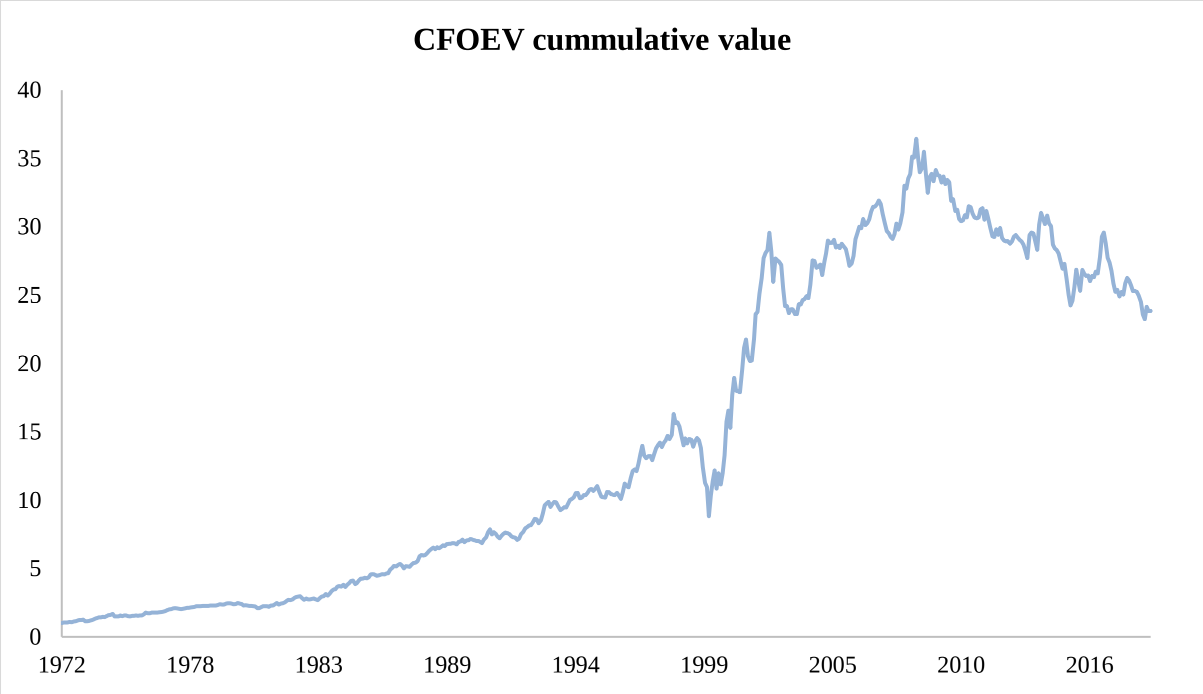

Figure 2. This figure shows the cumulative value of the equal-weighted value factor cash-flow-from-operations-to-enterprise-value, CFOEV, covering 1972 to 2018. Yearly rebalancing is used in portfolio formation.

Out of all the investigated metrics dating back to 1963, the traditional value factor—HML, in isolation, delivered the worst CAGR of 0.29%. From fig. (1) it is evident book-to-market have done especially poorly after it became popularized by (Fama & French, 1993). HML may have fallen victim to Goodhart's law, which dictates as phrased by (Strathern, 1997), "when a measure becomes a target, it ceases to be a good measure." It could also be underperforming because the other value metrics essentially are driven by exposures to different factor premia, that has done better recently.

Single factor five-factor model loadings and historical return performance

	FF-HML	P/B	REVEV	GPEV	EBITDAEV	EBITEV
α		0.01%	-0.14%	0.02%	0.05%	0.03%
		(0.17)	(-1.68)	(0.26)	(0.74)	(0.51)
MKT-RF		-0.06	0.14	0.12	-0.01	-0.07
		(-3.70)	(6.85)	(5.55)	(-0.57)	(-4.42)
SMB		0.07	0.39	0.42	0.07	0.01
		(2.94)	(13.68)	(13.95)	(3.08)	(0.62)
HML		0.93	0.66	0.46	0.83	0.74
		(28.70)	(16.64)	(10.80)	(27.43)	(22.56)
RMW		0.11	0.66	0.63	0.70	0.83
		(3.47)	(16.62)	(14.89)	(23.14)	(24.96)
CMA		0.24	0.09	0.12	0.12	0.01
		(5.00)	(1.45)	(1.97)	(2.64)	(0.18)
R^2		77.3%	58.6%	47.4%	78.6%	75.5%
HWM	9.65	14.86	18.45	31.90	38.86	17.19
CUMV	6.68	9.32	13.00	25.60	25.10	14.05
CAGR	0.29%	0.34%	0.39%	0.49%	0.49%	0.40%
MAXDD	40.89%	47.50%	57.56%	34.96%	56.23%	61.61%
CURRENTDD	30.71%	37.31%	29.51%	19.76%	35.40%	18.28%
SR	0.40	0.37	0.47	0.61	0.56	0.46

Table 1. This table reports the equal-weighted single factors' Fama French alpha, five-factor loadings and other return characteristics for the period 1963 to 2018, for which yearly rebalancing was used. The factor strategies included are the original Fama French book-to-market factor with 30th and 70th NYSE value percentile breakpoints, FF-HML; Our replication of the book-to-market factor with 20th and 80th percentile NYSE value breakpoints, P/B; Revenue-to-enterprise-value, REVEV; Gross-profits-to-enterprise-value, GPEV; EBITDA-to-enterprise-value, EBITDAEV; and EBIT-to-enterprise-value, EBITEV. Alpha is the estimated intercept is α. The Fama French five factors we regress returns on are: market return minus the risk-free rate, MKT-RF; the size factor, SMB; FF-HML and HML are the same; the profitability factor, RMW; and investment factor, CMA. The return characteristics are high water mark, HWM; cumulative value, CUMV; cumulative average gross return, CAGR; maximum drawdown, MAXDD; current drawdown, CURRENTDD; and Sharpe ratio, SR.

CFOEV five-factor model loadings and historical return performance and FF-HML comparison

	FF-HML	CFOEV
α		0.20%
		(2.83)
MKT-RF		-0.06
		(-3.31)
SMB		-0.07
		(-3.01)
HML		0.63
		(19.57)
RMW		0.62
		(18.92)
CMA		0.25
		(5.01)
R^2		77.4%
HWM	8.07	36.44
CUMV	5.59	23.86
CAGR	0.31%	0.57%
MAXDD	40.89%	45.82%
CURRENTDD	30.71%	34.53%
SR	0.42	0.65

Table 2. This table reports the equal-weighted cash-flow-from-operations-to-enterprise-value, CFOEV factor Fama French alpha, five-factor loadings and other return characteristics for the period 1972 to 2018, for which yearly rebalancing was used. CFOEV is also compared to the original Fama French book-to-market factor with 30th and 70th NYSE value percentile breakpoints, FF-HML; Alpha is the estimated intercept α. The five Fama French factors we regress returns on are: market return minus the risk-free rate, MKT-RF; the size factor, SMB; FF-HML and HML are the same; the profitability factor, RMW; and investment factor, CMA. The return characteristics are high water mark, HWM; cumulative value, CUMV; compound annual growth rate, CAGR; maximum drawdown, MAXDD; current drawdown, CURRENTDD; and Sharpe ratio, SR.

The CFOEV factor had the highest CAGR over the period with an astonishing 0.57%, almost twice the return of HML which was only 0.29%. To make our results more comparable to the HML factor we also replicated it with the non-standard 20th and 80th percentile NYSE breakpoints we used for the other factors, and we called this factor price-to-book. The price-to-book factor still failed to beat any of the profitable value metrics on a raw return base with a return of just 0.34%. It is, of course, essential to keep in mind that the estimation period for CFOEV is shorter compared to the other factors estimated. GPEV and EBITDAEV also did remarkably well, both with a CAGR of 0.49%. REVEV and EBIT beat HML and price-to-book by a smaller margin with their CAGR of 0.39% and 0.40% respectively.

HML had the second lowest maximum drawdown of just 40.89%, but our price-to-book measure, while still in the better half, had a worse drawdown of 47.50%. The price-to-book factor also had the worst current drawdown of 37.31%. Interestingly, GPEV had a very low maximum drawdown of only 34.96% and the second smallest current drawdown of 19.76%, while EBITEV had the worst drawdown as high as 61.61% and the second worst current drawdown of 35.40%. There is quite a difference between drawdowns for different value metrics and using profitable value instead of HML appears to provide no hedge against drawdowns.

To get a deeper understanding of the return drivers and whether the outperformance of individual profitable value metrics can be attributed to exposures to other factors, we test how they individually perform against the standard factor models.

5.1.1 Revenue to Enterprise Value (REVEV)

The REVEV factor historically had a CAGR of 0.39% which is the lowest of the alternative value strategies, but still higher than the 0.29% per month of the HML factor. The cumulative value of the REVEV factor is 13.00 which is substantially higher than the cumulative value of the traditional HML factor, which is only 6.68, and this indicates that investors would have been better off investing in the REVEV factor compared to the HML factor. The high water mark of the REVEV factor is 18.45 which is also substantially higher than the HWM of the HML factor which was only 9.65.

The maximum drawdown is, however, worse for the REVEV factor as it has a maximum drawdown of 57.56% compared to the HML factor’s maximum drawdown of 40.89%. The current drawdown of the REVEV factor is relatively similar to the HML factor’s current drawdown as the REVEV factor has a current drawdown of 29.51% compared to the HML factor’s current drawdown of 30.71%. The higher maximum drawdown of the REVEV factor indicates that the CAGR might be risk compensation for the risk of higher maximum drawdowns.

The Sharpe ratio is higher for the REVEV factor as it historically had a Sharpe ratio of 0.47 compared to the HML factor's Sharpe ratio of 0.40, which indicates that the REVEV factor historically had a better return per risk unit compared to the HML factor and that the strategy historically was a better stand-alone strategy.

The REVEV factor generates a negative five-factor alpha of -0.14% with a resulting p-value of 0.09. All else equal buying shares in firms with more revenue per price unit should perform better than firms with less revenue, and it is therefore surprising that the REVEV factor performs so poorly when regressed against the five-factor model. It appears that revenue is a poor predictor of future profitability compared to the existing factors in the model, perhaps because it does not account for either fixed or variable costs and thus provides little information about a firm's profit margin. It can include anything from an undifferentiated manufacturer with small margins, to a company with proprietary intellectual property rights with huge margins.

A large part of the variation of the REVEV factor is explained by the factor's exposure to the HML and RMW factors. From Table. (1) the strategy performs very similarly to the HML factor despite also having high loadings on the RMW and SMB factor. Because the factor has a negative alpha, investors would have been better off shorting the REVEV factor and buying the existing factors from the five-factor model to offset factor exposures.

REVEV q-factor loadings

	HMLGPA
α	0.10%
	(-0.81)
MKT-RF	0.10
	(3.45)
ME	0.28
	(0.04)
I/A	0.79
	(0.06)
ROE	0.15
	(0.05)
R^2	23.1%

Table 3. This table reports q-factor loadings and t-statistics for the REVEV factor. α is the estimated intercept. The q-factors we regress returns on are market return minus the risk-free rate, MKT-RF; the size factor, ME; the investment factor, I/A; and the profitability factor, ROE.

The REVEV factor's alpha is however insignificant when controlling for the q-factors as the alpha is only -0.10% with a resulting p-value of 0.42. That means the q-factor model is able to better explain the performance of the REVEV factor compared to the five-factor model.

REVEV long- and short-side five-factor model loadings and historical return performance

	Long REVEV	Short REVEV
α	-0.06%	0.08%
	(-0.92)	(1.54)
MKT-RF	1.09	0.95
	(71.90)	(70.76)
SMB	0.68	0.29
	(32.41)	(15.65)
HML	0.37	-0.29
	(12.78)	(-11.08)
RMW	0.28	-0.38
	(9.53)	(-14.71)
CMA	0.11	0.02
	(2.47)	(0.56)
R^2	92.7%	93.5%
HWM	2027.38	121.68
CUMV	1610.32	103.40
CAGR	1.11%	0.70%
MAXDD	60.41%	61.92%
CURRENTDD	20.57%	15.02%
SR	0.80	0.60

Table 4. This table reports the equal-weighted single factors' Fama French alpha, five-factor loadings and other return characteristics for the period 1963 to 2018, for which yearly rebalancing was used. The factor strategies included are the long-portfolio used in the construction of REVEV, long REVEV; and the short portfolio used in the construction of REVEV, short REVEV. α is the estimated intercept. The Fama French five-factors we regress returns on are: market return minus the risk-free rate, MKT-RF; the size factor, SMB; FF-HML and HML are the same; the profitability factor, RMW; and investment factor, CMA. The return characteristics are high water mark, HWM; cumulative value, CUMV; cumulative average gross return, CAGR; maximum drawdown, MAXDD; current drawdown, CURRENTDD; and Sharpe ratio, SR.

We then tested if it was possible to captures this alpha by decomposing the factor. We found that the negative alpha is driven by a combination of the two sides of the factor as both sides of the factor were insignificant when regressed against the five-factor model. The long-side of the factor generated an alpha of -0.06% per month with a resulting p-value of 0.36, while the short-side of the factor generated an alpha of 0.08% with a p-value of 0.12

5.1.2 Gross Profits to Enterprise Value (GPEV)

The GPEV factor had a CAGR of 0.49% per month which is higher than the CAGR of 0.29% for HML. The current value of the GPEV factor is 25.60 which is also substantially higher than the current value of the HML factor which is only 6.68. This indicates that investors who invested in the GPEV factor would have 3.83 times as much money as investors who invested in the HML factor proposed by (Fama & French, 1993).

The GPEV factor also has a substantially higher HWM of 31.90 compared to the 9.65 experienced by the HML factor, while simultaneously having a lower maximum drawdown compared to the traditional HML factor as the GPEV factor only have a maximum drawdown of 34.96% compared to the HML factor’s maximum drawdown of 40.89%. The current drawdown is 19.76% which is better than the drawdown of the HML factor which currently is experiencing a drawdown of 30.71%. The lower current and maximum drawdowns of the GPEV factor indicates that the GPEV factor performs better in periods where value investment strategies underperform growth stock, and this lower maximum drawdown also makes the strategy more compelling to more risk-averse investors relative to HML. The higher CAGR, therefore, does not seem to be a compensation for maximum drawdown risk.

The GPEV factor also has a higher Sharpe ratio compared to HML and price-to-book. The GPEV factor has a Sharpe ratio of 0.61 for the period compared to 0.41 for HML indicating that the GPEV factor has a substantially higher return per risk unit. Just because GPEV is superior in isolation, it may not be superior in combination with the other factors from the five-factor model as it has high loadings on both the HML, SMB, and RMW factors. The factor generates a positive alpha, but this alpha is insignificant when controlling for five-factor exposures, as it only has an alpha of 0.2% per month with a p-value as high as 0.80. Existing factors, therefore, explain the GPEV factor's returns and the positive alpha can easily be attributed to randomness. The GPEV factor is also highly correlated with the RMW factor which contributes to the insignificance of its alpha, and this makes sense as they are both based on income statement items.

5.1.3 EBITDA to Enterprise Value (EBITDAEV)

The EBITDAEV factor has a CAGR of 0.49% per month which is substantially higher than the CAGR of the HML factor. The current value of the factor is also higher for the EBITDAEV factor as it has a current value of 25.10 compared to the HML factor’s cumulative value of 6.68, and this indicates that investors would have 3.75 times as much money, had they invested in the EBITDAEV factor instead of the HML factor.

The high water mark is also better for the EBITDAEV factor which has a high water mark of 38.86 compared to the HML factor’s high water mark of 9.65. The maximum drawdown of the EBITDAEV factor is however more substantial than for the EBITDAEV factor as it has a maximum drawdown of 56.23% compared to the 40.89% experienced by the HML factor. The current drawdown is also worse for the EBITDAEV factor as it has a drawdown of 35.40% while the HML factor only has a current drawdown of 30.71%. The higher maximum drawdown indicates that the EBITDAEV factor performs worse in periods where value strategies underperform and that the higher CAGR might be a compensation for maximum drawdown risk. The EBITDAEV factor does, however, have a significantly higher Sharpe ratio compared to the HML Factor. The EBITDAEV had a Sharpe ratio of 0.56 compared to the HML factor’s Sharpe ratio of just 0.40 and price-to-book Sharpe ratio of 0.37. The higher Sharpe ratio indicates that the return per risk unit is better for the EBITDAEV factor and this implies that the better performance does not seem to be a compensation for the total risk but could be a compensation for skewed returns with higher drawdowns.

The factor has an alpha of 0.05% when regressed against the five-factor model and a p-value of 0.46 which is insignificant at all the standard significance levels. The reason why this factor can deliver a higher CAGR compared to the HML factor is likely due to the EBITDAEV factor having a loading of 0.70 on the RMW factor while also having a loading of 0.83 on the HML factor. The factor is therefore not able to better capture a premium beyond what can be explained by existing factors in the five-factor model. It makes sense that the EBITDAEV factor is unable to generate alpha, as EBITDA is closely related to the operating profit, which the RMW factor is based on.

5.1.4 EBIT to Enterprise Value (EBITEV)

The EBITEV factor has a CAGR of 0.40% per month which is higher than the historical CAGR of the HML factor which was only 0.29% per month in the same period. The EBITEV factor has a current value of 14 compared to just 6.68 for HML factor. This means that investors also would have been better of investing in EBITEV compared to HML.

The high water mark is substantially higher than the HML factor's as it has a high water mark of 17.19 compared to the 9.65 delivered by the HML. The maximum drawdown of the EBITEV factor is however substantially higher compared to the HML factor as it had a maximum drawdown of 61.61% compared to the HML factor's maximum drawdown of 40.89%, and this indicates that the EBITEV factor performs substantially worse in periods where value investment strategies underperform. The current drawdown of the EBITEV factor is, however, better as it only has a current drawdown of 18.28% while the HML factor has a drawdown of 30.71%. The higher maximum drawdown of the strategy indicates that the higher return of the factor might again be compensation for the risk of higher maximum drawdowns.

The Sharpe ratio is also higher for the EBITEV factor with a Sharpe of 0.46, slightly ahead of the HML factor's Sharpe ratio of 0.41. This indicates that the higher CAGR is not a result of the strategy being highly volatile, as it has a higher return per risk unit compared to the HML factor, and the higher CAGR again might be compensation for a skewed return distribution.

When tested against the five-factor model, the EBITDAEV factor has an alpha of 0.03% per month with a p-value of 0.61. The factor fails to deliver significant alpha as the EBITDAEV factor is highly correlated with the RMW factor. The outperformance of the EBITDAEV factor compared to the HML factor is likely a result of the high loadings on RMW and SMB which indicates that the EBITDA factor is highly correlated with existing factors. With the insignificant alpha of EBITDAEV, there is not enough evidence to tell whether it captures a premium beyond the factors in the five-factor model, despite the Sharpe ratio surpassing that of the traditional HML factor.

5.1.5 Cash Flow from Operations to Enterprise Value (CFOEV)

The CFOEV factor has a CAGR of 0.57% which is substantially higher than the 0.31% delivered by the HML factor over the same period. The current value of the strategy is also appreciably higher compared to the HML factor as it has produced a cumulative value of 23.86 over the period while the HML only delivered 5.59.

The CFOEV factor also performed significantly better over the period compared to the other single metric factors, as it had a high water mark of 36.44. This high water mark is close to the highest water mark achieved by the other single metric factors where the best was 38.86 delivered by the EBITDAEV factor, and which were especially impressive as the factor is formed nine years after the other factors.

The CFOEV factor had a maximum drawdown of 45.82% which were worse than the 40.89% experienced by the HML factor. The CFOEV factor also has a higher current drawdown of 34.53% by the end of 2018, compared to the Fama and French HML factor's current drawdown of 30.71%, which indicates the higher CAGR might be a compensation for the higher maximum drawdown risk as the CFOEV factor does substantially worse in periods where value investment strategies underperform.

The CFOEV factor has a Sharpe ratio of 0.65, substantially higher than the HML factor's Sharpe ratio of just 0.42 for the same period. The higher Sharpe ratio indicates that the CFOEV factor might be a better stand-alone strategy compared to the HML factor. This higher Sharpe ratio also suggests that the higher CAGR is not a compensation for the total risk as the return per risk unit is higher for the CFOEV factor compared to the HML factor. It is however also questionable if the CFOEV is still generating excess returns as the strategy has barely recovered since the factor topped out by the end of December 2008 which is ten years before the last estimated returns for the CFOEV factor.

The CFOEV factor is the first and only single metric factor that can generate significant positive alpha when regressed against the five-factor model. The CFOEV factor's alpha in relation to the five-factor model is 0.20% with a p-value of approximately 0.

CFOEV q-factor loadings

	CFOEV
α	0.24%
	(2.20)
MKT-RF	-0.11
	(-4.24)
ME	-0.14
	(-3.90)
I/A	0.98
	(16.70)
ROE	0.20
	(4.78)
R^2	47.4%

Table 5. This table reports q-factor loadings and t-statistics for the cash-flow-from-operations-to-enterprise-value factor, CFOEV. α is the estimated intercept. The q-factors we regress returns on are market return minus the risk-free rate, MKT-RF; the size factor, ME; the investment factor, I/A; and the profitability factor, ROE.

We then tested the model against the q-factor model where the CFOEV factor generated an alpha of 0.24% and is still significant with a t-statistic of 2.20 and a p-value of 0.03. The most significant loading is the investment factor loading (I/A) of 0.98. Finally, we test whether the alpha of CFOEV survives the six-factor model.

CFOEV six-factor loadings

	CFOEV
α	0.04%
	(0.51)
MKT-RF	0.01
	(0.65)
SMB	-0.07
	(-2.45)
HML	0.84
	(27.64)
BAB	0.06
	(2.51)
QMJ	0.57
	(13.01)
UMD	0.01
	(0.68)
R^2	73.6%

Table 6. This table reports six-factor loadings and t-statistics for the cash-flow-from-operations-to-enterprise-value factor, CFOEV. α is the estimated intercept. The six-factors we regress returns on are: market return minus the risk-free rate, MKT-RF; the size factor, SMB; the Fama French book-to-market factor, HML; the betting-against-beta factor, BAB; the quality-minus-junk factor, QMJ; and the momentum factor, UMD.

The factor's alpha does, however, become insignificant when regressed against the six-factor model as the factor only generates an alpha of 0.04% with a resulting p-value of 0.61. Nonetheless, we still decomposed the factor to investigate if the factor's alpha on the five-factor model was driven by the long-side, short-side or a combination of the two. We found that the short-side of the factor was able to generate significant negative alpha when regressed against the five-factor model with an alpha estimate of -0.20% with a resulting p-value of 0.002. The long-side of the strategy was however insignificant with an alpha estimate of 0.02% with a p-value of 0.74. The alpha of the short-side of the factor was however not significant when regressed against the six-factor model, and it can therefore not be concluded that the short-side of the strategy can generate abnormal returns in relations to the six-factor model.

We tested the robustness of the findings in relation to the five-factor model. We first updated the cut off points from 20% and 80% to 10% and 90%, to investigate if the strategy is still able to generate alpha when regressed against the five-factor model with different cut off points. The strategy's alpha when controlling for the five-factor model is 0.26%, which has a resulting p-value of approximately 0. We then updated the cut off points to 30% and 70%, and this updated factor had a significant alpha of 0.15% which had a resulting p-value of approximately 0.01. The CFOEV factor, therefore, seems rather robust against the five-factor model despite not being able to generate excess returns when regressed against the six-factor model.

Moreover, the factor also survives the q-factor model at the 10% significance level when updating the weights to 10% and 30%, as the factor generates an alpha of 0.29% with a t-statistic of 1.95 when using 10% and 90% cutoffs. And does even better when using 30% and 70% cutoffs as it survives the q-factor model at the 5% significance level with a t-statistic of 2.15 and generates an alpha of 0.20%. The CFOEV factor therefore also seem fairly robust when regressed against the q-factor model (see tab. (34) in appendix).

5.1.6 Conclusion on Single Metrics

All the profitable value factors had higher, CAGR and Sharpe ratios compared to HML factor which indicates that these factors are better as a stand-alone strategy compared to the HML factor. If an investor could only invest in one value factor, profitable value metrics are preferable to price-to-book.

Profitable value factors performed similarly to price-to-book on drawdowns, and surprisingly GPEV with the best returns of the estimated factors starting in 1963 also had the lowest maximum drawdown out of the tested factors, while CFOEV and GPEV also beat the HML factor on maximum drawdown.

Conclusively, most of the profitable value strategies were explained by the existing asset pricing models—five-factor, q-factor, and six-factor models, apart from the CFOEV factor that was able to generate significant alpha unexplained by the five-factor and q-factor models. The alpha was however driven by the short-side, meaning that investors would have to be able to effectively short low-scoring CFOEV stocks to exploit the anomaly. The factor's alpha was however insignificant when regressed against the six-factor model.

5.2 Compound Value Factor Models

Compound factors five-factor model loadings and historical return performance

	CPIRE	CPEXRE	CPBEEX
α	0.04%	0.07	0.06%
	(0.60)	(1.04)	(0.92)
MKT-RF	-0.00	-0.03	-0.04
	(-0.19)	(-1.71)	(-2.82)
SMB	0.17	0.13	0.15
	(6.59)	(5.47)	(6.54)
HML	0.84	0.84	0.90
	(25.20)	(25.95)	(29.48)
RMW	0.61	0.62	0.37
	(18.01)	(18.93)	12.04
CMA	0.13	0.13	0.22
	(2.53)	(2.53)	4.73
R^2	77.88%	79.48%	82.0%
HWM	24.53	30.96	22.93
CUMV	15.18	18.64	15.18
CAGR	0.49%	0.53%	0.49%
MAXDD	51.91%	49.76%	48.59%
CURRENTDD	38.13%	37.18%	33.79%
SR	0.56	0.58	0.54

Table 7. This table reports the equal-weighted compound factors' Fama French alpha, five-factor loadings and other return characteristics for the period 1972 to 2018, for which yearly rebalancing was used. The factor strategies included are: the compound factor of all single factors, CPIRE; the compound factor excluding revenue-to-enterprise-value, CPEXRE; the compound with 50% in book-to-market and 50% in the compound excluding book-to-market, CPBEEX. α is the estimated intercept. The Fama French five factors we regress returns on are: market return minus the risk-free rate, MKT-RF; the size factor, SMB; FF-HML and HML are the same; the profitability factor, RMW; and investment factor, CMA. The return characteristics are high water mark, HWM; cumulative value, CUMV; compound annual growth rate, CAGR; maximum drawdown, MAXDD; current drawdown, CURRENTDD; and Sharpe ratio, SR.

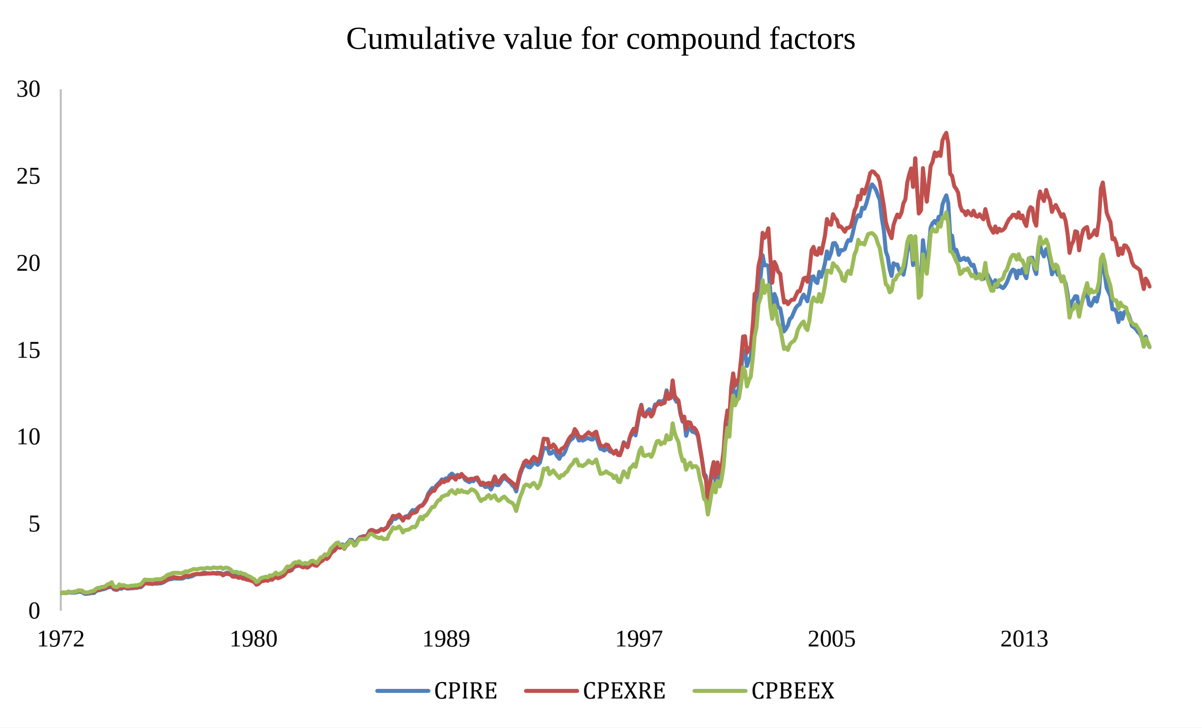

Figure 3. This figure shows the cumulative value for the equal-weighted compound factors' Fama French alpha, five-factor loadings and other return characteristics for the period 1972 to 2018, for which yearly rebalancing was used. The factor strategies included are: the compound factor of all single factors, CPIRE; the compound factor excluding revenue-to-enterprise-value, CPEXRE; the compound based 50% on book-to-market and based 50% on an equal-weighted rank on the GPEV, EBITEV, EBITDAEV, CFOEV metrics, CPBEEX.

To test the second hypothesis, we investigated if a compound measure of multiple value metrics were able to generate abnormal returns unexplained by the popular factor models. The compound factor is constructed by giving each company a score based on each value metrics.

The score a company gets is based on its percentile so the company with the best score on a metric get 100 and the worst get a score of zero on that metric. The score from each parameter is then combined and the factor formed by going long firms above the 80th percentile of compound scores and shorting firms below the 20th percentile. The portfolio holdings are equally weighted at formation with yearly rebalancing.

The ranked score methodology is similar to how (Stambaugh, Yu and Yuan 2015) constructed their mispricing measure, but we base the compound on value metrics. They argue that averaging rankings across anomalies can diversify anomaly-specific noise and increasing a measure's ability to identify anomalies. The compound measure can only be tested back to 1972, as we include the operating cash flow metric, which is not available before 1972. Below we present the results from the standard compound metric (CPIRE) as described above. After that, we test the compound (CPEXRE) which is formed by assigning 50% of the weight to the price-to-book ratio and equally weighting all of the other metrics excluding REVEV.

5.2.1 Equal-weighted Compound Including Revenue-to-Enterprise (CPIRE)

We formed a compound factor that includes all the individual value metrics, book-to-market, gross profits-to-enterprise value, EBITDA-to-enterprise-value, EBIT-to-enterprise value, and cash-flow-from-operations-to-enterprise value. The CPIRE factor has a CAGR of 0.49% which is substantially higher than the CAGR of 0.31% delivered by the HML factor in the same period. The CPIRE factor has achieved a current cumulative value of 15.18, whereas the HML factor only delivered a current cumulative value of 5.59 over the same period.

The CPIRE factor also had a substantially higher high water mark of 24.53 which was also considerably higher than the high water mark of the HML factor which was only 8.07. The CPIRE had a maximum drawdown of 51.91% which is substantially worse than the HML factor’s maximum drawdown which was only 40.89% in the same period. The current drawdown is also worse for the CPRIE factor having lost 38.13% of its value since it reached its high water mark, while the HML factor only has lost 30.71%, indicating the high CAGR might be compensation for the factor being riskier compared to the HML factor.

The Sharpe ratio is, however, substantially better for the CPRIE factor with 0.56 compared to the lower HML Sharpe ratio of 0.42, indicating that the CPRIE would be a better stand-alone strategy compared to the HML factor, which also suggests that the higher return is not a risk compensation for the total risk, but might be a compensation for skewed returns. The CPIRE factor did, however, have a worse Sharpe ratio compared to the CFOEV and GPEV factors indicating that these two strategies have a better risk-return tradeoff individually.

The factor is however insignificant when regressed against the five-factor model as it has an alpha of only 0.04% and a p-value of 0.55. The factor’s returns are likely explained by the fact that the factor has a loading of 0.61 on the RMW factor while at the same time having a loading of 0.84 on the HML factor, which again might be a result of the components of the factor being relatively similar to the existing factors in the model.

5.2.2 Equal-weighted Compound Excluding Revenue-to-Enterprise-Value (CPEXRE)

Based on the poor performance of REVEV measured on alpha and the idea that revenue may be too noisy a measure for future predictability attributed to the high position on the income statement, we investigated how the compound fared without it. We created the compound factor excluding the REVEV component. The CPEXRE had a CAGR of 0.53% in the period from June 1972 to December 2018. The factor delivered a current cumulative value of 18.64 over the period with a high water mark of 30.96 which is substantially higher than the current cumulative value achieved by the HML factor which was only 5.59 with a high water mark of 8.07.

The maximum drawdown of the strategy is 49.76% which was relatively similar to the compound measure which included the revenue component which had a maximum drawdown of 51.91%. This maximum drawdown was also substantially worse compared to the HML factor which only had a maximum drawdown of 40.89%. The current drawdown of the two compound strategies was very as the CPEXRE factor had a current drawdown of 37.18% while the CPRIE factor had a current drawdown of 38.14%, which were also worse than the current drawdown of the HML which had 30.71%. The worse drawdowns support the risk-based hypothesis as the higher CAGR might be compensation for the strategies being more exposed to big crashes.

Interestingly, the factor still had a high Sharpe ratio of 0.58 compared to the HML factor's Sharpe ratio of just 0.42. The Sharpe ratio of the CPEXRE was only slightly higher than the Sharpe ratio of CPIRE factor, which was 0.56, indicating that removing the revenue component did not alter the performance of the compound factor substantially. Interestingly was the CPEXRE factor's higher CAGR, not a result of having a higher total risk as the return per risk unit is higher for the factor compared to the HML factor.

The factor's alpha was also insignificant when regressed against the five-factor model with just 0.07% and a p-value of 0.30. Due to its insignificance at any of the standard significance levels, it is not possible to conclude that the CPEXRE caries a premium unexplained by the five-factor model. Unsurprisingly, the factor had high loadings on both the HML and RMW factors as it was constructed using scores on book-to-market and various profitable value metrics. The CPEXRE did not only fail to deliver alpha when regressed against the five-factor model but also performed significantly worse than the HML factor in periods when value underperformed growth.

5.2.3 Half Compound Half Book-to-price (CPBEEX)

Because the profitable value metrics all focus on buying cheap earnings power and are very similar, since they combine elements from profitability and price while at the same time do not take into account the amount of book equity in a firm, we tested a compound measure with a 50% weight in book-to-market and 50% in the other alternative metrics excluding revenue. The updated weights got included to investigate if a compound measure that weights the profitability and price-to-book measures equally were better at capturing value premia. We still excluded REVEV as the CAGR of the CPEXRE was higher compared to the compound factor which included REVEV. The CPBEEX factor had a CAGR of 0.49% per month which was worse than the equally weighted compound measure that excluded revenue and had a CAGR of 0.53% per month. This indicates that the CPEXRE historically were not better at capturing the value premium compared to the CPEXRE factor. The CPBEEX factor's cumulative value was almost identical to the cumulative value of the CPIRE factor as they both had a cumulative value of approximately 15.18, indicating the CPBEEX factor is unable to generate a significantly better performance compared to the CPIRE factor. The factor also underperformed the CPEXRE factor which again suggests that the new weight does not improve performance compared to the methodology used in the other two compound factors. The high water mark is also lower as the CPBEEX factor only had a high water mark of 22.93 which is lower than the HMW of the CPIRE factor's which had 24.53 and the CPEXRE factor's HWM of 30.96.

The maximum drawdown was also similar to the CPEXRE, as the CPBEEX factor had a maximum drawdown of 48.59% compared to the CPEXRE factor's maximum drawdown of 49.76%. The current drawdown is, however, better for the CPBEEX factor as it only has a current drawdown of 33.79% compared to the CPEXRE factor's current drawdown of 37.18% and this again indicates that the higher CAGR might be compensation for the maximum drawdown risk as the maximum drawdown were still substantially higher for the compound factor compared to the HML factor's maximum drawdown.

The Sharpe ratio of the CPBEEX factor is 0.54, substantially higher than the Sharpe ratio of the HML factor, but lower than the Sharpe ratio of the CPEXRE of 0.58. The factor's alpha was however still insignificant when regressed against the five-factor model as it only generated an alpha of 0.06% with a p-value of 0.36. CPBEEX was also fully explained by the five-factor model. The loading on RMW for CPBEEX dropped a lot in comparison to CPIRE and CPEXRE, as it became 0.37, where CPIRE and CPEXRE had respective loadings of 0.61 and 0.62. As anticipated with the higher weight in book-to-market the loading on HML increased. The HML loading was almost 0.90, an increase of about 0.06 compared to both CPIRE and CPEXRE.

5.2.4 Concluding Remarks on Compound Value Factors

Neither one of the compound factors generated alpha beyond what could be explained by the five-factor model. The CPEXRE factor had the best performance in terms of CAGR, while the CPBEEX factor had the best performance in terms of maximum drawdown, but generally, the compound factor models had very similar performance in terms of both return, drawdowns and Sharpe ratios. The higher Sharpe ratios of the compound value factors compared to the HML factor could be explained by loadings on RMW, CMA, and SMB in the five-factor model. The compound factors performed substantially better than the HML factor as stand-alone strategies.

All the compound factors we tested only beat three out of the six single metrics on a maximum drawdown basis. Thus, the compound factors do not necessarily improve maximum drawdowns in comparison to the single metrics but have a tendency to perform more like the average of the individual metrics, which intuitively makes sense. On a CAGR return basis, the compound factor that included all single metrics (except REVEV) and the 50% book-to-price factor did slightly better than the median for single metric as all compound metrics outperformed REVEV, EBITEV, and price-to-book. On a CAGR basis, they also performed equivalent to GPEV and EBITDAEV. The compound factor excluding REVEV was only beaten by the CFOEV factor. On a Sharpe ratio basis, the compound factors generally did better, as the combination of all value metrics had a Sharpe ratio of 0.56, and the one excluding REVEV had a Sharpe ratio of 0.58, only surpassed by CFOEV's Sharpe ratio of 0.65. The 50% book-to-price compound metric was only exceeded by EBITDAEV and CFOEV. In conclusion, the compound factors were unable to improve the five-factor model.

5.3 Are Profitable Value Metrics Driven by a High Profitability Subsection?

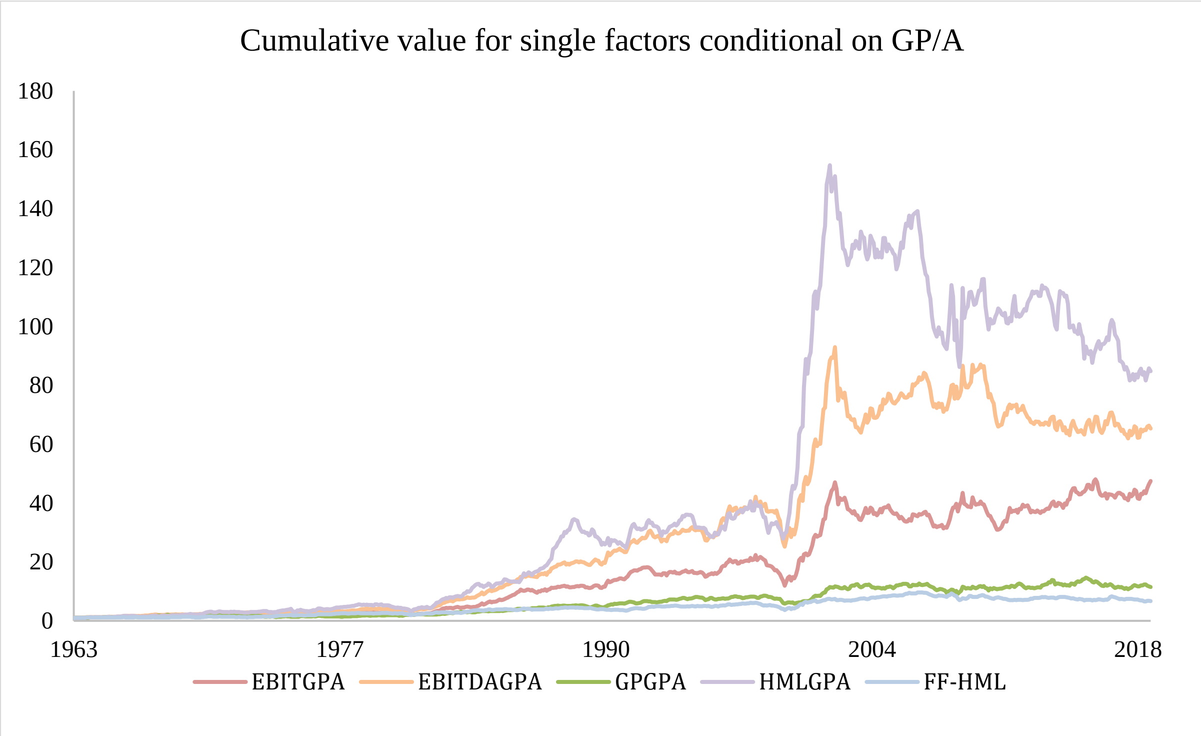

Figure 4. This figure shows the cumulative value for the single factors conditional on GP/A. They were formed based on the six value GP/A sorted portfolios, by going long the high-value-high-GP/A portfolio and short the low-value-low-GP/A portfolio. They cover the period from 1963 to 2018.

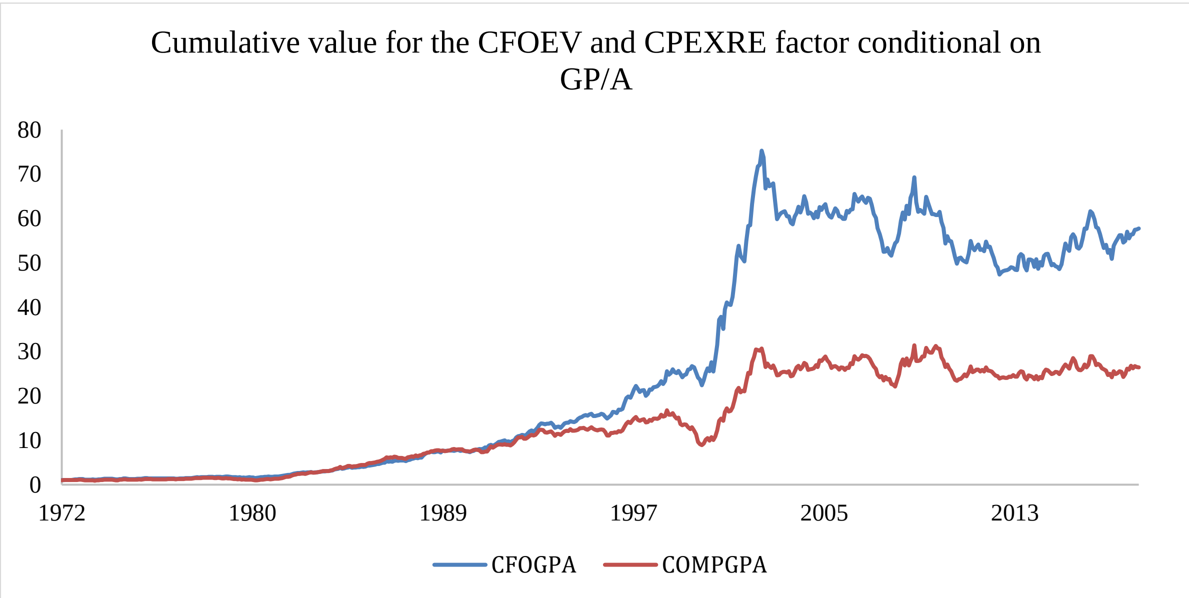

Figure 5. This figure shows the cumulative value for the CFOEV factor conditional on GP/A, CFOGPA and the compound factor that excluded REVEV conditional on GP/A, COMPGPA. They were formed based on the six value GP/A sorted portfolios, by going long the high-value-high-GP/A portfolio and

short the low-value-low-GP/A portfolio. They are shown together as they both cover the period from 1972 to 2018.

Below are the results for various tests of whether the profitable value metrics are driven by a subsection of firms with high gross profitability-to-asset (GP/A) and low valuations and vice versa for the short-side. These tests aim to address hypothesis (3) inspired by (Novy-Marx, 2013) claim that a subsection of firms drives the HML value premium. We chose to exclude the revenue-based factor, REVEV, as it performed so poorly when analyzed individually.

As shown in eq. (6) profitable value metrics already include an element of profitability, so we investigated whether a higher (lower) gross profits-to-asset was also beneficial (detrimental) if you paid the same per unit of gross profits, EBITDA, EBIT or cash flow from operations. At one hand, a higher profit margin means that firms are further away from being distressed, i.e., in danger of moving into negative profitability territory and may indicate that great growth opportunities lie ahead. On the other hand, higher margins could tend to mean revert downwards to the detriment of a profitable value GP/A combination.

We followed (Novy-Marx, 2013)'s methodology for testing portfolios of various GP/A ratios while controlling for the value metric under consideration, by making a conditional sort for each of the value metrics. The portfolios were formed using 20th and 80th percentile NYSE value breakpoints and tertiles rather than deciles. Within each value portfolio we then divided the firms into two portfolios, those with GP/A above the median and those with GP/A below, this gave us two portfolios with the same number of firms within each value category. We ended up with six-portfolios: high-value-high-GP/A, medium-value-high-GP/A, low-value-high-GP/A, high-value-low-GP/A, medium-value-low-GP/A, low-value-low-GP/A. This approach allows us to compare the magnitude of the improvement or deterioration when moving between high and low value or GP/A ratios as well as across the different combinations of GP/A and value metrics.

The reason why we did not use the standard double sort methodology here was that firms with the highest profitability and lowest valuations were scarce. This means that were we to use the standard method we would end up with undesirable portfolios as some portfolios would only hold a tiny number of firms, undesirable for investors as they would become exposed to idiosyncratic firm risk.

We constructed a factor for each value metric by going long high-value-high-GP/A and short low-value-low-GP/A. This makes it convenient to compare to the single metrics only sorted on a value metric. If this value GP/A factor improved the performance of single metrics, it implies that a factor which buys the high-value low GP/A portfolio and short the low-value high GP/A portfolio underperformed single metrics, we, therefore, do not need to make that direct comparison. We present both the statistics for the six-portfolios for each factor to be able to compare how the subsections behave differently within each value section, as well as the factor statistics described above.

Single factors conditioned on GP/A, five-factor model loadings and historical return performance

	HMLGPA	GPGPA	EBITDAGPA	EBITGPA
α	0.29%	0.05%	0.23%	0.29%
	(2.30)	(0.48)	(2.38)	(2.75)
MKT-RF	-0.08	0.16	-0.02	-0.08
	(-2.59)	(6.12)	(-0.83)	(-3.20)
SMB	0.43	0.47	0.27	0.18
	(9.85)	(12.69)	(8.09)	(4.94)
HML	0.84	0.00	0.44	0.33
	(14.01)	(0.00)	(9.41)	(6.63)
RMW	0.28	0.69	0.92	0.95
	(4.68)	(13.28)	(19.68)	(18.70)
CMA	0.23	-0.03	0.08	-0.00
	(2.59)	(-0.35)	(1.18)	(-0.05)
R^2	48.5%	33.1%	49.6%	44.63%
HWM	154.75	14.68	92.95	48.06
CUMV	84.83	11.49	65.32	47.50
CAGR	0.67%	0.37%	0.63%	0.58%
MAXDD	47.3%	39.40%	40.36%	46.64%
CURRENTDD	45.18%	21.72%	29.72%	1.18%
SR	0.62	0.45	0.71	0.64

Table 8. This table reports five-factor loadings and t-statistics for the factors that were formed based on the six value GP/A sorted portfolios, by going long the high-value-high-GP/A portfolio and short the low-value-low-GP/A portfolio. α is the estimated intercept. The Fama French five factors we regressed returns on were: market return minus the risk-free rate, MKT-RF; the size factor, SMB; FF-HML and HML are the same; the profitability factor, RMW; and investment factor, CMA. The return characteristics is the high water mark, HWM; cumulative value, CUMV; compound annual growth rate, CAGR; maximum drawdown, MAXDD; current drawdown, CURRENTDD; and Sharpe ratio, SR. The factor data for all the above factors start in 1963 and run until 2018.

CFOGPA, five-factor model loadings, and historical return performance

	CFOGPA	COMPGPA
α	0.42%	0.21%
	(3.97)	(1.73)
MKT-RF	-0.05	-0.07
	(-1.96)	(-2.37)
SMB	0.12	0.29
	(3.29)	(6.86)
HML	0.21	0.46
	(4.35)	(8.14)
RMW	0.75	0.78
	(15.08)	(13.65)
CMA	0.26	0.18
	(3.40)	(2.05)
R^2	42.9%	45.6%
HWM	75.26	31.41
CUMV	57.72	26.40
CAGR	0.73%	0.59%
MAXDD	37.10%	47.07%
CURRENTDD	23.30%	15.94%
SR	0.85	0.61

Table 9. This table reports five-factor loadings and t-statistics for the cash-flow-from-operations-to-enterprise-value factor that were formed based on the six value GP/A sorted portfolios, by going long the high-value-high-GP/A portfolio and short the low-value-low-GP/A portfolio, CFOGPA, as well as the compound factor excluding REVEV also conditional on GP/A, COMPGPA. α is the estimated intercept. The Fama French five factors we regressed returns on were: market return minus the risk-free rate, MKT-RF; the size factor, SMB; FF-HML and HML are the same; the profitability factor, RMW; and investment factor, CMA. The return characteristics is: high water mark, HWM; cumulative value, CUMV; compound annual growth rate, CAGR; maximum drawdown, MAXDD; current drawdown, CURRENTDD; and Sharpe ratio, SR. The factor data for all the above factors start in 1972 and run until 2018.

5.3.1 Book-to-Market and GP/A

Book-to-market and GP/A double sorted portfolio, five-factor alpha and return characteristics

		High value	Medium value	Low value
Low profitability	α	-0.03%	-0.17%	-0.27%
	α t-statistic	(0.36)	(2.95)	(3.66)
	CAGR	1.09%	0.85%	0.43%
	MAXDD	48.22%	49.24%	73.46%
	CURRENTDD	20.73%	16.60%	19.12%
	SR	0.58	0.42	0.14
High profitability	α	0.02%	-0.03%	0.09%
	α t-statisticistic	(0.22)	(0.61)	(1.75)
	CAGR	1.20%	1.06%	0.88%
	MAXDD	68.90%	54.61%	61.18%
	CURRENTDD	18.68%	18.90%	14.54%
	SR	0.56	0.54	0.40

Table 10. This table reports the portfolio characteristics of stocks that were first sorted on book-to-market with 20th and 80th NYSE breakpoints, and then within each portfolio, 50% of firms were sorted into a high (low) GP/A (profitability) portfolio. We ended up with six portfolios. α is the five-factor alpha followed by its t-statistic, α t-statistic. Return characteristics included are compound annual growth rate, CAGR; maximum drawdown, MAXDD; current drawdown, CURRENTDD; and Sharpe ratio, SR.

Tab. (10) shows the characteristics of the six value GP/A portfolios for book-to-market. There was a tendency for the CAGR to increase as the book-to-market and GP/A ratios, with the high-value-high-GP/A portfolio having had the highest CAGR over the period.

The highest monthly CAGR of 1.20% for the high-value-high-quality portfolio may, however, be a risk-premium investors require in compensation for maximum drawdown risk as it dropped an excruciating 68.90% while for the high-value-log-GP/A portfolio we only saw a drop of 48.22%. This was a change of more than 20% in maximum drawdown, while CAGR changed by just 0.11%. The high-value-high-GP/A portfolio may, therefore, have contained a large number of firms with high, but severely fragile profit margins that had a tendency to get walloped in economic downturns. This maximum drawdown of the high-value-high GP/A was also worse than any of the maximum drawdowns experienced by any of the single value metrics as well as the compound value portfolios, as the highest maximum drawdown for any of these factors was 61.61% experienced by the EBITEV factor.

These findings with detrimental implications for the long-only high-value-high-GP/A portfolio are were not addressed by (Novy-Marx 2013), who argued that GP/A almost ubiquitously improved the HML factor. Interestingly, were the maximum drawdown of the low-value-low-GP/A portfolio also extremely high with a maximum drawdown of 73.46%, and this may help offset the drawdown for the long-short factor portfolio assuming that the drawdowns for the long- and short-side tend to happen simultaneously. This low-value-low-GP/A also had the worst CAGR of 0.43% which was substantially lower than that of the low-value-high-GP/A portfolio of 0.88%. These results suggest that investors should avoid long positions in these unprofitable and expensive "growth trap" stocks by all means.

Regarding the Sharpe ratios of the six-portfolios, there was a tendency for the Sharpe ratio to increase when moving from low value to high value portfolios and low GP/A to high GP/A portfolios, with the exception of moving from the high-value-low-GP/A portfolio to the high-value-high-GP/A portfolio, where the Sharpe ratio slightly decreases from 0.58 to 0.56. The low-value-low-GP/A portfolio had an extremely low Sharpe ratio of 0.14 which were lower than any Sharpe ratio produced by any of the single and compound factors indicating that the return per risk unit is extremely low for the firms in this portfolio.

We regressed the portfolio's returns against the five-factor model to investigate if the existing factors drove the portfolio returns. Both the medium-value-low-GP/A and low-value-low-GP/A both had significantly negative alphas, which strengthens our suspicion that the short-side primarily drives (Novy-Marx, 2013)'s anomaly findings. It also became evident as the high-value-high-GP/A portfolio's five-factor alpha was insignificant. This results strongly suggest that the short-side of the strategy drove the findings of (Novy-Marx, 2013)**.**

These findings are consistent with (Stambaugh, Yu, & Yuan, 2012, 2014, 2015) that argue that most factor anomalies are driven by their short-leg, as investors face arbitrage risk and arbitrage asymmetry that makes it both challenging and risky to fully utilize the short-leg alpha. It is not unlikely for our case that these expensive and unprofitable portfolios were comprised of a large number of hard to short stocks and therefore driven by limits to arbitrage. Alternatively, investors may have been afraid to short them, because they contained popular growth stocks held by unsophisticated retail investors that may sustain a high price level for an extended time period or bid them up even higher. Answers to these speculations are beyond the scope of this paper but interesting for future research.

We generated the HMLGPA factor which bought the high-value-high-GP/A portfolio and sold the low-value-low-GP/A portfolio. The factor generated a CAGR of 0.67% per month which was substantially higher than the CAGR generated by the HML factor which only generated a CAGR of 0.29% per month. The maximum drawdown of the factor is 47.3% which was a bit higher than that experienced by the traditional HML factor which only had a maximum drawdown of 40.89%. The factor, however, had a substantially larger current drawdown of 45.18% compared to the HML factor's current drawdown of 30.71%. This indicates that the conditional factor performs worse in periods where value strategies underperform.

The HMLGPA factor had a Sharpe ratio of 0.62 which were significantly better than the Sharpe ratio of the HML factor alone, which were only 0.40. The HMLGPA factor, however, had a higher maximum drawdown of 47.3%, significantly higher than the 40.89% experienced by the HML factor. The current drawdown of the HMLGPA factor is also substantially worse with 45.18% compared to just 30.71% for the HML factor. This is counterintuitive as higher profit margins seem like insurance against distress risk; however, these results support the mean reversion hypothesis.

We find that HMLGPA were able to generate a significant five-factor alpha which is interesting as Novy-Marx only tested the model against the Fama French three-factor model. This indicates that the RMW and CMA factors are also unable to explain the performance of the conditional sort on gross-profits-to-assets. The factor generated a highly significant alpha of 0.29% with a p-value of 0.02 and a t-statistic of 2.30, which however did not clear the t-statistic hurdle of 3 suggested by (Harvey et al., 2016) as the threshold for new anomalies. These findings validate (Novy-Marx, 2013) claim, that GP/A improves the performance of the HML factor significantly. The factor had a loading of 0.84 on the HML factor indicating the returns were highly driven by the traditional HML factor.

HMLGPA q-factor loadings

	HMLGPA
α	0.47%
	(3.08)
MKT-RF	-0.16
	(-4.40)
ME	0.36
	(7.13)
I/A	1.03
	(12.52)
ROE	-0.25
	(-4.14)
R^2	32.4%

Table 11. This table reports q-factor loadings and t-statistics for the price-to-book factor that were formed based on the six value GP/A sorted portfolios, by going long the high-value-high-GP/A portfolio and short the low-value-low-GP/A portfolio. Alpha is the estimated intercept, α. The q-factors we regressed returns on were: market return minus the risk-free rate, MKT-RF; the size factor, ME; the investment factor, I/A; and the profitability factor, ROE.

We tested if the HMLGPA factor could survive the q-factor model. When regressed on the q-factors and found the HMLGPA factor had a monthly alpha of 0.47% and a p-value of 0.002 and thus became even more significant than when regressed against the five-factor model. The factor loaded highly on the I/A factor with a factor loading of 1.03, indicating the factor has a high correlation with the investment factor from the q-factor model. This indicates that the firms captured through this conditional sort are very similar to the firms captured through the investment factor. The factor did however still not clear the threshold of a t-statistic of 3. It is therefore hard to conclude if the HMLGPA factor gives the q-factor model incremental explanatory power.

HMLGPA six-factor loadings

	HMLGPA
α	0.19%
	(1.53)
MKT-RF	0.02
	(0.60)
SMB	0.53
	(12.05)
HML	1.11
	(22.37)
BAB	-0.19
	(-4.62)
QMJ	0.62
	(8.56)
UMD	-0.04
	(-1.17)
R^2	52.3%

Table 12. This table reports six-factor loadings and t-statistics for the price-to-book factor that were formed based on the six value GP/A sorted portfolios, by going long the high-value-high-GP/A portfolio and short the low-value-low-GP/A portfolio. Alpha is the estimated intercept, α. The six-factors we regressed returns on were: market return minus the risk-free rate, MKT-RF; the size factor, SMB; the Fama French book-to-market factor, HML; the betting-against-beta factor, BAB; the quality-minus-junk factor, QMJ; and the momentum factor, UMD.

The HMLGPA factor did, however, fail to deliver a significant alpha when regressed against the six-factor model as it only generated an alpha of 0.19% with a resulting p-value of 0.13. The factor had a factor loading of 1.11 on the HML factor loading indicating the conditional sort is closely related to the HML factor. Interestingly, the conditional sort loaded negatively on the BAB factor which indicates that the factor tends to buy high beta stocks and short, low beta stocks.

In conclusion, the factor was not able to significantly improve the six-factor model, but the conditional sort was able to improve the performance of the five-factor and q-factor models at a 5% significance level, but were unable to beat the higher threshold of a t-statistic of 3 proposed by (Harvey et al., 2016). It is therefore hard to conclude if the factor should be added to the q- and five-factor models.

5.3.2 Gross-Profit-to-Enterprise-Value and GP/A

GPEV and GP/A double sorted portfolio, five-factor alpha and return characteristics

		High value	Medium value	Low value
Low profitability	α	0.03%	-0.13%	-0.05%
	α t-statistic	(0.38)	(-2.30)	(-0.61)
	CAGR	1.11%	0.97%	0.68%
	MAXDD	60.86%	52.11%	56.50%
	CURRENTDD	20.45%	19.87%	10.16%
	SR	0.53	0.48	0.30
High profitability	α	0.00%	0.00%	-0.15%
	α t-statistic	(0.06)	(-0.02)	(-2.01)
	CAGR	1.06%	0.98%	0.69%
	MAXDD	57.43%	55.72%	51.42%
	CURRENTDD	16.47%	17.49%	14.26%
	SR	0.52	0.48	0.30

Table 13. This table reports the portfolio characteristics of stocks that were first sorted on gross-profits-to-enterprise-value (GPEV) with 20^{th} and 80^{th} NYSE breakpoints, and then within each portfolio, 50% of firms were sorted into a high (low) GP/A (profitability) portfolio. We ended up with six portfolios. α is the five-factor alpha followed by its t-statistic, α t-statistic. Return characteristics are also included for compound annual growth rate, CAGR; maximum drawdown, MAXDD; current drawdown, CURRENTDD; and Sharpe ratio, SR.

The same six-portfolios was formed as before, but by first sorting the securities on gross-profits-to-enterprise-value instead of the book-to-market ratio. The CAGR of the portfolios did not improve the high-GP/A portfolios compared to the low-GP/A portfolios, except for the low-value-low-GP/A portfolio where it increased by just 0.05% points per month. Moreover, the CAGR deteriorated when going from the high-value-low-GP/A portfolio to the high-value-high-GP/A portfolio as the CAGR fell from 1.11% to 1.06%. This indicates that the conditional sort on GP/A does not have the same predictive power as found on the price-to-book ratio.

The maximum drawdowns were also relatively similar within the high and low GP/A portfolios indicating that the portfolios perform similarly in periods where value-investment strategies have negative returns. The current drawdown of the portfolios also did not differ substantially within the high and low GP/A portfolios indicating that investors who buy the high-GP/A portfolios should not expect a substantially different performance compared to the low-GP/A portfolios, as the portfolios seem highly correlated. The Sharpe ratios of the low-GP/A portfolios were also almost the same as for the high GP/A portfolios, as the Sharpe ratios barely changed between the high and low GP/A portfolios.

We still formed the factor GPGPA which buy the high-value-high-GP/A portfolio and short the low-value-low-GP/A portfolio to test if the conditional factor performed better than the unconditional GPEV factor. The factor generated a CAGR of 0.37% per month which were lower than the CAGR for the unconditional GPEV factor which was 0.49% in the same period.

The maximum drawdown of the GPGPA factor was 39.40%, relatively similar to the maximum drawdown of the HML and HMLGP factors. The conditional factor did, however, have a higher maximum drawdown compared to the GPEV factor which only had a maximum drawdown of 34.96%. This indicates that not only do the GPGPA factor have a lower CAGR compared to the unconditional GPEV factor but also performs worse in periods where value strategies have negative returns. The current drawdown of the GPGPA factor is 21.72% which is better than the current drawdown of the HML factor of 30.71%. The current drawdown is, however, lower for the unconditional GPEV factor as it only has a current drawdown of 19.76%, again indicating that the unconditional factor performs better than the conditional GPGPA factor.

The Sharpe ratio of the GPGPA factor was 0.45 which was also slightly higher than the HML factor's Sharpe ratio of just 0.40. The Sharpe ratio was however substantially lower for the conditional factor as it for the unconditional GPEV factor was 0.61. This indicates that the GPGPA factor performs better than the HML factor on a Sharpe ratio, maximum drawdown, and current drawdown basis. The conditional sort based on the gross-profit-to-asset ratio was therefore unable to improve the performance of the GPEV factor.

The GPGPA factor was also insignificant when regressed against the five-factor model with an alpha of just 0.05% and a p-value of 0.63. This was to be expected as neither of the long-only high-value-high-GP/A or low-value-low-GP/A alphas was significant individually. The factor interestingly had an insignificant loading on the HML factor and was mostly driven by the $(r_m - r_f)$, RMW and SMB factors, indicating that the factor capture a different subsection of firms compared to the HML factor.

Conclusively, were a conditional sort on the GP/A factor unable to significantly improve the performance of the GPEV factor, as it did not generate a significant alpha in relation to the five-factor model. Moreover, was the Sharpe ratio, maximum drawdown and current drawdown not better for the conditional factor compared to the unconditional GPEV factor again indicating that the conditional sort is not able to better capture the value premium.

5.3.3 EBITDA-to-EV and GP/A

EBITDA and GP/A double sorted portfolio, five-factor alpha and return characteristics

		High value	Medium value	Low value
Low profitability	α	-0.02%	-0.13%	-0.27%
	α t-statistic	(-0.32)	(-2.46)	(-3.61)
	CAGR	1.17%	0.97%	0.42%
	MAXDD	53.52%	48.82%	73.47%
	CURRENTDD	24.30%	16.56%	19.74%
	SR	0.87	0.80	0.34
High profitability	α	-0.04%	-0.04%	0.05%
	α t-statistic	(0.50)	(0.82)	(0.81)
	CAGR	1.13%	1.04%	0.72%
	MAXDD	60.01%	57.44%	62.78%
	CURRENTDD	19.25%	17.30%	17.35%
	SR	0.80	0.79	0.50

Table 14. This table reports the portfolio characteristics of stocks that were first sorted on EBITDA-to-enterprise-value (EBITDAEV) with 20th and 80th NYSE breakpoints, and then within each portfolio, 50% of firms were sorted into a high (low) GP/A (profitability) portfolio. We ended up with six portfolios. α is the five-factor alpha followed by its t-statistic, α t-statistic. Return characteristics are also included for compound annual growth rate, CAGR; maximum drawdown, MAXDD; current drawdown, CURRENTDD; and Sharpe ratio, SR.

Again six-portfolios were formed, this time with EBITDAEV as the value dimension and GP/A as the profitability dimension. Moving towards higher GP/A generally improved the CAGR but appeared to matter less for the medium value portfolio compared to the low-value category and even had a slightly negative influence on the high-value portfolio as the CAGR dropped from 1.17% to 1.13%.

The maximum drawdown of the low-value-low-GP/A was especially grim as the portfolio had a maximum drawdown of 73.47%. The high-value-high-GP/A portfolio also had a high maximum drawdown of 60.01% which was higher than the maximum drawdown of the high-value-low-GP/A portfolio, which was only 53.52%. The higher maximum drawdown indicates that the high-value-high-GP/A is slightly riskier in periods where value investment strategies underperform. The current drawdown is however slightly better for the high-value-high-GP/A portfolio as its current drawdown is only 19.25% compared to the high-value-low-GP/A portfolio's drawdown of 24%. The Sharpe ratio only improved for the low-value-portfolio when GP/A increased. This indicates that the conditional sort on gross-profits-to-assets is not able to improve the risk-return relationship for the high and medium value portfolios. It also indicates that the short-side of the factor drives any improvements of this conditional sort as the low-value-low-GP/A portfolio only had a Sharpe ratio of 0.34 compared to the low-value-high-GP/A portfolio's Sharpe ratio of 0.50

As with the HML factor, the alpha of the six-portfolios was only significant for the low-value-low-GP/A and medium-value-low-GP/A portfolios which indicates the added gain came from separating the low profitability growth firm, i.e., growth traps from the high profitability growth firms.

The EBITDAGPA factor is formed by going long the high-value-high-GP/A portfolio and short the low-value-low-GP/A portfolio. The EBITDAGPA factor had a CAGR of 0.63% per month which were substantially higher than the CAGR of the HML factor which was 0.29% and the EBITDAEV factor's CAGR of 0.49%. The maximum drawdown of the conditional factor was 40.36% which also was substantially better than the unconditional EBITDAEV factor's maximum drawdown of 56.23%, but relatively similar to the HML factor's maximum drawdown of 40.89%. This indicates that the conditional sort performs better than the unconditional EBITDAEV factor in periods where value investment strategies underperform. The current drawdown is also better for the EBITDAGPA factor as it currently has a drawdown of 29.72% compared to the unconditional EBITDAEV factor's current drawdown of 35.40%.

The EBITDAGPA factor had a Sharpe ratio of 0.71, which is also substantially higher than the Sharpe ratio of the HML factor which was only 0.40 in the same period. Moreover, were the EBITDAGPA factor's Sharpe ratio also higher than the unconditional EBITDAEV factor, which was only 0.56 in the same period, indicating that the conditional sort historically was able to improve the return per risk unit compared to the unconditional EBITDAEV factor.

While the unconditional EBITDAEV factor was unable to generate alpha against the five-factor model, the EBITDAGPA factor was able to generate a monthly alpha of 0.23% with a p-value of 0.017 and a t-statistic of 2.38, which as seen in tab. (14) were driven by the short-side again. The factor was however still not able to beat the t-statistic threshold of 3 indicating it is hard to conclude if the five-factor model should be updated with the EBITDAGPA factor.

EBITDAGPA q-factor loadings

	EBITDAGPA
α	0.19%
	(1.44)
MKT-RF	-0.07
	(-2.45)
ME	0.19
	(4.51)
I/A	0.54
	(7.79)
ROE	0.45
	(8.85)
R^2	22.4%

Table 15. This table reports q-factor loadings and t-statistics for the EBITDA-to-enterprise-value factor that were formed based on the six value GP/A sorted portfolios, by going long the high-value-high-GP/A portfolio and short the low-value-low-GP/A portfolio. α is the estimated intercept, alpha. The q-factors we regressed returns on were: market return minus the risk-free rate, MKT-RF; the size factor, ME; the investment factor, I/A; and the profitability factor, ROE.

The alpha of EBITDAGPA did not survive the q-factor as it only generated a monthly alpha of 0.19% with a resulting p-value of 0.15 when regressed against the q-factor model. The difference in the construction of the q-factor model made the q-factor model successful at explaining the returns of the EBITDAGPA factor. This was because EBITDAGPA factor loaded highly on the size, investment and return on equity factors in the q-factor model, whereas the five-factor model only captured part of the profitability, investment, and size premia that the EBITDAEV factor was able to generate.

EBITDAGPA six-factor loadings

	EBITDAGPA
α	-0.07%
	(-0.72)
MKT-RF	0.17
	(6.24)
SMB	0.37
	(10.44)
HML	0.64
	(16.19)
BAB	-0.03
	(-0.98)
QMJ	1.12
	(19.11)
UMD	-0.06
	(-2.50)
R^2	49.4%

Table 16 This table reports six-factor loadings and t-statistics for the EBITDA-to-enterprise-value factor that were formed based on the six value GP/A sorted portfolios, by going long the high-value-high-GP/A portfolio and short the low-value-low-GP/A portfolio. α is the estimated intercept, alpha. The six-factors we regressed returns on were: market return minus the risk-free rate, MKT-RF; the size factor, SMB; the Fama French book-to-market factor, HML; the betting-against-beta factor, BAB; the quality-minus-junk factor, QMJ; and the momentum factor, UMD.

Furthermore, the alpha of the factor became negative when the model was regressed against the six-factor model as it only generated an alpha of -0.07% per month which was insignificant with a p-value of 0.47. This indicates that the EBITDAGP factor does not significantly better capture the value premium compared to the factors included in the six-factor model.

Conclusively, the conditional sort on EBITDAEV was not able to generate excess returns in relations to the q- and six-factor models. The factor was however significant when regressed against the five-factor model indicating that the five-factor model is not able to explain the performance of the EBITDAGPA factor at a 5% significance level. The current drawdown, maximum drawdown, and Sharpe ratio were also better for the conditional factor compared to the unconditional EBITDAEV factor indicating that the conditional sort can improve these measures.

5.3.4 EBIT-to-EV and GP/A

EBITEV and GP/A double sorted portfolio, five-factor alpha and return characteristics

		High value	Medium value	Low value
Low profitability	α	-0.09%	-0.14%	-0.32%
	α t-statistic	(-1.39)	(-2.75)	(-3.84)
	CAGR	1.08%	0.96%	0.45%
	MAXDD	52.85%	48.87%	72.50%
	CURRENTDD	18.95%	15.96%	25.14%
	SR	0.55	0.50	0.15
High profitability	α	-0.03%	-0.02%	0.08%
	α t-statistic	(-0.37)	(-0.53)	(1.08)
	CAGR	1.15%	1.02%	0.75%
	MAXDD	54.93%	57.40%	68.79%
	CURRENTDD	18.52%	15.66%	19.21%
	SR	0.58	0.60	0.30

Table 17. This table reports the portfolio characteristics of stocks that were first sorted on EBIT-to-enterprise-value (EBITEV) with 20th and 80th NYSE breakpoints, and then within each portfolio, 50% of firms were sorted into a high (low) GP/A (profitability) portfolio. We ended up with six portfolios. α is the five-factor alpha followed by its t-statistic, α t-statistic. Return characteristics are also included for compound annual growth rate, CAGR; maximum drawdown, MAXDD; current drawdown, CURRENTDD; and Sharpe ratio, SR.

Again, we formed six portfolios, this time based on EBITEV as the value dimension and GP/A as the profitability dimension. Returns seem to increase as EBIT-to-Enterprise-value and GP/A increases. The EBITGPA factor was formed by going long the high-value-high-GP/A portfolio and short the low-value-low-GP/A portfolio. Especially the low-value-low-GP/A portfolio underperformed compared to the portfolio with high GP/A as the CAGR increased from 0.45% per month to 0.75% per month when moving from the low-value-low-GP/A portfolio to the low-value-high-GP/A portfolio.

The maximum drawdown of the low-value-low-GP/A portfolio was especially awful as it had a maximum drawdown of 72.50%, which indicates that this portfolio captures the worst growth traps that perform terribly. The Sharpe ratios also increased as both GP/A and EBIT-to-enterprise-value increased. Especially the low-value-low-GP/A portfolio had an atrocious Sharpe ratio of 0.15 which is substantially lower than the Sharpe ratio of the low-value-high-GP/A portfolio which is 0.30.

The high-value portfolios were however still not able to generate a significant alpha in relation to the five-factor model. The monthly five-factor alphas of the medium-value-low-GP/A and low-value-low-GP/A portfolios were -0.14% and -0.32% respectively and significantly negative on all standard significance levels. Again, the GP/A sort within the low-value category were good at spotting the terrible growth firms.

The EBITGPA factor historically had a CAGR of 0.58% per month which were substantially higher than the CAGR of the HML factor which only was 0.29% per month. The CAGR was also substantially higher than the unconditional EBITEV factor as it only had a CAGR of 0.40% in the period. This indicates that the conditional sort improves the performance of the EBITEV factor.

The maximum drawdown for the EBITGPA factor was 46.64%, and the conditional sort therefore substantially improved the maximum drawdown of 61.61% for the EBITEV factor but was still slightly higher than the HML's factors maximum drawdown of 40.89%. Interestingly, the EBITGPA factor has done quite well recently compared to other value factors with a current drawdown of just 1.18%, leaving the EBITEV factor behind which currently have a drawdown of 18.28%.

The EBITGPA factor regressed against the five-factor model had an alpha of 0.29% and p-value of 0.001 with a t-statistic of 2.75. The more profitable subsection thus improved the unconditional factor alpha substantially, which only delivered an alpha of 0.03. This indicates the five-factor model cannot explain the performance of the factor at a significance level of 1%. The factor did however still not beat the t-statistic threshold of 3 proposed by (Harvey et al., 2016). Moreover, the factor had a very high loading on the RMW factor of 0.95 which indicate that the RMW factor primarily drives the performance of the EBITGPA factor.

EBITGPA q-factor loadings

	EBITGPA
α	0.26%
	(1.93)
MKT-RF	-0.13
	(-4.27)
ME	0.11
	(2.48)
I/A	0.35
	(4.81)
ROE	0.51
	(9.54)
R^2	22.2%

Table 18. This table reports q-factor loadings and t-statistics for the EBIT-to-enterprise value factor that were formed based on the six value GP/A sorted portfolios, by going long the high-value-high-GP/A portfolio and short the low-value-low-GP/A portfolio. α is the estimated intercept alpha. The q-factors we regressed returns on were: market return minus the risk-free rate, MKT-RF; the size factor, ME; the investment factor, I/A; and the profitability factor, ROE.

The factor's alpha was however only significant at the 10% significance level when regressed against the q-factor model as the factor generated an alpha of just 0.26% with a p-value of 0.053. Interestingly, the loadings were not as tilted toward the profitability factor in the q-factor model as the factor only had a loading of 0.51 on the ROE factor, whereas the EBITGPA factor had a loading of 0.95 on the RMW factor in the five-factor model.

EBITGPA six-factor loadings

	EBITDAGPA
α	-0.09%
	(-0.84)
MKT-RF	0.14
	(4.95)
SMB	0.31
	(8.58)
HML	0.50
	(12.20)
BAB	-0.01
	(-0.21)
QMJ	1.27
	(21.00)
UMD	-0.11
	(-4.41)
R^2	0.50%

Table 19. This table reports six-factor loadings and t-statistics for the EBIT-to-enterprise-value factor that were formed based on the six value GP/A sorted portfolios, by going long the high-value-high-GP/A portfolio and short the low-value-low-GP/A portfolio. α is the estimated intercept. The six-factors we regressed returns on were: market return minus the risk-free rate, MKT-RF; the size factor, SMB; the Fama French book-to-market factor, HML; the betting-against-beta factor, BAB; the quality-minus-junk factor, QMJ; and the momentum factor, UMD.

The EBITGPA factor's alpha became negative when regressed against the six-factor model as it generates an alpha of -0.088% but was insignificant with a p-value of 0.40. The EBITGPA factor had a loading of 1.27 on the QMJ factor, again indicating that the profitability factors mainly drive the performance of the EBITGPA factor.

The EBITGPA factor was not able to generate an excess return in relation to the q- and six-factor models. The EBITGPA factor was however significant when regressed against the five-factor model. Moreover, was the factor also able to deliver a better CAGR, Sharpe ratio, Maximum Drawdown, and current drawdown compared EBITEV factor.

5.3.5 CFOEV and GP/A

CFOEV and GP/A double sorted portfolio, five-factor alpha and return characteristics

		High value	Medium value	Low value
Low profitability	α	-0.01%	-0.08%	-0.38%
	α t-statistic	(-0.19)	(-1.37)	(-4.41)
	CAGR	1.16%	1.00%	0.38%
	MAXDD	49.25%	49.09%	72.65%
	CURRENTDD	21.15%	16.54%	20.33%
	SR	0.61	0.53	0.10
High profitability	α	0.05%	0.00%	-0.13%
	α t-statistic	(0.62)	(0.03)	(-1.69)
	CAGR	1.18%	1.05%	0.56%
	MAXDD	56.04%	55.84%	63.02%
	CURRENTDD	18.19%	16.16%	20.10%
	SR	0.59	0.53	0.21

Table 20. This table reports the portfolio characteristics of stocks that were first sorted on cash-flow-from-operations-to-enterprise-value (CFOEV) with 20th and 80th NYSE breakpoints, and then within each portfolio, 50% of firms were sorted into a high (low) GP/A (profitability) portfolio. We ended up with six portfolios. α is the five-factor alpha followed by its t-statistic, α t-statistic. Return characteristics are also included for compound annual growth rate, CAGR; maximum drawdown, MAXDD; current drawdown, CURRENTDD; and Sharpe ratio, SR.

We formed six portfolios as above, but with CFOEV and GP/A instead. The CFOGPA factor was formed by going long the high-value-high-GP/A portfolio and short the low-value low-GP/A portfolio. As for the six HMLGPA portfolios, the six CFOGPA portfolio combinations consistently had better returns when we moved from low to high along both the value and GP/A dimensions. The returns increased substantially when GP/A increased in the low-value portfolio as the CAGR increased from 0.38% per month to 0.56%. The maximum drawdown of the low-value-low-GP/A portfolio was especially awful as the portfolio had a maximum drawdown of 72.65%. This indicates the conditional sort on gross-profit-to-assets is also able to improve the performance of the CFOEV factor.

The Sharpe ratio of the low-value-low-GP/A portfolio was particularly bad as the portfolio only generated an annual Sharpe ratio of 0.095. The low-value-high-GP/A portfolio had a relatively low Sharpe ratio of 0.21, but the short-leg of the factor was still able to capture the worst growth firms. Both low-value portfolios had significantly negative alphas when regressed against the five-factor model, but the low-value-low-GP/A portfolio was more significant with a t-statistic of -4.41 compared to the low-value-high-GP/A portfolio which had a t-statistic of -1.69, which was only significant at a significance level of 10%.

We also formed the CFOGPA factor as described above. The CFOGPA factor had a CAGR of 0.73% per month and a Sharpe ratio of 0.85 which was more than double the Sharpe ratio of the HML factor. The maximum drawdown was also lower for the CFOGPA factor with 37.10% compared to 40.89% for the HML factor, whereas the CFOEV factor had a maximum drawdown of 45.82% indicating that the conditional factor generally has lower drawdowns compared to the unconditional version of it and the HML factor. This suggests that the higher CAGR is not compensation for higher drawdowns. Interestingly, the current drawdown of the CFOGPA factor is substantially more attractive with a current drawdown of 23.30% compared to the HML factor's current drawdown of 30.71%. The current drawdown of the unconditional factor is 34.53% indicating that the conditional sort improves the current drawdown of the CFOEV factor.

The CFOGPA factor had a monthly alpha of 0.42% which was significant with a p-value of approximately zero and a t-statistic of 3.97. This is a substantial improvement compared to the corresponding single metric, as the CFOEV factor only had an alpha of 0.20%. The five-factor model is unable to explain the returns of the CFOGPA factor, and the factor also beat the threshold of a t-statistic of 3 proposed by (Harvey et al., 2016), indicating that the factor can improve the explanatory power of the five-factor model.

CFOGPA q-factor loadings

	CFGPA
α	0.41%
	(3.27)
MKT-RF	-0.10
	(-3.53)
ME	0.07
	(1.67)
I/A	0.55
	(8.16)
ROE	0.38
	(7.77)
R^2	26.0%

Table 21. This table reports q-factor loadings and t-statistics for the cash-flow-from-operations-to-enterprise-value factor that were formed based on the six value GP/A sorted portfolios, by going long the high-value-high-GP/A portfolio and short the low-value-low-GP/A portfolio. α is the estimated intercept. The q-factors we regressed returns on were: market return minus the risk-free rate, MKT-RF; the size factor, ME; the investment factor, I/A; and the profitability factor, ROE.

We tested the model against the q-factor model to investigate if the q-factor model was able to explain the factor's return. The strategy generated an alpha of 0.41% per month which had a resulting p-value of approximately 0. The factor loading on the profitability factor was substantially smaller as it only had a loading of 0.38 on the ROE factor, suggesting it is much less quality dependent in comparison to the CFOGPA factor's loading of 0.75 on the RMW factor in the five-factor model. The factor also beat the threshold of a t-statistic suggested by (Harvey et al., 2016) of 3, suggesting that the factor can improve the q-factor model.

CFOGPA six-factor loadings

	CFOGPA
α	0.10%
	(1.00)
MKT-RF	0.11
	(4.26)
SMB	0.22
	(6.07)
HML	0.53
	(13.37)
BAB	-0.03
	(-0.93)
QMJ	1.01
	(17.77)
UMD	0.00
	(0.02)
R^2	50.0%

Table 22. This table reports six-factor loadings and t-statistics for the cash-flow-from-operations-to-enterprise-value factor that were formed based on the six value GP/A sorted portfolios, by going long the high-value-high-GP/A portfolio and short the low-value-low-GP/A portfolio. α is the estimated intercept. The six-factors we regress returns on are: market return minus the risk-free rate, MKT-RF; the size factor, SMB; the Fama French book-to-market factor, HML; the betting-against-beta factor, BAB; the quality-minus-junk factor, QMJ; and the momentum factor, UMD.

The CFOGPA factor was however not able to generate a significant alpha when regressed against the six-factor model as it only generated an alpha of 0.10% with a p-value of 0.32 and a t-statistic of 1.00. It loaded significantly on both the SMB, HML, MKT-RF, and QMJ factors.

Conclusively the CFOGPA factor was unable to improve the six-factor model. The factor was, however, able to generate an impressive significant alpha when regressed against the q- and five-factor model which means these asset pricing models were unable to fully explain the performance of the factor. Moreover, was the conditional sort, able to improve the Sharpe ratio of the unconditional CFOEV factor, which indicates that the conditional sort is a better stand-alone strategy compared to the CFOEV factor.

5.3.6 Gross Profitability and Compound Excluding Revenue

CPEXRE and GP/A double sorted portfolio, five-factor alpha and return characteristics

		High value	Medium value	Low value
Low profitability	α	0.01%	-0.13%	-0.30%
	α t-statistic	(0.09)	(-2.14)	(-3.30)
	CAGR	1.20%	1.01%	0.39%
	MAXDD	53.77%	48.59%	70.41%
	CURRENTDD	21.99%	17.36%	18.05%
	SR	0.62	0.53	0.12
High profitability	α	0.30%	-0.04%	-0.03%
	α t-statistic	(3.48)	(-0.68)	(-0.38)
	CAGR	1.11%	1.04%	0.63%
	MAXDD	59.52%	56.64%	61.47%
	CURRENTDD	19.06%	17.41%	17.30%
	SR	0.54	0.52	0.25

Table 23. This table reports the portfolio characteristics of stocks that were first sorted on the compound metric excluding revenue-to-enterprise-value (CPEXRE) with 20th and 80th NYSE breakpoints, and then within each portfolio, 50% of firms were sorted into a high (low) GP/A (profitability) portfolio. We ended up with six portfolios. α is the five-factor alpha followed by its t-statistic, α t-statistic. Return characteristics are also included for compound annual growth rate, CAGR; maximum drawdown, MAXDD; current drawdown, CURRENTDD; and Sharpe ratio, SR.

We then tested if a conditional sort on the CPEXRE compound factor based on the GP/A ratio was able to improve the performance of the CPEXRE factor significantly. This compound factor was chosen as it had the best historical performance over the period.

The effect of moving along the GP/A dimension had a similar pattern to many of the previous combinations with value. The CAGR was again slightly reduced for the high-value category from 1.20% to 1.11%. The medium value exhibited a mediocre improvement while low value improved substantially from 0.39% to 0.63% when moving to the high GP/A portfolio.

Again, we see higher GP/A ratios are associated with more significant maximum drawdowns, and this is both true for the medium and high-value categories indicating that the stock of these high GP/A portfolios tend to perform worse in periods where value-investment strategies have negative returns. The opposite effect was however found in the low-value portfolios as the low-value-low-GP/A portfolio had a maximum drawdown of 70.41% compared to the maximum drawdown of 61.47% for the low value-high-GP/A-portfolio.

The Sharpe ratios also decreased in the high-value-high-GP/A portfolio compared to the low-value-low-GP/A portfolio which indicates that the conditional sort might not be better at capturing the top value firms. The conditional sort was, however, able to improve the Sharpe ratio for the low-value category factor which increased from 0.12 when moving from the low GP/A portfolio to the high GP/A portfolio as it rose to 0.25.

Interestingly, the high-value-high-GP/A portfolio was able to generate significantly positive alpha, and this was the first time we discovered a significant alpha for the long-side of any of the conditional factor sorts. The portfolio had an alpha of 0.30% with a p-value as low as 0.0005. Simultaneously, the low-value-low-GP/A portfolio generated a significant alpha of -0.30% with a p-value of 0.001. The fact that both portfolios are generated significant alphas should make the strategies excellent components in the construction of a long-short factor.

The COMPGPA factor was formed by going long the high-value-high-GP/A portfolio and short low-value-low-GP/A portfolio as before. The COMPGPA factor has a CAGR of 0.59% which was slightly higher than the CPEXRE factor which only had a CAGR of 0.53%. This indicates the conditional sort is slightly better at capturing the top value firms compared to the unconditional CPEXRE factor.

The maximum drawdown for the COMPGPA factor is 47.07% which was slightly worse than for the HML factor which only had a maximum drawdown of 40.89%. The maximum drawdown was however still relatively similar to the unconditional factor which has a maximum drawdown of 49.76% and the conditional sort, and therefore, failed to improve the unconditional factor's maximum drawdown. The current drawdown is surprisingly better for the COMPGPA factor as it only has a current drawdown of 15.94% compared to 30.71% for the HML factor. The COMPGPA factor also has a substantially lower current drawdown compared to the CPEXRE factor which has a current drawdown of 37.18%.

The factor also had a substantially higher Sharpe ratio of 0.61 compared to 0.42 for the HML factor for the period from June 1972 to December 2018. The Sharpe ratio performed similarly to the unconditional factor which had a Sharpe ratio of 0.58.

We regressed the model against the five-factor model and found that the factor only generated an alpha of 0.21% per month with a resulting p-value of 0.08 and therefore was only significant at a 10% significance level. Moreover, the factor had high loadings on the RMW and HML factors indicating that the returns for the factor are still primarily driven by factors in the five-factor model.

COMPGPA q-factor loadings

	CFGPA
α	0.24%
	(1.61)
MKT-RF	-0.14
	(-4.21)
ME	0.23
	(4.83)
I/A	0.71
	(8.91)
ROE	0.32
	(5.54)
R^2	24.9%

Table 24. This table reports q-factor loadings and t-statistics for the compound factor that excluded the revenue-to-enterprise-value factor and was formed based on the six value GP/A sorted portfolios, by going long the high-value-high-GP/A portfolio and short the low-value-low-GP/A portfolio. α is the estimated intercept. The q-factors we regressed returns on were: market return minus the risk-free rate, MKT-RF; the size factor, ME; the investment factor, I/A; and the profitability factor, ROE.

We regressed the model against the q-factor model to investigate if the factor were able to generate a significant alpha in relations to the q-factor model. The factor had an insignificant alpha of 0.24% with a p-value of 0.11 and therefore did not survive at the 10% significance level. The factor mainly loads on the I/A and ROE factor with a loading of 0.71 and 0.32 respectively.

Conclusively, the COMPGPA factor was unable to improve the performance of either the five- or q-factor model at a 5% significance level, despite both of the long- and short-portfolios generated alpha when regressed individually on the five-factor model. The conditional compound factor was also incapable of improving the Sharpe ratio substantially compared to the unconditional factor, as they almost had the same Sharpe ratio. The current drawdown of the strategy was substantially better for the conditional factor compared to the HML and CPEXRE factors.

5.3.7 Conclusion for GP/A Subsections of Single Metrics

The high profitability subsection of profitable value metrics had a strong tendency to improve returns for the medium and low-value categories. The same cannot be said for the high-value category, as most profitable value metrics experienced a deterioration of returns, Sharpe ratios, and an increased maximum drawdown. The HML high-value category also exhibited an increase in the maximum drawdown when going from the lower profitability to the higher profitability category. We speculate that this generally worse maximum drawdown can be attributed to either a stronger mean reversion effect for higher profit margins, a tendency for high-profit margin value stocks to have particularly fragile business models or a combination of the two effects.

We saw two of the factors, namely the HMLGPA and CFOGPA survive the q-factor model, with CFOGPA being highly significant with its t-statistic of 3.27 when regressed against the q-factor model, which even surpassed the threshold proposed by (Harvey et al., 2016) of a t-statistic of 3. None of the factors were, however, able to survive the six-factor model. The results are still impressive considering the BAB factor's recent criticism by (Novy-marx, 2018) for not being fully implementable.

The factor construction that used a conditional sort on GP/A substantially improved the five-factor alpha for all single metric factors except GPEV. The alpha was, however, for almost all factors, driven by the short-side, as the expected return was substantially worse for the low profitability low-value portfolio compared to the high profitability low-value portfolio. Indeed, all the low profitability low-value portfolios had significantly negative alpha, except for GP that was slightly better.

Because returns, alpha, and Sharpe ratios for the profitable value factors combined with GP/A—GPGPA, EBITDAGPA, EBITGPA, CFOGPA, COMPGPA saw substantial improvements, we cannot reject the fourth hypothesis.

5.4 Past Size Changes as an Explanation for the Value Premium

In this section, we present the results for the Fama-MacBeth regressions that were used to test hypothesis (4) and (5). Because (Gerakos & Linnainmaa, 2018) hold that the value premium associated with book-to-market primarily emanates from changes in firm size, we test whether this also is a plausible explanation for the value premium associated with profitable value metrics. The dataset is truncated as we only include firms starting from 1967 since the number of firms in the period before that met our requirements was deemed too small to form 20 portfolios.

Additionally, because of (Novy-Marx, 2013) claim that the more profitable subsection of firms drives the value premium, we also ran Fama-MacBeth regressions for all the value metrics under investigation, where we controlled for both changes in size and the profitability metric GP/A.

We followed the example of (Gerakos & Linnainmaa, 2018), and regression (1) in each table predicts the cross-section of monthly stock returns using log market equity (i.e., firm size), prior one-month return, prior one-year return skipping a month and finally the (log) value metric under investigation. In this way, we took into account changes in size for the past year which allowed us to distinguish between the size changes associated with the short-term reversal and one-year momentum.

Regression (2) added annual log changes in the market value of equity (dme_{t-s}), where the first—dme_{t-1}, is the December-to-December change over the calendar year *t-1* and so forth. (Gerakos & Linnainmaa, 2018) demonstrated that book-to-market became insignificant when they controlled for past five years of changes, but in their analysis controlled for size changes up to six years, so we decided to control for six years in regression (2). Regression (2) allowed us to answer the fourth hypothesis.

Finally, regression (3) added log gross profits to assets (log GP/A) that allowed us to investigate whether the value premium was driven by an overlap of relatively profitable firms with a sharp trend reversal. Alternatively, the value premium could consist of two subsections with little overlap. The more overlap between the firms with sharp trend reversal and firms with high profitability the smaller a change we expected to see when we would add the log

GP/A to the regression. If there is minimal overlap between the two subsections and we assume six years of past size changes and GP/A are both able to explain the value premium individually, we expected the value metric to turn significantly negative with the addition of log GP/A. In this way regression (3) aimed at answering hypothesis (5).

Our results may vary from (Gerakos & Linnainmaa, 2018) due to several differences in the data manipulation. Some changes that are due to the fact that we followed a similar methodology to (Fama & Macbeth, 1973) as they allocated stocks to 20 portfolios based on their beta, where we deviated by allocating stocks based on their value metric, as described in the methodology section. (Gerakos & Linnainmaa, 2018) supplemented their Compustat information with the Davis, Fama, and French (2000) database, secondly, we excluded several firms with negative GP/A to be able to use log values for the ratio. Third, we used data until more recently that ended in 2018, while their dataset ended in 2016.

5.4.1 Size Changes Book-to-Market

Average returns, book-to-market, changes in the market value of equity and GP/A

Independent variable	Regression (1)	(2)	(3)
		Average regression slopes (× 100)	
$log\, market\ equity$	12.3	1.87	-8.68
$r_{1,1}$	1.45	1.50	1.45
$r_{2,12}$	4.00	0.54	0.19
log *book-to-market*	2.72	1.90	2.66
dme_{t-1}		8.28	5.78
dme_{t-2}		-5.33	-8.00
dme_{t-3}		-4.39	-2.66
dme_{t-4}		4.48	2.83
dme_{t-5}		-5.92	-6.83
$log\, GP/A$			0.29
		t-values	
$log\, market\ equity$	1.18	0.13	-0.32
$r_{1,1}$	1.14	1.02	1.15
$r_{2,12}$	1.20	0.12	0.05
log *book-to-market*	0.38	0.35	0.37
dme_{t-1}		2.42	2.07
dme_{t-2}		-1.20	-1.28
dme_{t-3}		-0.96	-0.63
dme_{t-4}		0.80	0.44
dme_{t-5}		-1.30	-1.52
$log\, GP/A$			0.41
$Average\ R^2$	37.5%	59.5%	64.9%

Table 25. This table reports the average regression slopes and t-values from the Fama-MacBeth regressions predicting monthly returns to find factor premia. The regressions include the following variables: the *log market value of equity*, *log market equity*; prior one-month return, $r_{1,1}$; prior one-year return skipping a month, $r_{2,12}$; *log book-to-market*, *log book-to-market*; annual log changes in the market value of equity, dme_{t-s}; and log gross-profits-to-assets, *log GP/A*. dme_{t-1} is then the December to December change over calendar the year *t-1*. Some observations started in 1962, but all required variables were unavailable before year the 1968 as we required data on six years of past size changes.

When stocks were allocated to 20 portfolios for each period rather than analyzed on an individual firm level, we failed to see the same pattern as shown by (Gerakos & Linnainmaa, 2018). In regression (1) we got positive coefficients for all explanatory variables even though we expected to see a negative slope coefficient for *log market equity* and short-term momentum ($r_{1,1}$), i.e., short-term reversal. Both one-year momentum excluding last month ($r_{2,12}$) and *log book-to-market* had positive coefficients as anticipated but without any statistically significant t-statistics.

In regression (2) we saw that past size changes could explain some of *log book-to-market* as its coefficient dropped from 2.72 down to 1.9, we did however not observe a negative *log book-to-market* in regression (3) as *log book-to-market* increased to 2.66, while however remaining statistically insignificant. The size changes for the past six years had mixed results. We were expecting to see consistently negative coefficients as found by (Gerakos & Linnainmaa, 2018), we did, however, observe both positive and negative coefficients with the size change for year *t-1* being the only statistically significant size change in both regression (2) and (3). When we added the *log GP/A* in regression (3) which was slightly positive, we observed *log market equity* becoming negative, and this was perhaps because many small firms were unprofitable and therefore if you control for a quality metric, the size premium had a tendency to become evident. While one should be careful to draw inferences based on such statistically insignificant results, these results are consistent with the findings of (Asness, Frazzini, Israel, Moskowitz, & Pedersen, 2018).

It seems that we lost quite a bit explanatory power for the momentum and size change variables as the analysis was conducted on a portfolio level, where the size change firm characteristics aggregated over many stocks to a large extent canceled each other out.

Despite that, as we analyzed on a portfolio level, we expected to get more statistically significant results for book-to-market since stocks were allocated each year based on their book-to-price ventile. These results may be attributed to a combination of the recent poor performance of book-to-market as in fig. (1), and the book-to-market premium becoming less significant when controlling for momentum and short-term reversal.

5.4.2 Size Changes GPEV

Average returns, GPEV, changes in the market value of equity and GP/A

Independent variable	Regression (1)	(2)	(3)
	Average regression slopes (× 100)		
$log\,market\ equity$	-18.00	-20.00	-5.23
$r_{1,1}$	3.10	3.55	3.69
$r_{2,12}$	4.67	5.86	-0.38
$log\,GPEV$	24.34	25.84	13.82
dme_{t-1}		-18.45	-10.84
dme_{t-2}		8.74	8.20
dme_{t-3}		10.62	5.29
dme_{t-4}		-9.25	-0.97
dme_{t-5}		-3.07	-0.77
$log\,GP/A$			1.21
	t-values		
$log\,market\ equity$	-1.10	-1.18	-0.41
$r_{1,1}$	3.63	3.52	3.51
$r_{2,12}$	1.31	1.63	-0.12
$log\,GPEV$	3.50	3.98	2.81
dme_{t-1}		-4.52	-3.05
dme_{t-2}		2.16	1.68
dme_{t-3}		1.98	1.06
dme_{t-4}		-2.45	-0.28
dme_{t-5}		-0.96	-0.22
$log\,GP/A$			1.05
$Average\ R^2$	42.0%	61.9%	64.9%

Table 26. This table reports the average regression slopes and t-values from the Fama-MacBeth regressions predicting monthly returns to find factor premia. The regressions include the following variables: the log market value of equity, log *market equity*; prior one-month return, $r_{1,1}$; prior one-year return skipping a month, $r_{2,12}$; log gross-profit-to-enterprise-value, *log GPEV*; annual log changes in the market value of equity, dme_{t-s}; and log gross-profits-to-assets, *log GP/A*. dme_{t-1} is then the December to December change over calendar the year $t-1$. Some observations started in 1962, but all required variables were not available before the year 1968 as we required data on six years of size changes.

We found statistical significance for GPEV at both the 5% and 10% significance level in regression (1) as we expected. GPEV also had a much higher slope coefficient of 24.34 than the one found for book-to-market in tab. (25) of just 2.72. We expected to see a higher slope coefficient as the return for the single metric GPEV was also much higher than the return of the book-to-price single metric, where GPEV had a CAGR of 0.49% per month compared to the 0.34% for book-to-price. Surprisingly, we did not observe short-term reversal consistent with our expectations, but instead a slightly positive short-term momentum ($r_{1,1}$) which was significantly different from zero with a slope of 3.10 and t-statistic of 3.63.

The GPEV and short-term momentum premia remained statistically significant for regression (2) and (3). The GPEV factor premium did, however, drop to about half of its previous value when we controlled for *log GP/A* in regression (3). This makes sense as firms that score well on profitable value metrics naturally have a gravitation towards profitable firms as previously discussed. As in table. (25), we found the size change dme_{t-1} to be significant in both regression (2) and (3), while the results for the rest of the size changes were mixed. We cannot conclude that the GPEV value premium was fully explained by changes in past size based on our findings.

5.4.3 Size Changes EBITDAEV

Average returns, EBITDAEV, changes in the market value of equity and GP/A

Independent variable	Regression		
	(1)	(2)	(3)
	Average regression slopes (× 100)		
$log\, market\ equity$	-61.31	-49.84	-32.28
$r_{1,1}$	2.48	4.10	0.57
$r_{2,12}$	2.76	-1.65	-3.46
$log\, EBITDAEV$	12.36	6.15	0.20
dme_{t-1}		-12.07	-11.51
dme_{t-2}		-3.16	-6.83
dme_{t-3}		-11.66	-12.34
dme_{t-4}		5.46	9.36
dme_{t-5}		1.06	6.10
$log\, GP/A$			2.44
	t-values		
$log\, market\ equity$	-1.84	-1.36	-0.93
$r_{1,1}$	2.26	2.72	2.92
$r_{2,12}$	0.77	-0.46	-0.93
$log\, EBITDAEV$	1.64	0.78	0.03
dme_{t-1}		-2.80	-2.13
dme_{t-2}		-0.63	-1.00
dme_{t-3}		-2.39	-2.18
dme_{t-4}		1.45	2.03
dme_{t-5}		0.17	1.05
$log\, GP/A$			0.99
$Average\ R^2$	41.5%	61.3%	65.5%

Table 27. This table reports the average regression slopes and t-values from the Fama-MacBeth regressions predicting monthly returns to find factor premia. The regressions include the following variables: the log market value of equity, log *market equity*; prior one-month return, $r_{1,1}$; prior one-year

return skipping a month, $r_{2,12}$; log EBITDA-to-enterprise-value, *log EBITDAEV*; annual log changes in the market value of equity, dme_{t-s}; and log gross-profits-to-assets, *log GP/A*. dme_{t-1} is then the December to December change over the calendar year *t-1*. Some observations start in 1962, but all required variables are not available before the year 1968 as we require data on six years of size changes.

For regression (1) when we analyzed the EBITDAEV, *log market equity* became significantly negative at the 10% significance level for regression (1) with a t-statistic of -1.84, but EBITDAEV failed to become significant at both the 5% and 10% level with a t-statistic of just 1.64.

The *log EBITDAEV* coefficient slope dropped to about half when we controlled for size changes in regression (2) and got close to zero when we finally controlled for *log GP/A* as well. We found three of the size change coefficients became significant in regression (3), where dme_{t-1} and dme_{t-3} were negative, while dme_{t-4} unexpectedly was positive.

These results indicate that size changes and GP/A partly drive EBITDAEV. We expected to see the drop in the EBITDAEV coefficient slope when we controlled for *log GP/A* but were surprised to see a drop for the value premium in regression (2) when we controlled for size as we did not observe this pattern for the other value metrics analyzed above. While the findings were statistically weak, they are consistent with (Gerakos & Linnainmaa, 2018) in the sense that size changes did explain part of the value premium, but inconsistently as slope coefficients for past size changes were not consistently negative.

5.4.4 Size Changes EBITEV

Average returns, EBITEV, changes in the market value of equity and GP/A

Independent variable	Regression		
	(1)	(2)	(3)
	Average regression slopes (× 100)		
$log\,market\ equity$	-49.52	-66.01	-62.44
$r_{1,1}$	1.64	0.97	0.96
$r_{2,12}$	2.07	-0.48	-2.13
$log\ EBITEV$	11.83	12.17	14.53
dme_{t-1}		-12.72	-0.18
dme_{t-2}		7.6	6.86
dme_{t-3}		6.01	2.68
dme_{t-4}		-7.01	-7.69
dme_{t-5}		-10.43	-16.74
$log\ GP/A$			5.90
	t-values		
$log\,market\ equity$	-2.00	-2.58	-2.39
$r_{1,1}$	1.91	1.34	1.28
$r_{2,12}$	0.39	-0.09	-0.44
$log\ EBITEV$	1.59	2	1.77
dme_{t-1}		-2.65	-0.27
dme_{t-2}		2.33	2.21
dme_{t-3}		1.49	0.68
dme_{t-4}		-1.91	-2.01
dme_{t-5}		-1.71	-1.72
$log\ GP/A$			2.32
$Average\ R^2$	40.5%	60.7%	61.4%

Table 28. This table reports the average regression slopes and t-values from the Fama-MacBeth regressions predicting monthly returns to find factor premia. The regressions include the following variables: the log market value of equity, log *market equity*; prior one-month return, $r_{1,1}$; prior one-year return skipping a month, $r_{2,12}$; *log* EBIT-to-enterprise-value, *log EBITEV*; annual log changes in the market value of equity, dme_{t-s}; and log gross-profits-to-assets, *log GP/A*. dme_{t-1} is then the December to December change over the calendar year *t-1*. Some observations started in 1962, but all required variables were unavailable before the year 1968 as we required data on six years of size changes.

In regression (1) counter to expectations, we observed positive short-term momentum that was statistically significant at the 5% level. Log EBITEV had a positive slope coefficient but was insignificant.

EBITEV did not exhibit the same pattern as EBITDAEV did in tab. (27), as it increased when we controlled for past changes in size in regression (2) and then increased again once we controlled for GP/A but remained insignificant. *Log market equity* remained statistically negative across all three regressions congruent to our expectations. Again, we got mixed results for the size changes with both positive and negative statistically significant slope coefficients and therefore cannot conclude the value premium are driven by mean reversion from past size changes.

Interestingly, we found GP/A to be a significant predictor of returns in regression (3) with its positive slope coefficient and t-statistic of 2.32 as expected, but the *log EBITEV* coefficient only dropping slightly, while still being insignificant.

5.4.5 Size Changes CFOEV

Average returns, CFOEV, changes in the market value of equity and GP/A

Independent variable	Regression (1)	(2)	(3)
		Average regression slopes (× 100)	
$log\,market\,equity$	-39.09	-25.26	-17.28
$r_{1,1}$	1.78	0.31	1.51
$r_{2,12}$	1.23	4.44	6.93
$log\,CFEV$	12.2	9.46	10.91
dme_{t-1}		2.21	0.43
dme_{t-2}		-7.23	-7.77
dme_{t-3}		1.30	-0.30
dme_{t-4}		9.12	8.04
dme_{t-5}		-5.60	-4.94
$log\,GP/A$			1.16
		t-values	
$log\,market\,equity$	-1.96	-1.25	-1.04
$r_{1,1}$	1.73	0.22	1.09
$r_{2,12}$	0.35	1.33	1.80
$log\,CFEV$	2.50	2.15	2.24
dme_{t-1}		0.89	0.17
dme_{t-2}		-1.86	-1.83
dme_{t-3}		0.34	-0.08
dme_{t-4}		2.52	2.43
dme_{t-5}		-1.22	-1.19
$log\,GP/A$			0.85
$Average\ R^2$	38.5%	61.70%	64.9%

Table 29. This table reports the average regression slopes and t-values from the Fama-MacBeth regressions predicting monthly returns to find factor premia. The regressions include the following variables: the log market value of equity, *log market equity*; prior one-month return, $r_{1,1}$; prior one-year return skipping a month, $r_{2,12}$; log cash-flow-from-operations-to-enterprise-value, *log CFOEV*; annual log changes in the market value of equity, dme_{t-s}; and log gross-profits-to-assets, *log GP/A*. dme_{t-1} is then the December to December change over calendar the year $t-1$. Some observations started in

1962, but all required variables were not available before the year 1968 as we required data on six years of size changes.

In regression (1) we found a statistically significant negative *log market equity* coefficient, as it had a t-statistic of -1.96. The momentum factors were both positive, but insignificant. As anticipated, the *log CFOEV* was statistically significant with a t-statistic of 2.5.

Log CFOEV remained statistically significant for both regression (2) and (3), but we did not see it drop monotonically as expected when we controlled for size changes and later GP/A. While dme_t-4 were significant in regression (2) and (3), we did not observe a clear pattern between past size changes and returns as we expected based on the findings of (Gerakos & Linnainmaa, 2018).

5.4.6 Concluding Remarks for Fama-MacBeth Regressions

With our Fama-MacBeth regressions we did find the value premia to be consistently positive for all value metrics, and most of the size premia were positively associated with smaller firms, but all in all we failed to reject or confirm hypothesis (4) and (5) due to lack of statistical power associated with the data or our methodology. When we allocate firms to 20 portfolios rather than analyzing them on a firm level, we lose a substantial amount of predictability because as it turned out, the specific characteristics besides the value metric, which they were sorted on, tended to be very similar when aggregated across firms.

5.5 Timing Value

In this section, we present the results for our tests of hypothesis (6). We investigated if it was possible to time value factors using a similar approach to the one used in (Asness et al., 2017) where positions in the value factors were taken when the value spread exceeded the 80th percentile of the historical value spread measured on a prior rolling window of 10 years. These positions were then kept until the value spread became lower than the historical median spread, again based on a rolling window of the past ten years.

The strategy was not invested at all when the timed strategy was inactive, as the timed strategy was a self-financing strategy. To gauge whether we could successfully time a factor we did not compare returns period for the period as they would be the same when the timing strategy was active, but rather looked at differences between the Sharpe ratio of the factor that was active in the entire period and the timed factor.

This was just a rough measure as it was based on an unrealistic assumption that the returns between the timed strategy and the untimed strategy were independent across time. If the timed strategy happened to be invested in periods where value strategies did well, the outperformance may essentially have nothing to do with timing, but rather be due to mere chance.

Timed and untimed factors part one

	T-GPEV	UT-GPEV	T-EBITEV	UT-EBITEV
HWM	4.99	13.34	5.06	11.62
CUMV	4.36	10.71	3.83	9.50
CAGR	0.26%	0.43%	0.24%	0.40%
ACTIVECAGR	0.65%	0.43%	0.68%	0.40%
MAXDD	32.62%	34.96%	31.84%	61.61%
CURRENTDD	12.71%	19.76%	24.28%	18.28%
SR	0.12	0.15	0.12	0.13
t-statistic	0.77		0.35	

Table 30. This table reports part one of the return characteristics of the timed and untimed single factor strategies that started in 1963 and ended in 2018. The timed strategies use the notation T- while the untimed use the notation UT-. The return characteristics are high water mark, HWM; cumulative value, CUMV; compound annual growth rate, CAGR; compound annual growth rate when the strategy was active, ACTIVECAGR; maximum drawdown, MAXDD; current drawdown, CURRENTDD; Sharpe ratio, SR; and the t-statistic found using the HAC test for the difference in Sharpe ratios. They cover the period 1972 to 2018.

Timed and untimed factors part two

	T-EBITDAEV	UT-EBITDAEV	T-Book-to-market	UT-Book-to-market
HWM	9.00	23.34	7.67	10.27
CUMV	6.23	15.08	6.24	6.44
CAGR	0.33%	0.49%	0.33%	0.33%
ACTIVECAGR	0.91%	0.49%	0.71%	0.33%
MAXDD	45.67%	55.79%	37.38%	47.50%
CURRENTDD	30.71%	35.40%	18.65%	37.31%
SR	0.13	0.16	0.12	0.11
t-statistic	0.58		-0.30	

Table 31. This table reports part two of the return characteristics of the timed and untimed single factor strategies that started in 1963 and ended in 2018. The timed strategies use the notation T- while the untimed use the notation UT-. The return characteristics are high water mark, HWM; cumulative value, CUMV; compound annual growth rate, CAGR; compound annual growth rate when the strategy was

active, ACTIVECAGR; maximum drawdown, MAXDD; current drawdown, CURRENTDD; Sharpe ratio, SR; and the t-statistic found using the HAC test for the difference in Sharpe ratios. They cover the period 1972 to 2018.

Timed and untimed CFOEV

	T-CFOEV	UT-CFOEV
HWM	1.77	12.76
CUMV	1.43	8.36
CAGR	0.08%	0.49%
ACTIVECAGR	0.33%	0.49%
MAXDD	25.67%	45.82%
CURRENTDD	18.99%	34.53%
SR	0.06	0.16
t-statistic	1.76	

Table 32. This table reports the return characteristics of the timed and untimed cash-flow-from-operations factor that started in 1972 and ended in 2018. The timed strategies use the notation T- while the untimed use the notation UT-. The return characteristics are high water mark, HWM; cumulative value, CUMV; compound annual growth rate, CAGR; compound annual growth rate when the strategy was active, ACTIVECAGR; maximum drawdown, MAXDD; current drawdown, CURRENTDD; Sharpe ratio, SR; and the t-statistic found using the HAC test for the difference in Sharpe ratios. They cover the period 1982 to 2018.

5.5.1 Timing Book-to-Market

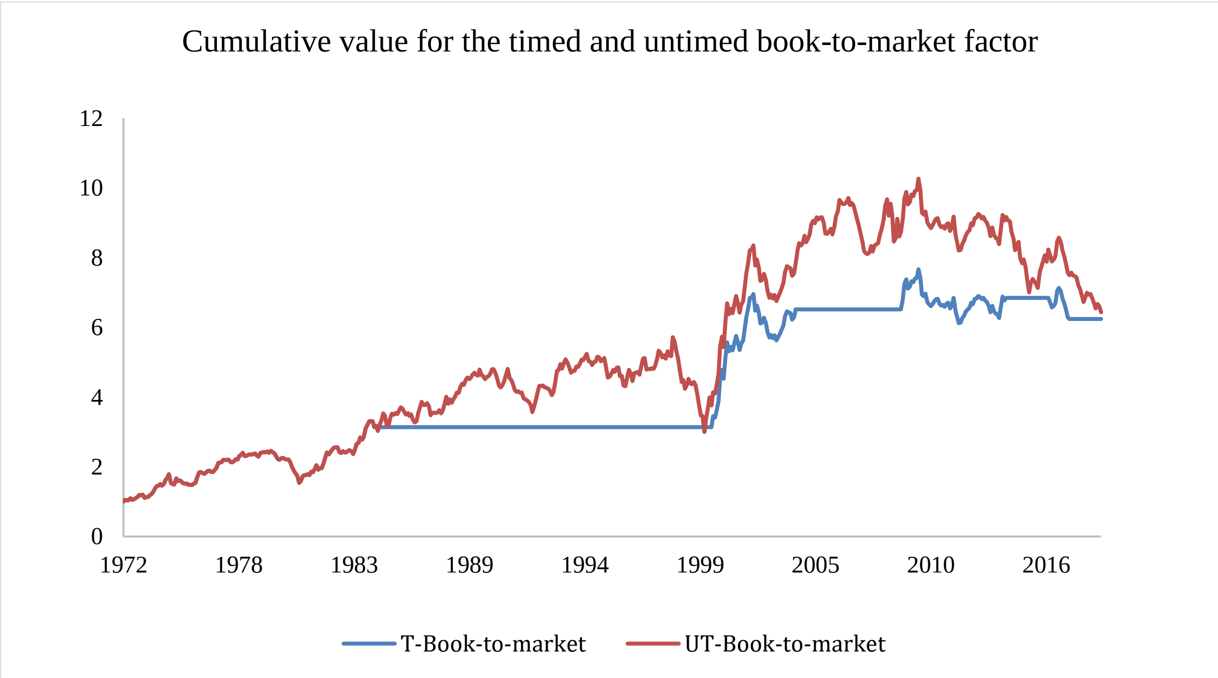

Figure 6. This figure shows the cumulative value for the timed and untimed cash-flow-from-operations-to-enterprise-value factor, book-to-market. The timed factor is T-Book-to-market and the untimed UT-Book-to-market. Yearly rebalancing was used in portfolio formation. It covers the period 1972 to 2018.

We first tested if it was possible to time the traditional price-to-book anomaly. The timed factor had a CAGR of 0.33% per month while the untimed factor has a CAGR of 0.33% per month. The CAGR of the timed factor when the strategy was active was 0.71% per month which indicates that the price-to-book factor earns higher returns in periods after a deep value episode.

The timed factor also had a smaller maximum drawdown compared to the untimed factor as the timed factor only had a maximum drawdown of 37.38% while the untimed factor had a maximum drawdown of 47.50%. The current drawdown of the factor is also better for the timed factor as it only has a current drawdown of 18.65% while the current drawdown for the untimed factor is 37.31%.

The cumulative value of the timed factor is 6.24 while for the untimed factor 6.44. The untimed factor does, however, have a high water mark of 10.27 while the timed factor only has a high water mark of 7.67.

We interestingly found that the Sharpe ratio of the timed factor was higher than for the untimed factor; this was despite the timed factor being active only 38.7% of the time. The monthly Sharpe ratio of the untimed factor was 0.11 while for the timed factor it was 0.12. Using the HAC test we found that the difference was insignificant with a p-value of 0.76. This insignificance was likely influenced by the factor’s inability to generate excess returns in 61.3% of the months. Nevertheless, we cannot conclude that the timed factor's Sharpe ratio was significantly better than the untimed factor's Sharpe ratio.

5.5.2 Timing Gross-Profits-to-Enterprise-Value (GPEV)

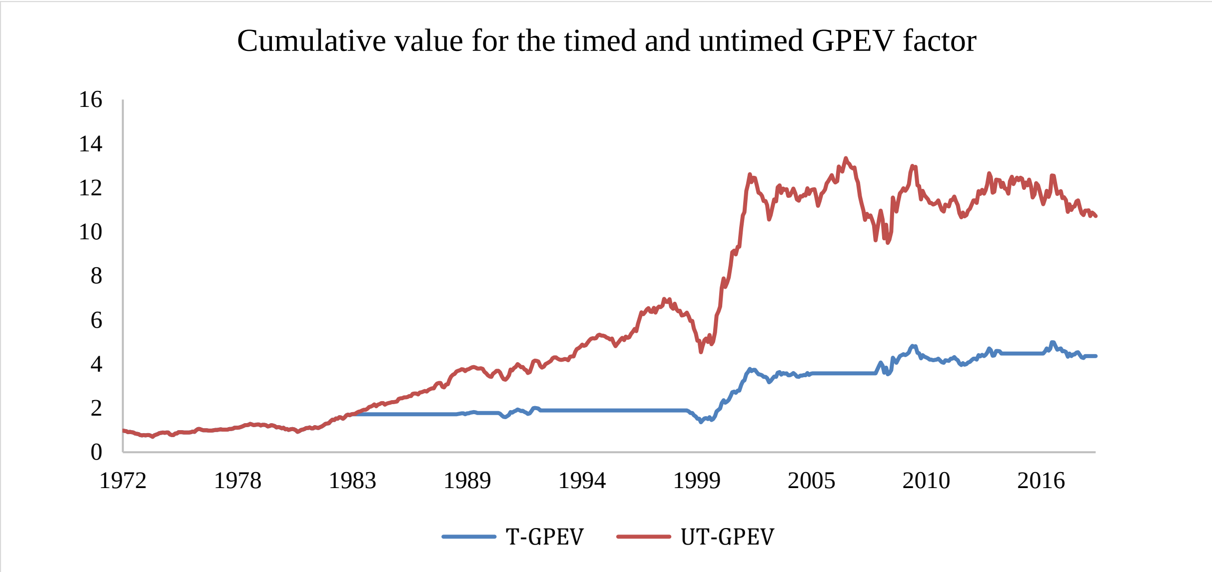

Figure 7. This figure shows the cumulative value for the timed and untimed cash-flow-from-operations-to-enterprise-value factor, GPEV. The timed factor is T-GPEV and the untimed UT-GPEV. Yearly rebalancing was used in portfolio formation. It covers the period 1972 to 2018.

We then tested if it was possible to time the GPEV factor. The CAGR of the timed factor was 0.26% per month which was substantially smaller than the CAGR of the GPEV factor which was 0.43% per month. The timed factor had a CAGR of 0.65% when it was active, indicating that the returns of the GPEV strategy are higher during deep value periods, but that the strategy is also generating positive returns in the periods where the timed factor is inactive.

The timed factor had a smaller maximum drawdown of 32.62% compared to the maximum drawdown of 34.96% associated with the untimed factor. Moreover, the strategy's current drawdown is also better for the timed factor as it currently only has a drawdown of 12.71% which is better than the current drawdown of the untimed factor which currently has a drawdown of 19.76%.

The cumulative value of the untimed factor is 10.71 by the end of 2018 whereas the timed factor only delivered a current cumulative value of 4.36. The high water mark of the two strategies was very dissimilar as the timed factor had a high water mark of 4.99 while the untimed factor had a high water mark of 13.34.

The untimed factor had a monthly Sharpe ratio of 0.15 which was higher than for the timed factor which only had a Sharpe ratio of 0.12. We then tested if the Sharpe ratio of the untimed factor was significantly higher than the timed factor using the HAC test. The difference in the two Sharpe ratios was however insignificant with a p-value of 0.44, and this indicates that the untimed factor's Sharpe ratio is not significantly better than the time factor which is interesting as the factor was only invested in the timed factor 30.1% of the time.

5.5.3 Timing EBITDA-to-Enterprise-Value (EBITDAEV)

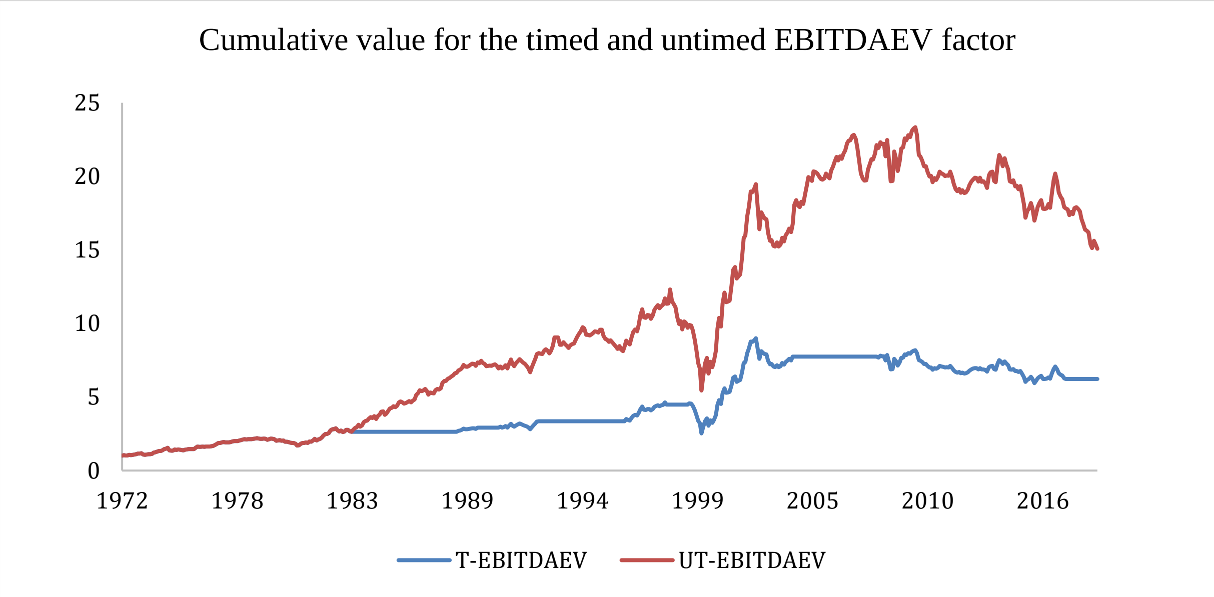

Figure 8. This figure shows the cumulative value for the timed and untimed cash-flow-from-operations-to-enterprise-value factor, EBITDAEV. The timed factor is T-EBITDAEV and the untimed UT-EBITDAEV. Yearly rebalancing was used in portfolio formation. It covers the period 1972 to 2018.

We then tested if it was possible to time the EBITDAEV factor. The untimed EBITDAEV factor had a CAGR of 0.49% per month which was relatively dissimilar to the realized CAGR of 0.33% per month for the timed EBITDAEV factor. The timed factor had a CAGR of 0.91% when isolating the periods where the strategy was active, and this is an indication that the EBITDAEV factor can generate substantially higher returns after deep value events. The timed strategy did however still perform worse over the period as the untimed factor still generated positive returns in the period where the strategy was inactive.

The maximum drawdown of the timed strategy was 45.67%, whereas the untimed strategy had a maximum drawdown of 55.79%. These results indicate that the investors who invest in the timed factor do not experience drawdowns as harsh as investors who invest in the untimed factor. The timed factor currently has a drawdown of 30.71 % which is slightly better than the current drawdown of the untimed factor which is 35.40%. The current cumulative value of the two strategies is highest for the untimed factor as it has a cumulative value of 15.08 compared to the cumulative value of 6.23 for the timed factor.

The high water mark of the untimed EBITDAEV strategy was also better for the untimed factor as it had a high water mark of 23.34 while the timed factor only had a high water mark of 9.00. We also found that the untimed factor had a higher Sharpe ratio compared to the timed factor. The Sharpe ratio for the timed factor is 0.13, and 0.16 for the untimed factor. The difference in the two Sharpe ratios was however insignificant with a p-value of 0.55, and it can therefore not be concluded that the traditional untimed factor had a higher monthly Sharpe ratio compared to the untimed strategy.

5.5.4 Timing EBIT-to-Enterprise-Value (EBITEV)

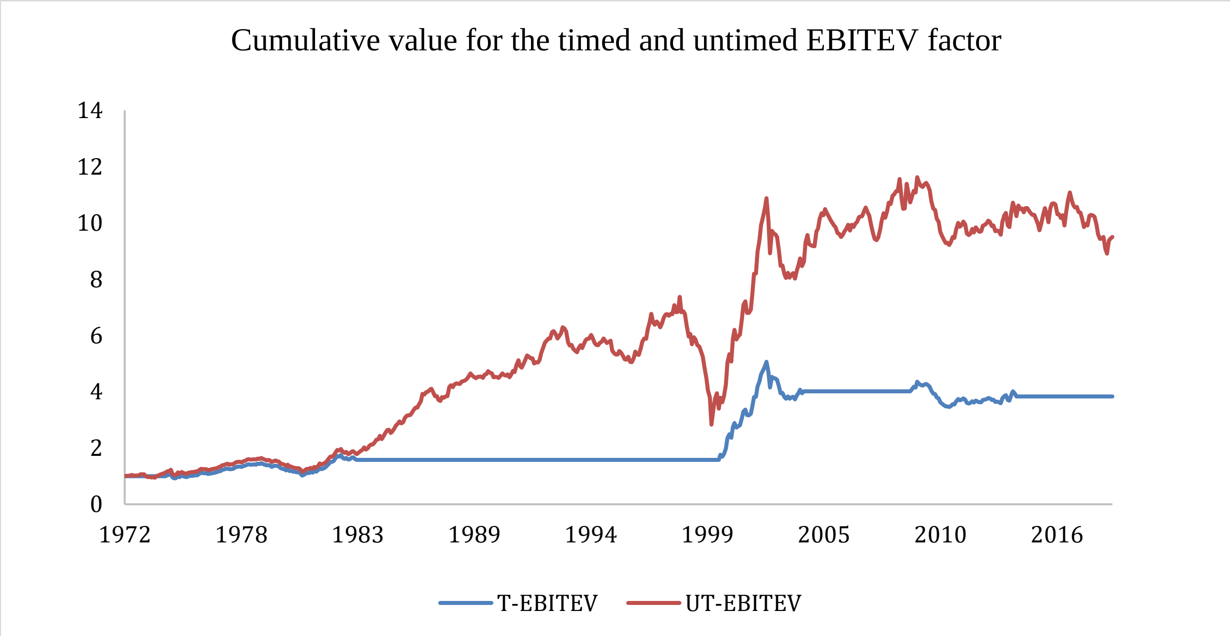

Figure 9. This figure shows the cumulative value for the timed and untimed cash-flow-from-operations-to-enterprise-value factor, EBITEV. The timed factor is T-EBITEV and the untimed UT-EBITEV. Yearly rebalancing was used in portfolio formation. It covers the period 1972 to 2018.

We then tested if it was possible to time the EBITEV factor. The untimed factor had a CAGR of 0.40% per month which was higher than the CAGR of the timed factor which was 0.24% per month for the period from June 1972 to the of December 2018. The timed strategy had a CAGR of 0.68% when the strategy was active indicating that the strategy earns higher returns after a deep value period.

The timed factor had a maximum drawdown of 31.84% which was substantially lower than the 61.61% maximum drawdown experienced by the untimed factor. This indicates that the timed factor is better at avoiding the large drawdowns which might make it a more attractive strategy compared to the untimed factor for risk-averse investors. The current drawdown of the timed-factor is, however, worse than for the untimed factor as the timed factor have a current drawdown of 24.28% whereas the untimed factor only has a current drawdown of 18.28%.

The current cumulative value is also lower for the timed factor as it currently has a value of 3.83 while the untimed factor currently has a cumulative value of 9.50. The high water mark was also worse for the timed strategy as it had a high water mark of 5.06 while the untimed factor had a high water mark of 11.62.

The timed factor also had a lower monthly Sharpe ratio of 0.12 compared to 0.13 for the factor that was active throughout the entire period from June 1972 to December 2018. The difference between the two Sharpe ratios was however insignificant when testing with the HAC test as it gave a p-value of 0.72. It can therefore not be concluded that the Sharpe ratio of the untimed factor was significantly better than the timed strategy. While the timed factor did not manage to significantly outperform the untimed factor measured on a Sharpe ratio basis, the insignificance was still interesting since the timed strategy still generated a higher CAGR in periods where the factor was active, and the factor was therefore heavily punished for not being invested 63.44% of the time.

5.5.5 Timing Cash-Flow-From-Operations-to-Enterprise-Value (CFOEV)

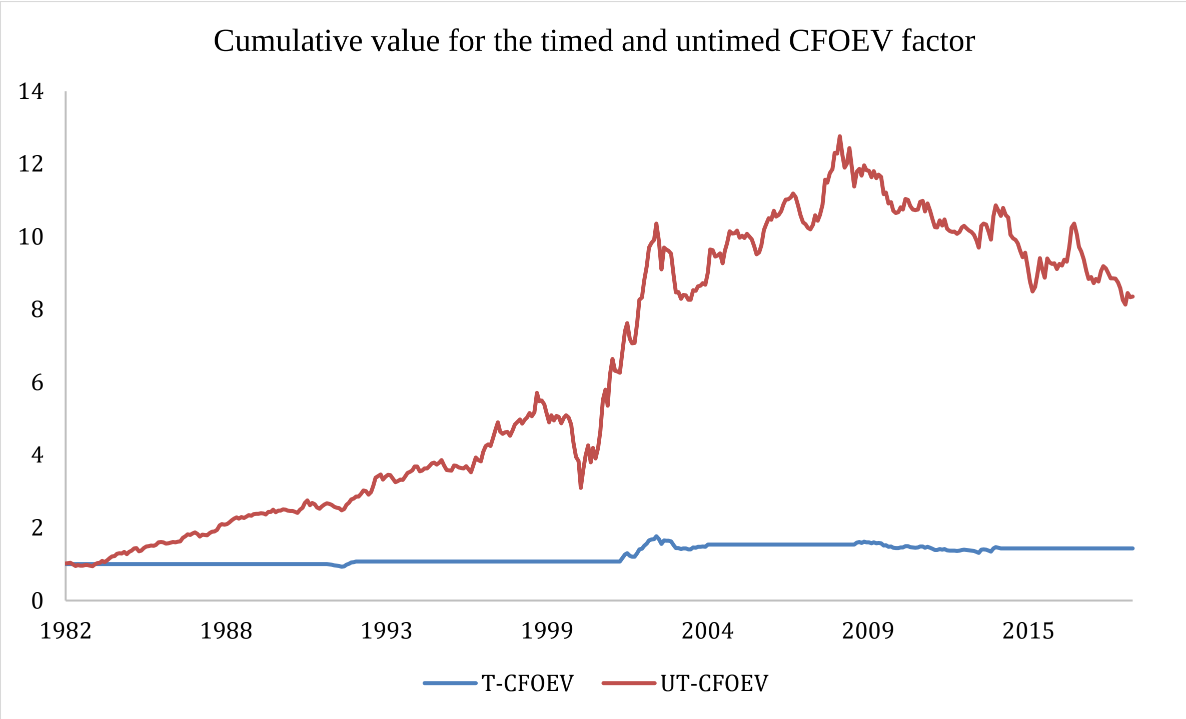

Figure 10. This figure shows the cumulative value for the timed and untimed cash-flow-from-operations-to-enterprise-value factor, CFOEV. The timed factor is T-CFOEV and the untimed UT-CFOEV. Yearly rebalancing was used in portfolio formation. It covers the period 1982 to 2018.

The timed CFOEV has a CAGR of 0.08% while the traditional cash flow factor had a CAGR of 0.49% in the same period. The timed factor, therefore, underperformed the untimed factor substantially throughout the period from June 1982 to December 2018 on a raw return basis. The CAGR of timed strategy when the strategy was active was 0.33% per month which indicates that the CFOEV performs worse in periods when the timed factor is active compared to the untimed factor.

The strategy's maximum drawdown was, however, better for the timed factor as it only had a drawdown of 25.67% while the untimed factor had a drawdown of 45.82%. The current drawdown for the timed factor is also better than for the untimed factor as the timed factor has a current drawdown of 18.99% while the untimed factor currently has a drawdown of 34.53%. It is not surprising that the drawdown for the active strategy is so much lower as the strategy is only active 24.7% of the time.

On a cumulative value basis, the timed factor underperformed terribly as it has a cumulative value of just 1.43 compared to the untimed factor's cumulative value of 8.36. The high water mark of the timed strategy was also substantially lower with only 1.77 compared to 12.76 for the untimed strategy.

The Sharpe ratio of the timed strategy of 0.06 per month was also lower than the Sharpe ratio of the untimed factor which was 0.16. Using the HAC test to test the difference in the Sharpe ratios we got a p-value of 0.08, which indicates that the Sharpe ratio of the untimed factor is significantly better than the timed factor at a 10% significance level. This suggests that the untimed factor is better at capturing the CFOEV factor premium. The massive outperformance of the untimed factor was due to the timed factor only being invested very shortly, while the strategy also, return-wise, did worse on average when active compared to the rest of the period.

We replicated the study using a rolling estimation window of 5 and 15 years, and similar results were found. All the timed strategies underperformed the untimed strategies which indicate that the factors are generating positive returns in periods where the strategy is inactive.

5.5.6 Conclusion on the Timed Value Factors

Timing value factors for US equities based on the market wide value spread and entry and exit strategy similar to (Asness et al., 2017) failed to improve the performance of all value factors we tested significantly.

Not only were the timed value strategies unable to deliver better cumulative returns over the period they also all underperformed on a Sharpe ratio basis, except for timed book-to-market which did have a slightly higher Sharpe ratio, but insignificantly different when compared to the Sharpe ratio of the untimed factor using the HAC test. The timed CFOEV severely underperformed the untimed factor with a p-value of approximately 0.08.

These findings indicate that value strategies can generate positive average returns even in periods with a relatively narrow value spread. For US equities we, therefore, do not find evidence that subsection of periods with large value spreads, where value stocks collectively become underpriced, drive the entire value premium.

As one would expect, the maximum drawdowns were lower for the timed factors which indicate that they avoid the worst losing periods by exiting before the strategies have bottomed out.

Notably, the timed CFOEV factor underperformed its untimed counterpart, as it were only able to generate a CAGR of 0.08% whereas the untimed factor generated a CAGR of 0.49%. The return of the timed CFOEV factor was so bad it also underperformed in the periods where the strategy was active. All the rest of the timed value strategies, however, had better average CAGR returns in the active period compared to the untimed factor CAGR for the whole period. Because timed strategies underperformed compared to the whole period, but often had higher average returns for the active periods compared to the inactive periods, the timed strategy as a standalone strategy did not look promising but combining timed strategies with other strategies such as momentum, to avoid long periods of inactivity may hold potential.

6 Conclusion

We showed some of the differences and similarities between the HML factor and alternative profitable value factors, and while most of them were explained by exposures to traditional factors we did find profitable value metrics performed substantially better as stand-alone strategies when compared to the traditional HML factor on both a CAGR, Sharpe ratio, and maximum drawdown basis. The CFOEV was an exceptionally strong anomaly that survived the q- and five-factor model on a stand-alone basis and further improved the performance when conditioned on GP/A with a q-factor alpha of 0.41% and a highly significant t-statistic of 3.97, easily passing (Harvey et al., 2016)'s suggested threshold for new anomalies of 3. We, therefore, suggest the addition of the conditional cash flow from the operations-based factor to the q-factor model and five-factor models.

(Novy-Marx, 2013) claim that the subsection of HML factor with high gross profits to assets almost ubiquitously improves its performance. However, when the long-leg is analyzed individually, we find it has a substantial maximum drawdown of 68.9% suggesting the improved performance of the HMLGPA over the HML factor may be maximum drawdown compensation. A plausible explanation may be that the long-leg of HMLGPA tend to contain many firms with high, but severely fragile profit margins that tend to get walloped in economic downturns and therefore are often subject to substantial mean reversion in these periods.

We also found that the subsection of profitable value metrics with high gross profits to assets ratios for most value metrics experienced a lowered return and increased maximum drawdown, and this is counterintuitive as higher profit margins seem like insurance against distress risks. Again, this effect can reasonably be explained by these relatively profitable value firms tend to fall victim to a strong profit mean reversion effect.

The alpha we found for single metrics and single metric GP/A combinations were consistently driven by the short-leg of the factor's consistent with (Stambaugh et al., 2012, 2014, 2015) that argue that the short-leg drives most anomalies due to arbitrage risk and arbitrage asymmetry. A likely explanation is that the expensive and unprofitable subsection often contains a large number of growth stocks held by the least sophisticated retail investors that may sustain the high price level for an extended period or bid them up even higher leading to considerable shorting risks.

Our compound factors comprised each period of the firms with the best overall score on single value factors failed to subsume all the single metrics on their own but did have a tendency to do better than the median of single factors, suggesting it may be a less risky approach to capture value premia compared to investing in the individual factors. The compound factors did, however, fail to deliver alpha when tested against the q-factor model and investors, therefore, may be better off with a q-factor mimicking portfolio.

With our Fama-MacBeth regression results, we were neither able to support or reject (Gerakos & Linnainmaa, 2018)'s findings, that past changes in size can explain the entire HML premium as too much information was lost when firm characteristics for size changes were aggregated across ventiles. It is therefore still a plausible explanation that value premia are driven by a subsection of mean reverting firms that have had negative size changes in the past.

Timing value factors for US equities based on the market-wide value spread measured using a rolling window, did not prove to be an effective standalone investment strategy. All the timed strategies underperformed the untimed strategies on a CAGR basis, and only the timed book-to-market factor beat the untimed equivalent on a Sharpe ratio basis but was statistically insignificant. The lousy performance is mainly due to the factor being a self-financing strategy and therefore naturally experiences long periods without being invested in any asset when the timing signal was inactive.

Our findings suggest that value strategies can generate positive average returns even in periods with a narrow value spread and therefore imply that the value premium cannot be attributed to subperiods with considerable value spreads alone. The CAGR of the factors was better for them in the active period when compared to the inactive period, except for CFOEV factor, which suggests that the returns are on average higher in periods following a deep value event.

While being difficult to choose a reasonable estimation window for measuring the value spread without engaging in data-mining, combining timed strategies with other strategies such as momentum to avoid long periods of inactivity could hold potential for investors.

7 Future Research

(Gray & Vogel, 2012) investigated whether long-term valuation metrics, meaning the average of a value metric over a couple of years prior could enhance the return predictability of the metric but found little evidence that for its efficacy. (Piotroski & So, 2012) find that a subsection of book-to-market value strategies performs better for a subsection of value firms with good fundamental momentum based on Piotroski's FSCORE and (Huang et al., 2019) find statistical significant alpha for fundamental momentum strategies based on various profitability measures when tested against both the five-factor model and the q-factor model. Because a long-term valuation strategy does not distinguish between firms that are trending downwards and upwards, it is likely to hold more firms that are trending downwards as their average value measure include fundamentals from a few years back is higher. Based on this notion, we suggest investigating whether long-term valuation metrics in combination with fundamental momentum can enhance value strategies. Such research may discover a subsection of firms that drive a significant proportion of value strategies.

(Asness et al., 2017) find that a strategy that allocates to subsections of industries, when the value spread within the industry is relatively high, can improve the performance for a book-to-market strategy. It could be interesting to investigate the performance of timed value strategies for other metrics in industry subsection or within other geographical subsections. It could also be interesting to investigate whether these timed strategies could be improved if one allocates to other self-financing strategies, such as momentum, when the timed value strategy is inactive.

Combining value strategies with a measure for clean balance sheet may hold potential in distinguishing firms that are too cheap for a reason from firms that are mispriced. Future research could try to develop such a measure. It may take into account if firms seem to move balance sheet items back and forth to appeal to shareholders and could include items such as continuously large extraordinary item reporting.

Another issue could be to investigate whether the value premia often are driven by the short-side because of shorting constraints as has been shown to often be the case by (Stambaugh et al., 2012, 2014, 2015) or whether it is because of a predictable behavioral bias.

It could also be interesting to investigate the performance of profitable value metrics in other markets, so see if the performance resembles the findings on the US market.

For future research, we also suggest analyzing whether size changes and GP/A drive the value premia of the profitable value metrics by running Fama-MacBeth regressions on individual firm level rather than forming portfolios. This approach could perhaps provide more insightful results revealing whether size changes and relatively profitable subsections entirely drive profitable value metrics premia.

8 Limitations

8.1 Data Issues

Despite the high quality of data on CRSP and Compustat, they are still subject to erroneous data. Examples of this include the tendency of machine-readable data to be mistreated due to stock splits or the occasional mispricings of digits (Rosenberg & Houglet, 1974). These errors can often become outliers and therefore also have a proportionally higher impact on the nature of the data, researchers should, therefore, be cautious by trying to reduce the dependence on outliers (Rosenberg & Houglet, 1974). While testing whether annual changes in size explain the value premium, we winsorize the data for explanatory variables that are not return based. However, in most of our study, we have chosen to follow most of the standard procedures of data modification and factor construction, that do not specifically address this concern to retain the direct comparability of our results to the well-known factor models. The databases are under continuous improvement and through the years many errors have been corrected (Rosenberg & Houglet, 1974), but significant influence of undetected erroneous data is nevertheless plausible.

The study is limited by the fact that a large number of firms have missing fundamentals in the periods, and this limited our ability to generalize the findings as the firms with missing data might not have had the same anomalies as companies with nonmissing fundamental data. This issue of missing data was also a problem for the study, as there were only a few firms with fundamental data in the initial period from 1963 which again might bias the study. Moreover, cash flow from operations was also not available before 1972 which limited our ability to conclude whether this variable was able to capture the value-premium better.

There was also a problem with the delisting returns as multiple securities had missing delisting returns. This issue was solved using the delisting returns proposed by (Shumway, 1997). This way of solving the missing delisting returns might not be optimal as the average delisting returns could have changed since Shumway published his paper in 1997. Moreover, the approach might also not be the best solution as delisting returns could be correlated with certain firm characteristics.

The study was also limited by our inability to adjust for structural changes, and this is important as quantitative factor investment strategies have become popular over the last decade. This indicates that these historically profitable factor investment strategies might have become flooded and thereby have made the market more efficient which eventually could eliminate many asset pricing anomalies. Investors should therefore not expect the same performance of the factor investments strategies as they have had historically. It is however likely risk-driven anomalies persist.

Since we formed our factors in June and held them for a year, we did not account for seasonality effects that could bias the results. This must be accounted for by practitioners that rely on our results.

8.2 Implementation Issues for Practitioners

The factors presented above are based on historical data, which can be a dubious indicator of future returns due to the extensive data-snooping by academics. There is, therefore, no certainty about whether investors can expect similar returns in the future. This is especially relevant as multiple of the value factors have had negative returns over the last 10 years. On the other hand, the factors have still generated abnormal returns over a long period, and it is, therefore, hard to conclude if the value investment strategies have stopped working.

Multiple factors formed in this study were able to generate alpha in relation to the q- and five-factor models. However, this alpha often came from the short-side of the factor, which is an important notion as investors often face shorting constraints, especially among securities with lower market capitalizations. Secondly, investors are also often subject to short sale cost. These might alter the factor strategies attractiveness as the profit from the factor strategies could be wiped out by the short sale costs. Investors would also have a hard time implementing the factors as the factors are formed based on the assumption that investors face no margin requirement which is unrealistic from a practical perspective. Brokers often demand a margin of certainty to avoid unwanted risk exposure (Pedersen, 2015). Investors should therefore not expect future returns equivalent to the historical backtests, as the factor strategies in practice face short sale cost, shorting constraint and margin requirements that will impact their returns negatively.

9 Appendix

CFOEV five-factor model loadings and robustness test

	CFOEV(10%)	CFOEV(30%)
α	0.26%	0.15%
	(2.66)	(2.63)
MKT-RF	-0.11	-0.04
	(-4.66)	(-2.72)
SMB	-0.13	-0.06
	(-3.72)	(-3.28)
HML	0.79	0.54
	(17.39)	(20.98)
RMW	0.82	0.54
	(17.75)	(20.41)
CMA	0.14	0.23
	(1.99)	(5.80)
R^2	72.8%	79.9%

Table 33. This table reports the five-factor robustness tests for the equal-weighted cash-flow-from-operations-to-enterprise-value factor, CFOEV. It includes the Fama French five-factor alpha, α; the five-factor loadings for the period 1972 to 2018, for which yearly rebalancing was used. The CFOEV (10%) were formed using 10^{th} and 90^{th} NYSE value percentile breakpoints, while the CFOEV (30%) were formed using 30^{th} and 70^{th} NYSE value percentile breakpoints. The Fama French five factors we regress returns on were: market return minus the risk-free rate, MKT-RF; the size factor, SMB; FF-HML and HML are the same; the profitability factor, RMW; and investment factor, CMA.

CFOEV q-factor loadings and robustness test

	CFOEV(10%)	CFOEV(30%)
α	0.29%	0.20%
	(1.95)	(2.15)
MKT-RF	-0.17	-0.08
	(-5.04)	(-3.87)
ME	-0.20	-0.13
	(-4.17)	(-4.33)
I/A	1.04	0.87
	(13.26)	(17.65)
ROE	0.35	0.15
	(6.24)	(4.32)
R^2	43.5%	49.4%

Table 34. This table reports the q-factor robustness tests for the equal-weighted cash-flow-from-operations-to-enterprise-value factor, CFOEV. It includes the Fama French five-factor alpha, α; the five-factor loadings for the period 1972 to 2018, for which yearly rebalancing was used. The CFOEV (10%) were formed using 10^{th} and 90^{th} NYSE value percentile breakpoints, while the CFOEV (30%) were formed using 30^{th} and 70^{th} NYSE value percentile breakpoints. The q-factors we regress returns on were: market return minus the risk-free rate, MKT-RF; the size factor, ME; the investment factor, I/A; and the profitability factor, ROE.

9.1.1 References

Aharoni, G., Grundy, B., & Zeng, Q. (2013). Stock returns and the Miller Modigliani valuation formula: Revisiting the Fama French analysis. *Journal of Financial Economics*, *110*(2), 347–357.

AQR. (2019). Data Sets. Retrieved March 4, 2019, from https://www.aqr.com/Insights/Datasets

Asness, C., Frazzini, A., Israel, R., Moskowitz, T. J., & Pedersen, L. H. (2018). Size matters, if you control your junk. *Journal of Financial Economics*, *129*, 479–509. https://doi.org/10.1016/j.jfineco.2018.05.006

Asness, C. S. (1994). *Variables that explain stock returns*.

Asness, C. S., Frazzini, A., & Pedersen, L. H. (2018). Quality minus junk. *Review of Accounting Studies*, *24*(November 2018), 34–112.

Asness, C. S., Liew, J. M., Pedersen, L. H., & Thapar, A. K. (2017). Deep Value. *Ssrn*, (December 2017). https://doi.org/10.2139/ssrn.3076181

Asness, C. S., Moskowitz, T. J., & Pedersen, L. H. (2013). Value and Momentum Everywhere. *Journal of Finance*, *68*(3), 929–985. https://doi.org/10.1111/jofi.12021

Ball, R., & Watts, R. (1977). Reply to Salamon and Smith. *The Journal of Finance*, *32*(5), 1802–1808.

Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics*, *9*, 3–18. https://doi.org/10.1016/0304-405X(81)90018-0

Bhandari, L. C. (1988). Debt/Equity Ratio and Expected Common Stock Returns: Empirical Evidence. *The Journal of Finance*, *43*(2), 507–5029.

Campbell, J. Y., & Shiller, R. J. (1998). Valuation ratios and long run stock market outlook.pdf. *The Journal of Portfolio Management*, *24*(2), 11–26.

Carhart, M. M. (1997). On Persistence in Mutual Fund Performance. *The Journal of Finance*, *52*(1), 57–82.

Carlson, M., Fisher, A., & Giammarino, R. (2004). Corporate Investment and Asset Price Dynamics : Implications for the Cross-section of Returns. *The Journal of Finance*, *59*(6), 2577–2603.

Chopra, N., Lakonishok, J., & Ritter, J. R. (1992). Measuring abnormal performance: do stock overreact? *Journal of Financial Economics*.

Cochrane, J. H. (2011). Presidential Address : Discount Rates with Slides. *Journal of Finance*, *66*(4), 1047–1108. https://doi.org/10.1111/j.1540-6261.2011.01671.x

De Bondt, W. F., & Thaler, R. H. (1987). Further Evidence on Investor Overreaction and Stock Market Seasonality. *The Journal of Finance*, *42*(3).

Fama, E., & French, K. (2006). The Value Premium and the CAPM. *Journal of Finance*, *61*(5), 2163–2185.

Fama, E F, & French, K. R. (2008). Dissecting Anomalies. *The Journal of Finance*, *63*(4), 1653–

1678.

Fama, Eugene F., & French, K. R. (1992). The Cross-Section of Expected Stock Returns. *The Journal of Finance*, *47*(2), 427–465. https://doi.org/doi:10.1111/j.1540-6261.1992.tb04398.x

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

Fama, Eugene F., & French, K. R. (2006). Profitability, investment and average returns. *Journal of Financial Economics*, *82*, 491–518. https://doi.org/10.1016/j.jfineco.2005.09.009

Fama, Eugene F., & French, K. R. (2011). Q&A: Why Use Book Value to Sort Stocks? Retrieved March 28, 2019, from https://famafrench.dimensional.com/questions-answers/qa-why-use-book-value-to-sort-stocks.aspx

Fama, Eugene F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, *116*, 1–22. https://doi.org/10.1016/j.jfineco.2014.10.010

Fama, Eugene F., & French, K. R. (2016). Dissecting Anomalies with a Five-Factor Model. *Review of Financial Studies*, *29*(1), 69–103. https://doi.org/10.1093/rfs/hhv043

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

Favilukis, J., & Lin, X. (2016). Wage Rigidity: A Quantitative Solution to Several Asset Pricing Puzzles. *The Review of Financial Studies*, *29*(1), 148–192. https://doi.org/10.1093/rfs/hhv041

Frazzini, A., Kabiller, D., & Pedersen, L. H. (2013). *Buffett's Alpha*.

Frazzini, A., & Pedersen, L. H. (2014). Betting against beta. *Journal of Financial Economics*, *111*, 1–25. https://doi.org/10.1016/j.jfineco.2013.10.005

French, K. R. (2019). Current Research Returns. Retrieved March 2, 2019, from http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

Gerakos, J., & Linnainmaa, J. T. (2018). Decomposing value. *Review of Financial Studies*, *31*(5), 1825–1854. https://doi.org/10.1093/rfs/hhx118

Gray, W. R., & Vogel, J. (2012). Analyzing Valuation Measures: A Performance Horse-Race Over the Past 40 Years. *Ssrn*. https://doi.org/10.2139/ssrn.1970693

Greenblatt, J. (2010). *The Little Book That Still Beats the Market*. Hoboken: John Wiley & Sons, Inc.

Harvey, C. R., Liu, Y., & Zhu, H. (2016). ⋯ and the Cross-Section of Expected Returns. *Review of Financial Studies*, *29*(1), 5–68. https://doi.org/10.1093/rfs/hhv059

Hou, K., Mo, H., Xue, C., & Zhang, L. (2018). Which factors? *NBER Working Paper Series, No. 20682*(November 2014, Revised July 2018), 1–50.

Hou, K., Xue, C., & Zhang, L. (2015). Digesting anomalies: An investment approach. *The Review of Financial Studies*, *28*(3), 650–705. https://doi.org/10.1093/rfs/hhu068

Huang, D., Zhang, H., Zhou, G., & Zhu, Y. (2019). *Twin Momentum: Fundamental Trends Matter**.

Jegadeesh, N., & Titman, S. (1993). Returns to Buying Winners and Selling Losers: Implications for

Stock Market Efficiency. *The Journal of Finance*, *48*(1), 65–91.

Karolyi, G. A. (2015). The Cross-Section of Expected Returns: Where We Stand Today. *Oxford University Press*, 2–4. https://doi.org/10.1093/rfs/hhv124

Kogan, L., & Papanikolaou, D. (2013). Firm characteristics and stock returns: The role of investment-specific shocks. *Oxford University Press*, *26*(11), 2718–2759. https://doi.org/10.1093/rfs/hht026

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

Ledoit, O., & Wolf, M. (2008). Robust performance hypothesis testing with the Sharpe ratio. *Journal of Empirical Finance*, *15*, 850–859. https://doi.org/10.1016/j.jemp

Lettau, M., & Wachter, J. A. (2007). Why is long-horizon equity less risky? A duration-based explanation of the value premium. *Journal of Finance*, *62*(1), 55–62.

Linnainmaa, J. T., & Roberts, M. R. (2016). The History of the cross section of stock returns. *NBER Working Paper Series*, *No. 22894*(December 2016), 1–72.

Lo, A. W., & MacKinlay, A. C. (1990). When Are Contrarian Profits Due To Stock Market Overreaction? *Review of Financial Studies*, *3*(2), 175–205. https://doi.org/10.1093/rfs/3.2.175

Memmel, C. (2003). Performance Hypothesis Testing with the Sharpe Ratio. *Finance Letters*, (1), 21–23.

Modigliani, F., & Miller, H. M. (1958). The Cost of Capital, Corporation Finance and the Theory of Investment. *The American Economic Review*, *48*(3), 261–297. https://doi.org/10.1136/bmj.2.3594.952

Modigliani, F., & Miller, M. H. (1963). Corporate Income Taxes and the Cost of Capital : A Correction. *The American Economic*, *53*(3), 433–443.

Novy-marx, R. (2018). *Betting Against Betting Against Beta**.

Novy-Marx, R. (2013). The other side of value: The gross profitability premium. *Journal of Financial Economics*, *108*, 1–28. https://doi.org/10.1016/j.jfineco.2013.01.003

Pedersen, L. H. (2015). *Efficiently Inefficient*. Princeton: Princeton University Press.

Petkova, R., & Zhang, L. (2005). Is value riskier than growth? *Journal of Financial Economics*, *78*, 187–202. https://doi.org/10.1016/j.jfineco.2004.12.001

Piotroski, J. D., & So, E. C. (2012). Identifying expectation errors in value/glamour strategies: A fundamental analysis approach. *Oxford University Press*, *25*(9), 2841–2875. https://doi.org/10.1093/rfs/hhs061

Rosenberg, B., & Houglet, M. (1974). Error Rates in CRSP and Compustat Data Bases and their Implications. *The Journal of Finance*, *29*(4), 1303–1310.

Rosenberg, B., Reid, K., & Lanstein, R. (1985). Persuasive evidence of market inefficiency. *The Journal of Portfolio Management*, *11*(3).

Sharpe, W. F. (1964). of FINANCE. *The Journal of Finance*, *19*(3), 425–442.

Shumway, T. (1997). The Delisting Bias in CRSP Data. *The Journal of Finance*, *52*(1), 327–340. https://doi.org/10.2307/2329566

Shumway, T., & Warther, V. A. (1999). American Finance Association The Delisting Bias in CRSP ' s Nasdaq Data and Its Implications for the Size Effect. *The Journal of Finance*, *54*(6), 2361–2379.

Sloan, R. G. (1996). Do Stock Prices Fully Reflect Information in Accruals and Cash Flows about Future Do Stock Prices Fully Reflect Information in Accruals and Cash Flows About Future Earnings? *The Accounting Review*, *71*(3), 289–315.

Stambaugh, R. F., Yu, J., & Yuan, Y. (2012). The short of it : Investor sentiment and anomalies $. *Journal of Financial Economics*, *104*, 288–302. https://doi.org/10.1016/j.jfineco.2011.12.001

Stambaugh, R. F., Yu, J., & Yuan, Y. (2014). The long of it: Odds that investor sentiment spuriously predicts anomaly returns. *Journal of Financial Economics*, *114*, 613–619. https://doi.org/10.1016/j.jfineco.2014.07.008

Stambaugh, R. F., Yu, J., & Yuan, Y. U. (2015). Arbitrage Asymmetry and the Idiosyncratic Volatility Puzzle. *The Journal of Finance*, *70*(5), 1903–1948. https://doi.org/10.1111/jofi.12286

Stattman, D. (1980). Book values and stock returns. *The Chicago MBA: A Journal of Selected Papers*.

Strathern, M. (1997). "Improving ratings": audit in the British University system. *European Review*, *5*(3), 305–321.

Tjero-Pech, C. O., Welden, R. N., & House, L. A. (2008). Earnings, Accruals, Cash Flows, and EBITDA for Agribusiness Firms. *Agricultural Finance Review*, *68*(2).

Yan, X. (Sterling), & Zheng, L. (2017). Fundamental analysis and the cross-section of stock returns: A data-mining approach. *Oxford University Press*, *30*(4), 1382–1423. https://doi.org/10.1093/rfs/hhx001