Does the Incongruence of Market Expectations with Fundamentals explain Stock Return Patterns?

Bernd Huefner, Marcel Rueenaufer*

Friedrich Schiller University Jena

Martin Boesch

Ernst-Abbe-Hochschule Jena

Version as of: August 20, 2022

Keywords: Reverse Engineering, Residual Income Valuation, Financial Statement Analysis, Behavioural Finance, Value Investing

JEL Classification: M41, G11, G12, G14, G17, G32

* Corresponding author at: Friedrich-Schiller University Jena, Carl-Zeiss-Str. 3, 07743 Jena, Germany.
Tel.: +49 3641 9-43154.
E-mail address: marcel.rueenaufer@uni-jena.de (M. Rueenaufer).

Acknowledgements: We thank Stephen Penman, two anonymous reviewers, Harald Jansen and all participants of the FAcT research seminars in Jena for helpful comments.

Does the Incongruence of Market Expectations with Fundamentals explain Stock Return Patterns?

Abstract

This paper uses reverse engineering of market expectations to identify potentially mispriced stocks. Building upon the "errors-in-expectations" hypothesis, we develop a theoretically funded yet practical tool for stock screening in this paper. We use the Ohlson (1995) valuation model to connect total stock returns to accounting-based fundamentals and (changes in) expected residual income levels, both in the short-term and long-term future. In a similar manner to Piotroski & So (2012), we construct a scoring index – VScore. VScore includes both fundamental data and short-term market expectations that stem from the theoretical framework we provide beforehand, where short-term expectations perform a verifying function for historical fundamentals for the determination of quality. We document that the book-to-price (B/P) effect is concentrated among firms for which long-term speculation is simultaneously incongruent to the underlying fundamentals and short-term expectation (stocks that combine cheapness with factors of quality), indicating that those stocks are interesting subjects for a more extensive fundamental analysis.

Contents

1. *Introduction*

Stock screening research has tried to find characteristics that predict future stock returns to eliminate the left tail of the return distribution. That is typically done by classifying firms into portfolios based on several predictive characteristics, such as accounting measures or cash flows, management or strategy aspects. In the context of reverse engineering (RE) in accounting-based valuation, one stream of literature stands out: stock screening based on fundamentals and multiples. In this field of research, it is understood – and consistently demonstrated – that stocks with low multiples of book-to-price (B/P) and earnings-to-price (or forward earnings-to-price, E/P) yield lower returns compared to stocks with high multiples (as shown in Fama & French (1992), Fama & French (1995), Penman (2011), Penman & Reggiani (2013) and Penman & Reggiani (2018) for example).[1]

The debate between the so-called glamour (or growth) stocks with low multiples and value stocks with high multiples has been long ongoing. Indeed, there has been a large number of contributions over the last five decades that examine the characteristics and relation of B/P and E/P with subsequent returns, such as Basu (1977), Fairfield (1994), La Porta (1996), Penman (1996), La Porta et al. (1997), Bartov & Kim (2004), Petkova & Zhang (2005), Cohen et al. (2009), Campbell et al. (2010), Piotroski & So (2012), Penman & Reggiani (2013), Lee (2015), Penman & Reggiani (2018) and Asness et al. (2019). They have in common that they (implicitly or explicitly) infer a relation between market expectations and returns through reverse engineering; high multiples indicate low expectations and vice-versa. In the literature, the return difference has been explained by risk-related aspects (e.g. due to a higher share of financially distressed firms among value firms, as Fama & French (1992) state) or behavioural aspects (as shown in Jiang et al. (2005) and Zhang (2006) for example), often implying forms of mispricing as a source of returns.

In fact, previous literature presents evidence that supports forms of the so-called "errors-in-expectations" hypothesis (EiE). Among others, Sloan (1996), Lakonishok et al. (1994), La Porta et al. (1997), Piotroski (2000), Ali et al. (2003), Bartov & Kim (2004), Piotroski & So (2012) and Li & Mohanram (2019) demonstrate that among value stocks, firms that exhibit stronger characteristics of fundamental strength or "quality" yield higher returns on portfolio levels , particularly when they simultaneously have cheap prices. Despite the large amount of research, we identified two remaining issues: first, existing stock screens include too many ad-hoc choices and assumptions. This leads to empirical assessments that prove to be methodically valid, but lack a clear and concise link to valuation theory (e.g. Piotroski (2000) and Mohanram (2005)). Second, many existing approaches are not parsimonious enough to be practical. They often require market-wide data at any given point in time or apply a large set of rather complex binaries (e.g. Li & Mohanram (2019) and Asness et al. (2019)). We tackle these issues by providing a currently lacking theoretical fundament for stock screens based on the errors-in-expectations hypothesis, anchoring on accounting-based valuation theory. We derive a scoring system that employs drivers of value as lined out by Nissim & Penman (2001) and Lee (2015) instead of ad-hoc measures of financial performance. We combine cheapness with quality through a single (and rather simple to use) scoring metric with enhanced practicality.

The remainder of this paper is structured as follows: In the ensuing Section 2, we outline loose strings on existing theory and bring them together to an overarching framework. In Section 3, we use this framework to operationalize

---

[1] In recent years, several literature contributions examine whether "value investing is dead" with mixed results. See e.g. Israel et al. (2020) or Arnott et al. (2021) and references therein for more insights on that debate. Based on our data, using subsamples for different timeframes (e.g. only the last ten years) does not change any of the main conclusions from the analysis.

both cheapness and quality and form our main empirical predictions. Afterwards in Section 4, we describe our research design, with an emphasis on the structure of the empirical analysis and the underlying data. Section 5 then includes our main analyses, consisting of four steps. First, we test the empirical validity of VScore by comparing past ROEs and expected ROEs to actual ROEs that were realized post-portfolio formation. Second, we examine whether our central empirical predictions are supported by the data. We do so by testing for differences in mean returns and their standard deviations across portfolios with varying degrees of quality and cheapness. Third, we perform regressions to control for other factors of influence (like size, momentum, and risk). Finally, we empirically link firm performance to the changes in expectations over time and investigate the drivers of ex-post returns for varying degrees of quality and cheapness.

Before we begin with the theoretical fundament, we want to make the reader aware that the goal of our paper is not to derive a standalone investment strategy, even though our results are consistent with strategies presented in previous studies (e.g. Piotroski & So (2012) or Asness et al. (2019)). We provide investors with a framework that is useful as a starting point for fundamental analysis – i.e. for stock screening – but is not meant to replace it. Kothari (2001) states – with good reason – that there are several problems in longer-run event studies that require consideration, like a potential misspecification of risk factors, a lack of a suitable specifications for the "normal" return and data problems. Therefore, a more thorough analysis of each individual firm is necessary to control for these limitations as good as possible.

2. *Theoretical Fundament*

2.1. Model Development Based on the Buildings Blocks of Valuation

Since accounting-based stock screens anchor on the (mostly implicit) extraction of market expectations, we make use of the insights on the three steps of reverse engineering for our stock screen: modelling, conceptualisation, and operationalization lined out by Huefner & Rueenaufer (2020). This section concentrates on the modelling and conceptualisation, while the subsequent section covers the operationalization. Building upon the Present Value of Expected Dividends (PVED)-condition, we employ a version of the Residual Income Model (RIM) that includes the assumptions of a constant long-term earnings growth rate and payout ratio after the precise forecast horizon (proof is delivered in Huefner & Rueenaufer (2020)):[2]

$$\mathrm{P_0 = B_0 + \underbrace{\sum_{i=1}^{T-1} \frac{E[ROE_t - r_e]B_{t-1}}{(1+r_e)^t}}_{E[STE_0]} + \underbrace{\frac{E[ROE_T - r_e]B_{T-1}}{(1+r_e)^{T-1}\big(r_e - ROE_L(1-k_L)\big)}}_{E[LTS_0]}}. \tag{1}$$

Here, $\mathrm{B_0}$ is the anchoring book value, $\mathrm{r_e}$ is the cost of capital, $\mathrm{(ROE_t - r_e)B_{t-1}}$ is the residual income for the year, $\mathrm{ROE_T}$ is the return on book value of equity for the terminal year, and $\mathrm{ROE_L}$ and $\mathrm{k_L}$ are the long-run expected ROE and payout ratio after the precise forecast horizon.[3] Within this model, price is split into an anchoring book value ($\mathrm{B_0}$), short-term expectations ($\mathrm{STE_0}$) and long-term speculation ($\mathrm{LTS_0}$).[4] Under uncertainty and imperfect

---

[2] Going concern further requires $\mathrm{E_T}$ to be positive and $\mathrm{k_L}$ to remain between 0 and 1, respectively.

[3] For the implications of such a modeling on Dividend Policy Irrelevance (DPI), we point the reader to the discussion that Huefner & Rueenaufer (2020) provide in their paper.

[4] A similar decomposition is done by Penman & Reggiani (2013), but they utilize a risk-free rate and a different terminal value expression. Further, they use capitalized residual income for the short-term expression, similar to capitalized terminal residual income in our model.

(or asymmetric) information of market participants, current expectations may be revised in future periods, leading to potential price changes.[5] As such, in our model framework prices are considered to reflect the "timeline" of ROEs that are observed (past ROEs) as well as currently expected to occur in the future; changes in prices reflect the changes in such expectations. The observed (or past) ROEs form the anchoring book value, i.e. the value from earnings over book value that the firm was permitted to realize to the present date.[6] Analogously, the currently expected ROEs consist of the short-term expectations (short-term residual income), the starting point of the convergence process (forming capitalized terminal residual income) and the long-term expected earnings growth level.[7] Through $\mathrm{ROE_L}$ and $\mathrm{k_L}$, we can infer the currently implied long-term earnings growth rate ($\mathrm{g_L}$) as follows:

$$\mathrm{g_L} = \mathrm{ROE_T}(1 - \mathrm{k_L}) = \mathrm{ROE_L}(1 - \mathrm{k_L}). \tag{2}$$

This is particularly useful because it allows the investor to get an idea of the contribution of earnings growth to long-term value creation that market participants expect from the firm. Evidently, payout also has a role to play there, which is thoroughly discussed in Huefner & Rueenaufer (2020). As a summary of the section, Figure 1 shows a systematic decomposition of stock price into building blocks, depending on the degree of speculation they include.

Figure 1

Decomposition of the Building Blocks of Valuation

<table>
<tr><th rowspan="2" colspan="2"></th><th colspan="3">Level of Decomposition</th><th colspan="2">Value Spread</th></tr>
<tr><th>Low</th><th></th><th>High</th><th>Determinant(s) of Value</th><th>Benchmark</th></tr>
<tr><td rowspan="3">Degree of Speculation</td><td>High</td><td rowspan="3">Price $\mathbf{P_0}$</td><td rowspan="2">Equity Premium $\mathbf{EP_0}$</td><td>Long-Term Speculation $\mathbf{LTS_0}$</td><td>$\mathbf{ROE_L(1 - k_L)}$ or $\mathbf{g_L}$</td><td>$\mathbf{r_e(1 - k_L)}$</td></tr>
<tr><td></td><td>Short-Term Expectation $\mathbf{STE_0}$</td><td>$\mathbf{ROE_1, ROE_2, \ldots, ROE_{T-1}}$</td><td>$\mathbf{r_e}$</td></tr>
<tr><td>Low</td><td>Book Value $\mathbf{B_0}$</td><td>Book Value $\mathbf{B_0}$</td><td>$\mathbf{ROE_{-1}, ROE_{-2}, \ldots, ROE_{-\infty}}$</td><td>$\mathbf{r_e}$</td></tr>
</table>

2.2. Moving from Prices to Stock Returns

Assuming clean-surplus accounting (CSA), the standard RIM-based expected total stock return decomposition – as shown in Shroff (1995), Penman (2016) or Penman & Yehuda (2019) for example - looks as follows:

[5] This stands in contrast to the semi-strong and strong levels of the efficient market hypothesis (EMH) lined out by Fama (1970). In a sense, the EMH is paradox: in a market setting where everyone knows everything, there is no incentive for anyone to learn anything. See Grossman & Stiglitz (1980) for further arguments on the matter.

[6] It is important to note that past payout behavior affects observed ROEs, so the fundamental investor should consider the impact of shareholder transactions on ROEs not only in the future, but also in the past to evaluate the quality of the anchoring book value as an indicator for financial performance.

[7] Lee (2015) also argues that the RIM aids the task of fundamental analysis by providing the investor with drivers of value to focus on through expected ROEs in relation to "capital in place".

$$\frac{P_1 - P_0}{P_0} + \frac{D_1}{P_0} = TSR_1 = \frac{E_1}{P_0} + \frac{(P_1 - B_1) - (P_0 - B_0)}{P_0}, \quad (3)$$

where P_t and B_t are the stock price and book value of equity at date t and D_t and E_t are dividends and earnings for the period t. This identity indicates that expected total stock returns are equal to the forward earnings yield if there is no expected change in the equity premium. There is, however, no necessity that realized returns follow this identity, because clean surplus may be violated for the subsequent year or realized earnings and dividends may not equal the expected values. Revisiting Equation (3), we observe that the evaluation of realized total stock returns requires a slight alteration regarding the first part of the equation. The realized total stock return (or cum-dividend stock return) for any period t relative to the starting point t-1 can be expressed as follows:

$$TSR_t = \underbrace{\frac{P_t - P_{t-1}}{P_{t-1}}}_{R_t} + \underbrace{\frac{D_t}{P_{t-1}}}_{DY_t} = \frac{D_t}{P_{t-1}} + \frac{B_t - B_{t-1}}{P_{t-1}} + \frac{EP_t - EP_{t-1}}{P_{t-1}}, \quad (4)$$

where EP_t denotes the equity premium at date t. Technically, the dividend yield (DY_t) in this equation is redundant as it is simply added to the buy-and-hold return (R_t) on both sides.[8] We retain it here to transparently account for the capital inflow that a dividend provides to the investor over the period.[9] The focal point of forecasting stock returns, however, lies in the prediction of the change in the equity premium (EP_t). In our model framework, a change in the equity premium implies a change in the expectations of discounted residual income that a firm faces over its current book value. Besides *ceteris paribus* considerations, the equity premium is, however, typically taken as it is without further decomposition, for example in Easton et al. (1992), Shroff (1995) or Penman (2016).
We can try to get a better understanding of the change in the equity premium by “merging” Equation (1) and Equation (4). We do so by splitting the change in the equity premium into the sum of changes in short-term expectations and long-term speculation:

$$TSR_t = \frac{D_t}{P_{t-1}} + \frac{B_t - B_{t-1}}{P_{t-1}} + \frac{STE_t - STE_{t-1}}{P_{t-1}} + \frac{LTS_t - LTS_{t-1}}{P_{t-1}}. \quad (5)$$

Equation (5) relates total stock returns to four factors of expectations on the RHS, where the factors differ in the degree of uncertainty they include (increasing from left to right). In the context of our paper, this equation provides a useful decomposition, because it allows the investor to explicitly analyse what ex-post returns were driven by for portfolios with different characteristics. Speaking of uncertainty, a risk-averse investor might recognize that the (mean) TSR for a portfolio is only one side of the coin; it lacks any specification about the risk included in the investment. There are various measures that investigate the matter. A straightforward way to analyse the risk-reward trade-off (ex-post) for a portfolio is given by the Sharpe Ratio, which we adapt as follows:[10]

[8] Note that the effect of different taxation for dividends and capital gains is neglected here for simplicity, but may have an impact in practice.
[9] Since dividends are typically positive, our return statistics may show higher values compared to similar studies that do not consider cum-dividend returns. We also report raw buy-and-hold returns whenever we deem it to be adequate.
[10] Several studies in the field also employ Sharpe ratios to demonstrate this, e.g. Asness et al. (2019).

$$\mathrm{SR}_t = \frac{\mu\left(\mathrm{TSR}_t - \mathrm{r}_{\mathrm{f}_t}\right)}{\sigma(\mathrm{TSR})}. \tag{6}$$

For a risk-averse investor, portfolios that maximize this ratio seem more appealing, because they yield a higher return in relation to the risk they bear. One part of our endeavour is the analysis of the risk-reward trade-off that investors face when considering different portfolios. But what portfolios to look for? We have yet to provide an answer to this question.

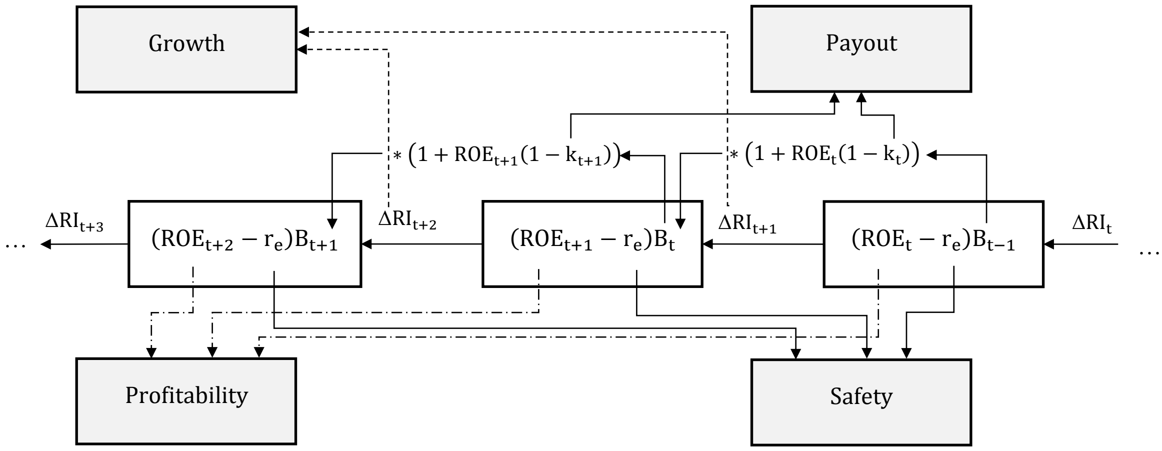
2.3. Identification of Quality Drivers

Many approaches to fundamental analysis prove to be empirically valid, but – as Nissim & Penman (2001) already criticized two decades ago – remain theoretically arbitrary. In order to identify drivers of quality with a concise link to valuation theory, a definition of quality is needed. Based on the literature, we define quality firms as those that show indicators of probable positive future business outcomes, while junk firms do not.[11] Recent literature further suggests a subcategorization of quality factors into profitability, growth, safety and payout (as listed in Lee (2015) and Asness et al. (2019) and references therein). Taking a closer look at the time-series of the payoff in the RIM – residual income – provides direct links to quality.[12] Figure 2 shows a decomposition of the time-series of changes in residual income (ΔRI) relative to date t and connects its components to the four drivers of quality: profitability, growth, safety and payout. It is evident from Figure 2 that *ceteris paribus,* increasing profitability (through higher ROE), growth (through higher ΔRI), safety (through lower $\mathrm{r_e}$) and payout (through higher $\mathrm{k_t}$) lead to increased quality.[13] Anchoring quality on the RIM also shows why ROEs are a natural starting point for valuation (and provides reasoning for the popularity of B/P and E/P).

Figure 2

Decomposition of Time-Series Changes in Residual Income with Links to Quality Drivers

[11] Such a definition is used by Li & Mohanram (2019) for example.
[12] Examples for similar DuPont-decompositions are found in Nissim & Penman (2001), Fairfield & Yohn (2001) and Palepu et al. (2019). In this equation, all stocks are beginning balances (ending values of the previous fiscal year).
[13] It is visible that payout policy influences residual income through the displacement of book values; it requires the investor to actively include a correction mechanism for earnings based on changes in payout to retain DPI. Firms paying dividends justify higher prices than non-payers if they achieve similar earnings (and earnings growth) levels. It is also important to note that payout policy can include both cash dividend payments and share repurchases, with the sum of both representing total payout, similar to Brav et al. (2005).



This figure shows a decomposition of time-series changes in residual income with links to the four quality drivers described by Lee (2015) relative to date t. Here, ΔRI_t is the period-to-period change in residual income, with residual income defined as $RI_t = (ROE_t - r_e)B_{t-1}$. B_t is the ending book value of equity for the fiscal year t, ROE_t is the return on equity which equals E_t/B_{t-1} with E_t being the earnings for the fiscal period t. r_e is the cost of equity capital and k_t is the payout ratio, which equals D_t/E_t, with D_t representing dividends for the period t.

2.4. Identification of Cheapness Drivers

In line with Lee (2015), we define cheapness as the relation between current market price and quality, i.e. as a comparative attribute. In other words, stocks are cheap (expensive) if investors pay a low (high) price for a high-quality (low-quality) asset. Logically, quality stocks should justify higher prices than junk stocks. In that case, it is essential to understand whether that inference is reflected by current market prices. Applying reverse engineering to the building blocks of valuation can aid the investor in identifying potential indicators of cheapness. As is common in the field, Piotroski & So (2012) employ B/P as a single indicator of implied market expectations and thus cheapness. In the context of the RIM, we can observe that B/P indicates the share of the market price that is currently justified by the present value of expected future residual income (i.e. the equity premium):

$$\frac{B_0}{P_0} = 1 - \frac{EP_0}{P_0}. \tag{7}$$

The EP as the target parameter however does not reveal much about the timing and uncertainty of expected residual income, as markets could expect an infinite number of different "timelines" of residual income for a given B/P. As suggested by Huefner & Rueenaufer (2020), we can use a time-based decomposition of the equity premium into two separate components, STE and LTS. They argue that the investor strives to challenge the part of prices where market participants are most likely to make mistakes: the share of price explained by long-term speculation, which is measured as LTS/P. In consequence, STE contributes to quality as a second factor on top of historical fundamentals, while LTS indicates cheapness:[14]

[14] This also sets us apart from other studies employing analyst forecasts (see Frankel & Lee (1998) for example), because they tend to simply extrapolate analyst forecasts into the long-term future instead of reverse engineering aspects of long-term speculation explicitly. Moreover, including forecast data from analysts could provide additional benefits especially for firms with information that justifies a change in price, but is not reflected in accounting figures yet. See Sloan (2019) for a more elaborate discussion.

$$\frac{\mathrm{LTS}_0}{\mathrm{P}_0} = 1 - \frac{\mathrm{B}_0 + \mathrm{STE}_0}{\mathrm{P}_0}. \tag{8}$$

2.5. Errors-in-Expectations and the Incongruence of Expectations

So far, we outlined loose strings of research that explained drivers of quality and cheapness. However, we have not provided a theoretical background for why the investor could expect a stock screen that focuses on these drivers to work. A foundation for such an investment strategy is supplied by the errors-in-expectations (EiE) hypothesis (as explained in Lakonishok et al. (1994), La Porta (1996) or Piotroski & So (2012) for example).

The errors-in-expectations hypothesis grounds on behavioural aspects, in particular on research examining the reaction of markets to a variety of signals that have a bearing on future cash flows (i.e. ROEs and residual income). Evidently, these signals can be divided into the four factors of quality lined out above. First off, markets tend to underreact to signals of profitability and growth, e.g. earnings and revenue innovations, changes in fundamental ratios or accrual reversals (see Ou & Penman (1989), Abarbanell & Bushee (1998), Sloan (1996), Piotroski (2000), Beneish et al. (2001), Doyle et al. (2003), Doyle et al. (2006), Jegadeesh & Livnat (2006) and Balakrishnan et al. (2010) for example). An outcome of that underreaction is the empirically documented post-earnings announcement drift (PEAD, see Bernard and Thomas (1989) and Bernard and Thomas (1990) for more details). Markets also tend to underreact to signals of safety and payout, such as changes in capital structure through debt and payout policy (e.g. repurchases, splits and issuances of stock, as Spiess & Affleck-Graves (1995), Ikenberry et al. (1996) and Dichev & Piotroski (1999) show) or changes in perceived distress risk (such as bond ratings, as Dichev & Piotroski (2001) show). This evidence hints to – as Piotroski & So (2012) argue – a reluctance of market participants to incorporate quality-related information that contrasts previous beliefs into their valuations, at least temporarily.

Such a reluctance will produce a market environment with temporarily biased prices.[15] These biases are then corrected over time as market participants recognize that their previous expectations were misjudged. Piotroski & So (2012) draw a crucial inference from that dynamic: if returns are (at least partially) driven by errors-in-expectations, chances for positive price revisions are highest in cases where expectations implied by prices do not align with fundamentals (i.e. where quality meets cheapness), and they provide supporting empirical evidence for their claim on portfolio levels.[16] In the context of the value-growth debate, their analysis shows that growth (value) stocks with weak (strong) fundamentals have shown the highest probability of being overvalued (undervalued) in the past.[17] The major drawbacks we see in the analysis of Piotroski & So (2012) are the theoretically rather arbitrary nature of their quality-index (FScore) and the way they measure cheapness (through B/P alone). The following section provides an improved operationalization for the key identifiers of quality and cheapness to alleviate those concerns.

---

[15] Lee (2015) also observes this and puts it into a Bayesian context: "*Apparently, Bayesian updating is a difficult cognitive task.*" An implicit assumption in all these studies is that there exists an idea of the "appropriate" reaction to newly arriving information, which implies rational and efficient Bayesian behavior, as DeBondt & Thaler (1985) argue.

[16] Lakonishok et al. (1994) label this as "contrarian investing".

[17] Such a strategy is not at all revolutionary; popular fundamental investors (e.g. Benjamin Graham, Warren Buffett and Joel Greenblatt) have been using similar strategies for a long time by buying stocks that showed attributes of cheapness and quality.

3. *Operationalization and Central Empirical Predictions*

3.1. Operationalization of Cheapness

In the previous section, we used the building blocks of valuation to infer proxies for cheapness through reverse engineering. We also provided theoretical arguments as for why a further decomposition of the equity premium could provide a better understanding of cheapness. In order to demonstrate this, we can take a closer look at Equation (8) and separate the book value and short-term expectations on the RHS:

$$\underbrace{\frac{\mathrm{LTS}_0}{\mathrm{P}_0}}_{\text{Cheapness}} = 1 - \underbrace{\left(\frac{\mathrm{B}_0}{\mathrm{P}_0} + \frac{\mathrm{STE}_0}{\mathrm{P}_0}\right)}_{\text{Quality}}. \tag{9}$$

In this equation, the RHS measures (the inverse of) quality and the LHS indicates the speculative degree of the price in relation to the quality, i.e. (the inverse of) cheapness. If the weight shifts more towards the LHS (RHS), there is a high (low) degree of speculation embedded in the price, together with a low (high) degree of quality.[18] It is inevitable to make some assumption that discriminates between the short-term and the long-term horizon (i.e. decide where the cutoff between STE and LTS is). A short horizon leads to a small effect of STE on the overall results, while a long horizon reduces the objectivity and availability of the forecasts to insert. More importantly though, it may very well be that LTS/P does not capture EiE because other factors (e.g. conservatism) influence the results. There are firms where the current long-term speculation is justified because the market recognizes growth that materializes afterwards (ignoring potential over- or underestimations of risk in this case). Then, there are no EiE to exploit and an investment strategy focusing on these errors is not fruitful.

In order to isolate LTS/P, the investor needs to get a grasp on STE/P, the short-term expectations and include it as a factor of quality. At this point, the investor might hesitate and ask: does "*expectations investing*" not tell me to challenge expectations instead of accepting them as is? The reason for the inclusion of STE/P as another indicator of quality is simple: historically, the tendency of short-term market expectations has been correct on aggregate levels.[19] Rejecting proxies for short-term market expectations entirely, especially in the short-term, would be just as naïve as accepting them without any hesitation. Furthermore, in cases where LTS/P is simultaneously incongruent with fundamentals and STE/P, market participants do not only speculate against historical fundamentals, but also short-term expectations.[20]

3.2. Operationalization of Quality

Instead of combining ad-hoc strategies to an extensive framework, our analysis employs a single-digit number of variables (like Piotroski's FScore) that are supplied by valuation theory directly (unlike Piotroski's FScore). Analogously to the scoring systems provided in previous literature, our index is constructed as the sum of binary

[18] Obviously, the book value is not necessarily a measure of quality per se, which is why a scoring system can improve the reliability of the book value. A high book value could be driven by high earnings and healthy payout, but also by excessive acquisition paired with low goodwill impairment, accounting rules or large amounts of share issues, to name just a few negative examples.

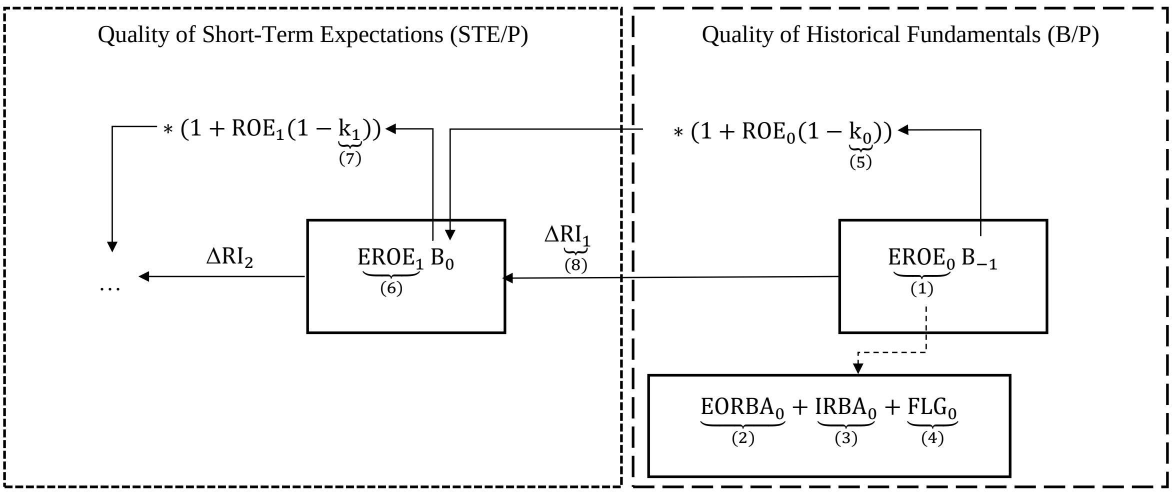
[19] Mean forecast errors of analysts for the next fiscal year are relatively small for example; Doukas et al. (2002) and Bradshaw et al. (2012) provide evidence.

[20] This is obviously limited to the extent that the investor is able to find suitable proxies for short-term market expectations; literature suggests that analyst forecasts proxy for the overall direction of short-term market expectations despite their empirically documented drawbacks (e.g. an optimistic bias and a reluctance for revisions of expectations).

variables, with a high score indicating quality and low scores indicating junk. As the previous section demonstrated, measures that indicate quality should have a bearing on future financial performance through profitability, growth, safety and payout. To the extent that historical fundamentals predict future fundamentals, they can function as the fundament for our first set of binaries.[21] Equation (9) shows that a proxy for quality should consider not only historical fundamentals, but also short-term expectations. In turn, VScore consists of eight binary variables, out of which five represent historical fundamentals (relating to B/P) and three represent short-term expectations (relating to STE/P), as shown in Figure 3.[22] The exact calculations for the quality drivers and conditions for the respective binary variable are shown in Appendix 1.

Figure 3

Selection of Quality Drivers for VScore



This figure shows the drivers of quality included in VScore as binaries. The exact calculations for the quality drivers and conditions for the respective binary variable are shown in Appendix 1. $EROE_t$ is ROE_t less r_e, which is the cost of equity capital. k_t is the payout ratio, which equals dividends scaled by earnings. $EORBA_t$ is the operating return on business assets less r_e. $IRBA_t$ is the investment return on business assets and FLG_t is the financial leverage gain to ROE_t. All stock items used for the ratios are ending balances of the previous fiscal year.

For these eight binaries, we examine whether they signal quality or not.[23] That is, we check whether the binaries are within specific intervals that indicate quality and set the variable to one if that is the case and zero otherwise. For the binaries (1) to (4), we employ the profitability measures $EROE_0$, $EORBA_0$, $IRBA_0$ and FLG_0, anchoring on a DuPont-decomposition of ROE.[24] We set the respective variable to one whenever it is positive because that indicates enough profitability to cover the cost of equity capital, well-functioning operations, a profitable financial

[21] Asness et al. (2019) provide supporting evidence. Penman & Reggiani (2018) also provide evidence for this claim. They show that for a given E/P, high B/P indicates a higher likelihood for growth not to materialize, i.e. that firms with a currently low ROE are less likely to increase it in the future and vice-versa.

[22] There is empirical evidence that suggests other measures than accounting earnings as better indicators of persistent financial performance. Sloan (1996) shows that the cash flow components of earnings is more persistent than accrual earnings; in this model however, cash accounting occurs only as a special case. Novy-Marx (2013) further shows that gross profit scaled by total assets has similar capabilities in predicting cross-sectional stock returns. Nonetheless, these insights can be used to further extend the scoring index by a cash-flow measure or alternative profitability measures.

[23] At first glance, the simultaneous inclusion of EROE and its components may seem redundant. The reason why we include both is that they express a hierarchical structure, where looking at the aggregate is helpful on its own, but it yields further insights to disaggregate. EROE alone only shows that the firm is able to exceed the cost of equity capital, but it does not explain why and how.

[24] See e.g. Fairfield & Yohn (2001), Nissim & Penman (2001) or Palepu et al. (2019) for similar decompositions.

investment strategy and/or a financing policy that contributes positively to profitability.[25] For the short-term expectations, we employ a similar strategy to the fundamentals using analyst forecasts for the next fiscal year.[26] Analyst forecasts have the advantage that they may include additional information that already affects intrinsic value, but is not reflected in historical fundamentals, e.g. more recent information concerning the business strategy or accounting quality. Ideally, we would also conduct an analogous decomposition of ROEs, but analysts often do not supply the data that would be necessary to do so, at least not publicly available. In consequence, we concentrate on the measures that are commonly predicted by analysts to retain a higher coverage of firms.

In line with the drivers of value in the RIM, analyst forecasts mainly consist of earnings and dividends, so we add EROE for the next fiscal year (EROE_1) as Signal (6) to VScore.[27] As part of the growth in book value shown in Figure 3, we employ the payout ratios for the previous fiscal year (k_0) and the expected payout ratio for the fiscal year after portfolio formation (k_1) as the basis for the binaries (5) and (7). The payout ratio is less straightforward in terms of its quality interval because it requires assumptions concerning healthy payout behaviour. We assume that a firm shows a healthy and sustainable payout behaviour if $\mathrm{k}_0 \in (0; 1)$ holds, since that indicates positive earnings and a growing book value. As a result, this variable is set to one if the firm pays dividends and/or repurchases common shares but does not overpay (whenever total payout is positive, but less than earnings). Finally, we include short-term growth in residual income ($\Delta\mathrm{RI}_1$) as Signal (8) to include an indicator of growth as well. We label this as a quality signal because growth in residual income is a value driver in the RIM-framework.[28] Similar to the profitability measures, this binary is set to one whenever it is positive. In summary, our scoring index VScore for each firm at the date of portfolio formation looks as follows:

$$\mathrm{VScore}_0 = \underbrace{\underbrace{\mathrm{V\_EROE}_0}_{(1)} + \underbrace{\mathrm{V\_EORBA}_0}_{(2)} + \underbrace{\mathrm{V\_IRBA}_0}_{(3)} + \underbrace{\mathrm{V\_FLG}_0}_{(4)} + \underbrace{\mathrm{V\_k}_0}_{(5)}}_{\text{a) Historical Fundamentals}} + \underbrace{\underbrace{\mathrm{V\_EROE}_1}_{(6)} + \underbrace{\mathrm{V\_k}_1}_{(7)} + \underbrace{\mathrm{V\_\Delta RI}_1}_{(8)}}_{\text{b) Short-Term Expectations}}. \tag{10}$$

As VScore consists of eight signals, it ranges from zero to eight. In our analysis, we classify firms into three portfolios based on VScore. Firms with a low VScore show rather negative signals for their historical financial performance, and they exhibit negative short-term expectations – they fulfil the characteristics of junk. For these firms, market participants project the financially weak performance to continue in the short-term future. Firms with a high VScore behave the opposite way; they exhibit both a strong past financial performance as well as positive short-term expectations – they fulfil the characteristics of quality. The middle portfolio consists of all firms that achieve a middle VScore, showing conflicting signals.[29]

[25] It must be noted that there may be firms that do not have financial leverage and have a positive ORBA, for which the binary is set to zero. Not having leverage does not have to be a bad sign in those cases, but we argue that for maximizing the ROE (and thus its equity value), the firm might be better off to utilize its positive spread.

[26] The focus on analyst forecasts limits the analysis to firms that are covered by analysts, obviously. As such, our final sample is significantly smaller than the sample that Piotroski & So (2012) use in their study.

[27] We limit the analysis of short-term expectations to the next fiscal year because analyst forecasts remain superior to time-series forecasts only for a limited forecast horizon (up to one year), based on empirical evidence provided by e.g. Bradshaw et al. (2012). Since time-series forecasting models typically extrapolate historical fundamentals to the future or require large pools of cross-sectional data, they would not add much value to our score. We see analyst forecasts as particularly advantageous to proxy market expectations because they are able to incorporate "other information" that is not included in historical or cross-sectional fundamentals, such as the quality of management and aspects of strategy.

[28] Aside from the obvious relevance of growth in residual income in the RIM, Ohlson & Juettner-Nauroth (2005) demonstrate this in their valuation model.

[29] The boundaries for the portfolio create narrow portfolios on both ends, reducing the number of stocks that are subjects for a further analysis.

3.3. Central Empirical Predictions

In isolation, we do not expect VScore to consistently predict stock returns. EiE implies that returns are determined by the deviation of market prices from intrinsic values, so determinants of intrinsic value alone are (expectedly) no clear indicator for future returns. While portfolios formed based on VScore will include different degrees of quality and it may have some positive effect on portfolio returns – by filtering out the financial "bad apples" in form of junk firms – we expect it to have more bearing on the risk of the investment.[30] The central supportive arguments for that claim are straightforward when re-examining the scoring system and adding in some empirical evidence from previous studies.

First, firms that are labelled with "quality" exhibit lower earnings and cash flow variability (more stable, positive EROEs and earnings growth), which in turn reduces earnings forecast errors, earnings surprises and market reactions to them.[31] Second, characteristics of quality also go hand in hand with characteristics of safety, such as low betas, low volatility, low leverage and a low risk for financial distress.[32] Since we estimate r_e through the CAPM, betas directly affect the VScore for each firm. A firm with lower beta *ceteris paribus* has an easier time reaching a positive EORBA (and EROE) and deserves a higher equity premium. Third, firms with a positive EORBA are likely to simultaneously achieve a positive FLG, so leverage does not add as much risk to their balance sheet as it does for junk firms.[33]

Finally, quality firms have a healthy payout ratio, at least either for the past or the next fiscal year. Due to a reluctance of firms to cut payout (dividends more so than repurchases), we expect payout to be fairly persistent over time.[34] As such, firms without any form of payout need to compensate the missing dividend through higher capital gains (in theory) to reach the same level of TSR.[35] This inference fits perfectly into Equation (5), since capital gains are subject to much higher uncertainty (they require the prediction of STE/P and LTS/P) than forward dividends. Combined with the difference in factors of fundamental risk, we expect the difference in the structure of TSR to lead to a lower standard deviation for quality firms compared to junk firms.

Hypothesis 1: Portfolios that include quality stocks have a lower standard deviation of TSR than portfolios including junk stocks.

Under the EiE-hypothesis, revisions of the bias in prices should be concentrated among cheap (low LTS/P) and expensive (high LTS/P) stocks for which the long-term speculation does not align with fundamentals and short-term expectations (measured by VScore). Similar to P&S, this results in a concentration of expectation errors and subsequent revisions on the top-left and bottom-right corners of Figure 4. As such, we expect TSR to increase diagonally, such that all overvalued firms yield lower returns compared to all undervalued firms. Note that unless

---

[30] Multiple previous measures of quality show significant differences in size-adjusted returns across portfolios, such as the FScore used by Piotroski (2000) and P&S and the GScore derived by Mohanram (2005). For further investigations of the effect of the FScore on returns see Turtle & Wang (2017) and Walkshäusl (2020), for example.

[31] For example, Jiang et al. (2005) and Zhang (2006) provide evidence that shows that market expectations for firms with high information uncertainty (or "value ambiguity") tend to be overvalued and riskier compared to firms with a more stable and predictable financial situation.

[32] See Lee (2015) and references therein for a more detailed analysis of safety measures and their relation to stock returns. Asness et al. (2019) provide recent evidence based on multiple factor-models and fundamental data that supports high quality stocks being safer.

[33] For high quality firms, we expect the cost of debt capital to be lower than for low quality firms, reinforcing that effect.

[34] This can, of course, turn against the investor if a firm is struggling, but has not initiated a cut in payout. Because of this, it is important to see payout in relation to the respective earnings level.

[35] Appendix 2 provides data that is in line with this prediction.

specified otherwise, we refer to all portfolios that are labelled as potentially or likely undervalued (overvalued) as undervalued (overvalued) – indicated by the three different colours in Figure 4.

Hypothesis 2: Portfolios that include undervalued stocks yield a higher TSR than portfolios including overvalued stocks.

Our last (and main) hypothesis is a logical consequence of the first two hypotheses and considers both risk and reward. In order to measure this empirically, we employ Sharpe ratios as shown in Equation (6). If risk behaves disproportionally to quality and returns follow the pattern predicted by EiE, the Sharpe Ratios increase diagonally in a similar manner to returns.

Hypothesis 3: Portfolios that include undervalued stocks have higher Sharpe ratios than portfolios including overvalued stocks; they yield higher returns for the risk they bear.

We conclude this section with a caveat: even if we discover that the empirical findings support the main theoretical predictions, that does not necessarily imply that errors-in-expectations are responsible for the return difference. Previous empirical evidence indicates that other factors may also be related to anomalous return differences, mainly effects of size, momentum, market-wide movement or a compensation for higher risk.[36] While it is possible to control for size, market-wide-movement and momentum, it is impossible to rule out the compensation for risk entirely, because it could be risk that no common asset pricing model (such as the CAPM or other factor models) recognizes. Nonetheless, it is important to include commonly applied risk measures and test whether the results change if they are taken into account.

Figure 4

Errors-in-Expectations Matrix

	High LTS/P (Expensive)	**Middle LTS/P**	**Low LTS/P (Cheap)**
Low VScore (Junk)	E[LTS/P] > E[VScore] Likely Overvalued	E[LTS/P] > E[VScore] Potentially Overvalued	E[LTS/P] = E[VScore] Aligned
Middle VScore	E[LTS/P] > E[VScore] Potentially Overvalued	E[LTS/P] = E[VScore] Aligned	E[LTS/P] < E[VScore] Potentially Undervalued
High VScore (Quality)	E[LTS/P] < E[VScore] Aligned	E[LTS/P] < E[VScore] Potentially Undervalued	E[LTS/P] < E[VScore] Likely Undervalued

4. *Data, Research Design and Descriptive Statistics*

4.1. Data and Research Design

Our empirical study consists of four main sections: first, we provide evidence on the validity of VScore as a proxy for quality by examining the differences across past ROEs, ROEs implied by analysts' earnings forecasts and actual ex-post ROEs that occurred after portfolio formation. Second, we analyse the return behaviour for portfolios

[36] For literature on the momentum, size and market anomalies see Fama & French (1992), Jegadeesh & Titman (1993), Fama & French (1995) and Carhart (1997).

that are sorted based on VScore and long-term speculation implied by LTS/P to test our three main hypotheses. We do so with separate sorts on VScore and LTS/P and joint sorts on both characteristics. Third, we test whether the main return differences we find may be attributable to other documented effects by adjusting returns for factors of size, momentum, and risk. In the last part, we analyse the decomposition of TSR for the portfolios based on Equation (5) in order to investigate the uncertainty embedded in stock returns.

Our sample consists of all US-firms for which the necessary data to estimate VScore and LTS/P are available on Thomson Reuters Worldscope and the unadjusted I/B/E/S file at portfolio formation three months after fiscal year end. As is common in the literature (see Penman & Reggiani (2013) for example), we impose that lag to assure that financial statement information for the preceding fiscal year has been largely incorporated in the stock price. We eliminate all firms with a negative book value of equity (beginning and ending balance of preceding fiscal year), stock price smaller than 3$ and all firms with missing financial data for VScore or stock price data from the sample.[37] Firms also need to have financial analyst forecasts of earnings and dividends available for the next fiscal year. In case only the dividend is missing, we set it to zero. We measure returns as annual forward TSR starting at portfolio formation, as described in Equation (5). This includes the buy-and-hold return for the subsequent year and all dividends (per share) that occurred in the timeframe.[38] We focus on annual returns here because fundamental investors typically perform longer-horizon trading and the convergence of price towards intrinsic value may take time (as e.g. Penman (2011) states).[39]

The minimum TSR and buy-and-hold return is -100%, which we apply to all firms that delist within the year after portfolio formation. We winsorize TSR at the upper percentile to reduce the impact of a few extreme outliers (all returns greater than ~230%) on the mean statistics. The cost of equity capital for VScore is computed using the CAPM with 60 monthly returns prior to portfolio formation.[40] If return data is missing or betas are negative, we employ an industry median beta using SIC industry codes. If the industry median beta cannot be calculated at the time or is also negative, we use a beta of one. The risk-free rate is equal to the 10-year US-treasury bill rate at the time and market risk premia are set to six percent market-wide for the entire sample. This leaves a total number of 39,410 annual observations in the final sample for the forecast data, ranging from 1998 to 2017.[41] For the ex-post return and ROE data, we had to exclude all observations for which the data for the next one or two years was not available. We also winsorize ROEs at -100% and 100% to reduce the impact of outliers on the statistics for this sample. This leaves 36,111 observations in the sample for the data on actuals.

An important issue to grapple with in a portfolio analysis is the determination of classification intervals. Our two-way sorting requires intervals for both VScore and LTS/P. In the literature, it is common to assume relative boundaries based on quantiles for the measurement of cheapness (B/P).[42] For the sake of simplicity and practicability, we apply absolute boundaries – 30% for the lower portfolio and 70% for the upper portfolio – that

[37] Ball et al. (1995) provide reasoning for the exclusion of low-dollar stocks. Consistent with their findings, altering the cutoff point for exclusion (to 1, 5 or 10 dollars for example) largely affects junk stocks in our sample. It significantly changes the mean return they yield, indicating that a large share of returns on them is due to extreme returns caused by low dollar value stocks that are rather illiquid. The returns of quality firms are robust to alternative cutoff points.

[38] We do not consider the effects of different taxation of capital gains and dividends in our paper.

[39] It is worth mentioning that in our sample, the return differences we document are weaker for shorter trading horizons, e.g. one month or three months after portfolio formation.

[40] The monthly returns go back until the beginning of 1994 (60 months preceding 1998).

[41] It is important to acknowledge that the inclusion of the fiscal year 2018 in the sample would have a significant negative impact on the returns for all portfolios. A large portion of returns would cover the period from the end of march 2019 to the end of march 2020, which was a peak of the downward stock market movements caused by the COVID-19 crisis.

[42] Piotroski & So (2012) use a 30/40/30 split for the classification of B/P, for example.

are time-invariant for the classification of cheapness. This approach is admittedly ad-hoc and ignores potential market-wide changes in expectations over time due to e.g. macroeconomic factors (such as COVID) or investor sentiment. The relative approach might improve the empirical validity of the portfolio strategy, but requires investors to have market-wide data available at all times.[43] For VScore, we deliberately chose to operate with a broad portfolio in the middle – all scores from 2 to 6 – and narrower outer portfolios. A stock screen that anchors on EiE and creates broad outer portfolios is less useful, since investors will still have a very large number of firms that qualify for a fundamental analysis. Broader outer portfolios would lead to a larger heterogeneity of quality within the potential investment choices. Investors typically have limited time on their hands, so having fewer "candidates" with a higher potential likelihood for subsequent price revisions seems preferable.

4.2. Descriptive Statistics and Basic Correlations

Prior to the empirical analysis, Table 1 gives statistics on the distribution of the main input variables and correlations between them. Panel A reports the means, standard deviations, and quartiles of the main input variables, namely the scoring variables of VScore, the parts of TSR and the three building blocks of valuation. It is visible that the average (median) TSR in the forecast is 7.65% (6.70%), indicating that over the last 20 years, the stock market has been growing in most years. As is typical, the distribution on returns is skewed to the right, with a maximum value of 226.15%. In comparison to other studies, the statistics include no indicators for anomalous behaviour that would require further treatment in our empirical setting.

We see that some of the parts of VScore include a set of large outliers in both directions (e.g. EROE_0 with a large range of over 400% or a high mean of EROE_1 caused by a few strongly positive values). But given that these outliers are converted into binary variables, there should be no measurable impact on the overall reliability of the results. In terms of the overall cheapness of stocks, we observe that LTS/P is positive overall, with a mean of 43.13% and a median of 54.24%. STE/P has a negligible effect overall, with both mean and median close to zero. Since we use a short forecast horizon for STE, that is not surprising. While the effect of STE on cheapness is therefore limited, we explicitly assign a higher priority to the elements of STE through VScore (three out of eight signals relate to STE). In line with Penman & Reggiani (2013), around half of the stock price is justified by long-term speculation.[44] The statistics are comparable for the actuals sample, but since we winsorize ROEs for that sample, the standard deviation of the respective components of VScore decreases significantly.

Panel B includes Pearson correlations between the main input variables. For the correlations, we included the parts of VScore as binaries to mitigate the impact of the beforementioned outliers on the correlations. As the construction of the building blocks suggests, B/P and LTS/P are strongly negatively correlated. More interestingly, B/P is also negatively correlated with STE/P, resulting in a positive correlation between LTS/P and STE/P. It seems that for firms with higher expected short-term earnings, the market also expects higher long-term earnings. Market participants thus tend to extrapolate short-term earnings to the long-term future, but the coefficient is only moderately high (0.293). Generally, the correlation between the quality characteristics and returns is rather small. They remain (mostly) statistically significant, but too small to draw a clear economic inference. The B/P effect on returns is present in the data, since B/P has a positive correlation with both TSR_1 and R_1. Consequentially, LTS/P

[43] Our results are robust to the application of time-varying relative boundaries with a 30/40/30 split as used by Piotroski & So (2012).

[44] Accounting conservatism induces positive residual income and has an impact on the input variables here. See Feltham & Ohlson (1995), Zhang (2000) and Skogsvik & Juettner-Nauroth (2013) for examples of contributions that discuss the effect of conservatism on valuation.

has a negative coefficient. The coefficients on B/P are larger than the coefficients for the quality characteristics but are also small (0.068 and 0.060 for TSR_1 and R_1, respectively).

Table 1

Descriptive Statistics of Main Input Variables and Pearson Correlations (Forecast Sample)

Panel A: Descriptive Statistics and Distribution of the Main Input Variables (Parts of VScore included as non-Binaries)

	TSR_1	R_1	DY_1	$EROE_0$	$EORBA_0$	$IRBA_0$	FLG_0	k_0	$EROE_1$	k_1	ΔRI_1	B/P	STE/P	LTS/P	r_e
μ	0.0765	0.0598	0.0168	-0.0016	-0.0171	0.0024	0.0137	0.6071	0.5154	0.1966	0.4505	0.5685	0.0002	0.4313	0.0894
σ	0.4704	0.4668	0.0384	3.1000	0.0319	2.7699	0.1528	1.3846	18.5661	29.3948	3.8736	1.1458	0.5365	0.0929	0.5019
Percentiles															
0%	-1.0000	-1.0000	0.0000	-141.4340	-30.4748	-16.4967	-120.6047	-1255.7322	-1.1103	-340.0000	-2.0000	0.0000	-5.0397	-8.8182	0.0169
25%	-0.1887	-0.2022	0.0000	-0.0713	-0.0723	0.0000	0.0000	0.0000	-0.0250	0.0000	-0.1506	0.2628	-0.0128	0.2923	0.0663
50%	0.0670	0.0446	0.0010	0.0154	-0.0116	0.0002	0.0076	0.1866	0.0400	0.0000	0.4333	0.4472	0.0160	0.5424	0.0853
75%	0.3030	0.2824	0.0229	0.0862	0.0412	0.0036	0.0398	0.8141	0.1105	0.2339	1.3739	0.7130	0.0346	0.7190	0.1076
100%	2.2615	2.2544	1.8551	374.7127	386.6731	6.1949	106.8887	2540.0000	2928.9251	240.2402	2.0000	9.9157	0.8907	0.9989	0.2066

Panel B: Pearson Correlations Between the Main Input Variables (Parts of VScore included as Binaries)

	TSR_1	R_1	DY_1	V_EROE_0	V_EORBA_0	V_IRBA_0	V_FLG_0	V_k_0	V_EROE_1	V_k_1	$V\_\Delta RI_1$	B/P	STE/P	LTS/P	r_e
TSR_1															
R_1	0.997***														
DY_1	0.135***	0.053***													
V_EROE_0	0.041***	0.035***	0.083***												
V_EORBA_0	0.043***	0.040***	0.036***	0.758***											
V_IRBA_0	-0.007	-0.006	-0.012*	0.005	0.001										
V_FLG_0	0.038***	0.030***	0.105***	0.523***	0.319***	-0.055***									
V_k_0	0.034***	0.030***	0.051***	0.297***	0.244***	-0.029***	0.271***								
V_EROE_1	0.024***	0.020***	0.060***	0.523***	0.444***	-0.032***	0.349***	0.290***							
V_k_1	0.029***	0.015**	0.171***	0.255***	0.206***	-0.014**	0.274***	0.328***	0.271***						
$V\_\Delta RI_1$	-0.020***	-0.015**	-0.061***	-0.177***	-0.139***	-0.031***	-0.118***	0.041***	0.260***	0.057***					
B/P	0.068***	0.060***	0.106***	-0.174***	-0.200***	-0.053***	-0.019***	-0.042***	-0.307***	-0.024***	-0.180***				
STE/P	0.022***	0.016**	0.077***	0.323***	0.279***	-0.027***	0.246***	0.206***	0.506***	0.199***	0.180***	-0.447***			
LTS/P	-0.077***	-0.067***	-0.127***	0.127***	0.162***	0.061***	-0.026***	0.006	0.234***	-0.012*	0.160***	-0.986***	0.293***		
r_e	-0.039***	-0.027***	-0.154***	-0.294***	-0.309***	0.024***	-0.137***	-0.096***	-0.317***	-0.144***	-0.033***	0.025***	-0.230***	0.016**	

This table shows basic descriptive statistics on the distribution and Pearson correlations of the main input variables of the analysis over the entire sample with 39,410 observations ranging from 1998 to 2017. TSR_1, R_1 and DY_1 are defined as shown in Equation (4). Panel A includes statistics on the central tendency and dispersion of the main input variables, where the parts of VScore are included as absolute numbers. $EROE_0$, $ERNOA_0$, $RNIA_0$ and FLG_0 are calculated as shown in Appendix 1. r_e is the cost of equity capital, obtained from CAPM regressions using 60 month betas for each firm, a 10-year treasury yield and a market risk premium of 6% for all firms. $k_0$and k_1 are total payout ratios for the last and forward fiscal year (k_1 is an expected ratio), calculated as the mean consensus forward dividend forecast divided by the mean consensus forward earnings forecast. $EROE_1$ is calculated as the mean consensus forward earnings forecast, divided by the most recent book value per share, less the cost of equity capital. ΔRI_1 is expected short-term growth in residual income from the last to the current fiscal year, computed as shown in Appendix 1. B/P and STE/P are calculated as shown in Equation (1) (with a one-year precise horizon) and LTS/P is obtained by reverse engineering the building blocks as shown in Equation (9). Panel B includes (pooled) Pearson correlations between the beforementioned variables and significance levels, with ***, ** and * denoting levels of 0.1%, 1% and 5%, respectively. In the correlations, the parts of VScore are included as binaries based on Equation (10) to mitigate the impact of outliers on the statistics and create consistency with the portfolio analysis.

5. *Empirical Evidence on Errors in Market Expectations and Subsequent Stock Returns*

5.1. Comparison of Actual and Expected Behaviour of ROE

Prior to the main hypothesis tests, it is necessary to empirically validate VScore as a measure for quality. We defined quality firms as firms with indicators of probable positive future economic outcomes in Section 2. In order to test VScore based on this definition, we focus on past and expected short-term EROE and compare them to actual ex-post EROEs one and two years after portfolio formation. If VScore proxies for quality, firms with a high score should achieve higher levels of EROE (as the main driver of value) after portfolio formation. Table 2 shows mean and median values of EROEs for each of the nine score portfolios. The input EROEs and actual EROEs show a rather consistent pattern; low (high) scores go hand in hand with low (high) actual EROEs – in both mean and median – indicating that VScore identifies firms that have a rather persistent negative (positive) performance after portfolio formation. However, it is important to note that these statistics exhibit a stronger survivorship bias than statistics derived from the forecast sample, since only firms that survived for at least two years after portfolio formation are included.

Table 2

Validity Test for VScore

		Scoring Input (ex-ante)				Actuals (ex-post)			
Score	**n**	**Historical $EROE_0$**		**Analyst Forecast $E[EROE_1]$**		**Actual $EROE_1$**		**Actual $EROE_2$**	
		Mean	Median	Mean	Median	Mean	Median	Mean	Median
0	603	-17.94%	-11.96%	-19.62%	-14.15%	-24.40%	-17.06%	-25.60%	-18.20%
1	2,931	-20.71%	-13.64%	-13.21%	-8.68%	-22.55%	-14.10%	-21.02%	-12.22%
2	4,871	-16.95%	-10.48%	-5.47%	-4.19%	-16.63%	-9.04%	-25.48%	-7.03%
3	4,572	-7.46%	-4.39%	1.90%	0.18%	-6.37%	-3.23%	-5.61%	-2.25%
4	4,868	1.82%	0.62%	5.03%	3.10%	-0.01%	0.62%	-0.46%	0.85%
5	5,822	7.81%	4.90%	9.55%	6.55%	6.09%	4.54%	5.52%	4.49%
6	5,957	10.96%	7.46%	12.75%	9.10%	9.75%	7.12%	8.19%	6.58%
7	4,347	12.23%	8.63%	14.75%	11.11%	11.91%	8.82%	11.62%	8.45%
8	1,476	12.11%	8.53%	15.11%	11.68%	12.50%	9.38%	12.10%	9.38%

This table shows a comparison of means and medians of scoring input and actuals EROEs within the sample including 36,111 annual observations from 1998 to 2017. Score is the sum of binary variables as indicated in Equation (25). N depicts the number of firms within the respective portfolio. $EROE_t$ is the past, expected or actual excess ROE over the cost of equity capital, respectively. For the actuals, we used data for the subsequent two years at portfolio formation date for all components of the EROE.

A one-sided t-test comparing EROE means of high-quality and low-quality firms indicates that the mean difference is statistically significant at the one percent level (t-Statistic ≈ -51.50, p-Value ≈ 0), validating VScore as a suitable proxy for quality. We also conducted extended validity tests based on the DuPont-Decomposition (not tabulated). These tests provide similar results and reinforce the impression that VScore is able to identify quality firms on portfolio levels and that quality measures are rather persistent over time.

5.2. Portfolio Analysis of Realized Total Stock Returns

Analysis of Main Hypothesis Tests

In order to test whether the data supports our main hypotheses, we perform a portfolio-based analysis of the risk-reward-trade-off in this section. The portfolios in this analysis are determined by the sorts on quality and cheapness outlined in Section 4. The first set of statistics are the standard deviation of returns within the respective portfolio (σ_{TSR_1}), the excess annual TSR over the risk-free rate, the Sharpe Ratio and the share of returns that turned out to be positive. Table 3 displays those statistics for joint and separate sorts with the respective significance test. Based on Hypothesis 1, we expected σ (TSR_1) to differ between low-quality and high-quality firms such that high-quality firms have a lower standard deviation. The results shown in Panel A support that expectation, since all low-quality portfolios have a higher standard deviation of TSR than high-quality portfolios, with mean values of 57.67% and 34.07% for all portfolios, respectively.[45] The one-sided F-test yields a statistic that is highly statistically significant, indicated by the p-Value close to zero. More importantly however, the values are economically significant. The risk inherent to low-quality portfolios is almost twice as high compared to high-quality portfolios, with slight variation across the sorts. Moreover, the standard deviation decreases consistently from top to bottom on all sorts on cheapness. It is further noticeable that the standard deviation of the cheap stocks (low LTS/P) is higher than the standard deviation of middle or high LTS/P stocks.

In an efficient market, one would expect low-quality portfolios with presumably higher risk to reward the investor with a higher return. When looking at Panel B, that is not the case. While cheap stocks yield higher returns than expensive stocks overall (10.70% versus 0.63%), high-quality firms yield higher returns compared to low-quality stocks, despite the increased risk on the low-quality end. This is a puzzling finding that other researchers have also documented in the past for various attributes of quality firms, e.g. Frankel & Lee (1998) or Campbell et al. (2008) among others. While one could argue that there is some compensation for risk when buying cheap stocks (as Fama & French (1992) argue), the compensation is neither consistent across varying degrees of quality nor strong enough to justify the assumption of overall market efficiency. This inference is reinforced by the test of Hypothesis 2. Under EiE, undervalued firms should yield higher returns than overvalued firms because prior expectations are revised upwards. Undervalued firms have strongly positive returns (mean of 9.91%), while the returns of overvalued firms are close to zero (mean of -0.51%). The one-sided t-test yields a large t-statistic and a p-Value that is close to zero as well. With a difference of around 10% in returns, the results support the second hypothesis that undervalued firms yield higher returns than overvalued firms – they are both economically and statistically significant. Since Hypothesis 3 is a logical consequence of Hypotheses 1 and Hypothesis 2, it is not surprising that the Sharpe Ratios follow the same pattern as the returns do. Panel C includes Sharpe Ratios for the portfolios and shows that the three undervalued portfolios exhibit the highest values (37.84%, 20.52% and 20.75%, respectively), whereas the overvalued portfolios have the lowest values (-12.07%, -3.83% and 1.85%, respectively).

[45] It is important to mention that this standard deviation does not equal the portfolio volatility, since it includes stocks at different times and does not consider the correlation between different stocks in the portfolio.

Table 3

Statistics on the Risk-Reward-Trade-off for Portfolios sorted on Quality and Cheapness and Main Hypothesis Tests (Forecast Sample)

Panel A: Standard Deviations and F-Test (Hypothesis 1)

	σ_{TSR_1}	Cheapness (LTS/P) Low (> 70%)	Middle (in between)	High (< 30%)	Mean	L-H
Quality (VScore)	Low (0 or 1)	61.54%	52.03%	58.92%	57.67%	-2.61%
	Middle (2 to 6)	48.00%	44.32%	52.17%	47.70%	4.17%
	High (7 or 8)	33.87%	33.07%	38.24%	34.07%	4.37%
	Mean	47.52%	43.31%	52.44%		
	H-L	-27.67%	-18.96%	-20.69%		

F-Test (One-sided for Junk > Quality)

σ (Junk)	**σ (Quality)**	**F-stat**	**p-Value**
57.67%	34.07%	2.86	< 0.0000

Panel B: Total Stock Returns and t-Test (Hypothesis 2)

	μ (TSR_1 - rf)	Cheapness (LTS/P) Low	Middle	High	Mean	L-H
Quality (VScore)	Low	-7.43%	-2.00%	8.31%	-0.03%	15.74%
	Middle	0.89%	2.63%	10.83%	4.41%	9.94%
	High	4.49%	6.79%	14.47%	6.82%	9.98%
	Mean	0.63%	2.98%	10.70%		
	H-L	11.92%	8.78%	6.16%		

t-Test (One-sided for Undervalued > Overvalued)

μ (Undervalued)	**μ (Overvalued)**	**t-stat**	**p-Value**
9.91%	-0.51%	15.90	< 0.0000

Panel C: Sharpe Ratios (Hypothesis 3)

	SR	Cheapness (LTS/P) Low	Middle	High	Mean	L-H
Quality (VScore)	Low	-12.07%	-3.83%	14.10%	-0.05%	26.17%
	Middle	1.85%	5.93%	20.75%	9.25%	18.90%
	High	13.24%	20.52%	37.84%	20.03%	24.59%
	Mean	1.33%	6.88%	20.40%		
	H-L	25.32%	24.36%	23.74%		

Panel D: Share of Positive Returns

	% of TSR_1> 0	Cheapness (LTS/P) Low	Middle	High	Mean	L-H
Quality (VScore)	Low	40.24%	48.33%	56.60%	48.82%	16.36%
	Middle	52.97%	55.64%	62.52%	56.82%	9.55%
	High	62.73%	68.87%	77.07%	67.62%	14.34%
	Mean	53.42%	57.39%	62.61%		
	H-L	22.50%	20.55%	20.47%		

This table shows statistics for nine portfolios sorted based on quality (VScore) and cheapness (LTS/P) for 39,410 firm-years from 1998 to 2017. The outer rows and columns display statistics for portfolios sorted separately based on Quality and Cheapness and the difference between the high and low portfolios for each sort Portfolios are formed three months after fiscal-year end based on absolute boundaries for VScore and LTS/P. High (low) quality firms are all firms with a VScore greater than six (lower than two). Cheap (expensive) firms are all firms with a LTS/P ratio smaller (greater) than 30% (70%). Middle VScore and LTS/P firms are all firms in between the outer portfolios. $TSR_1$is the total stock return for the year following portfolio formation, calculated as the sum of the dividend yield and the actual buy-and-hold return as depicted by Equation (4). σ_{TSR_1} is the standard deviation of TSR_1within the respective portfolio. μ (TSR_1- rf) is the excess TSR_1over the risk-free rate, where the risk-free rate is obtained as the 10-year US-treasury rate at portfolio formation for each firm. Delisting returns are set to -100% (dividends are included in delisting returns whenever applicable). SR is the Sharpe Ratio, calculated as the ratio of μ (TSR_1- rf) to σ_{TSR_1} as shown in Equation (6). The share of positive returns is calculated as the portfolio mean of a binary variable that is set to one if $TSR_1$is greater than zero and zero otherwise. The t-test (F-test) is one-sided and tests whether the mean of returns of undervalued firms (the variance of low-quality firms) is greater than the mean of returns of overvalued firms (the variance of high-quality firms). The respective mean values are weighted by N, which is the number of observations in the portfolios found in Panel C of Table 4.

As such, undervalued portfolios yield a much higher return for the risk they bear, at least in our sample. Interestingly, the low-quality cheap portfolio also has a fairly high Sharpe Ratio of 14.10%. Since the portfolio is rather risky, one could see the high return as a compensation for risk, which would support a higher level of market efficiency. Given that mean portfolio returns and standard deviations are not necessarily indicative of the ex-ante likelihood for a positive outcome, it might be helpful to investigate the share of positive returns within the portfolios. To do so, we introduced a binary variable that is set to one if TSR is positive and zero otherwise. Panel D reports the mean of this binary per portfolio, which further supports the main findings. First, it increases with EiE, as undervalued firms have a much higher share of positive returns than overvalued firms (77.07%, 68.87% and 62.52% versus 40.24%, 52.97% and 48.33%, respectively) and second, it increases with a lower degree of speculation per se (with 53.42%, 57.39% and 62.61% for the high, middle, and low LTS/P). This dynamic hints to another explanation for the higher returns on low LTS/P portfolios aside from risk: these stocks have an easier time beating the low a-priori market expectations they face, so the likelihood for positive returns increases with lower prior expectation levels. In summary, the evidence presented in Table 3 is not supportive of a fully semi-strong or strong level of market efficiency since the quality measures included in VScore were not fairly priced. It seems that over the last 20 years, the U.S. market included noticeable EiE that were subsequently revised. Unless the control regressions indicate the opposite, it is reasonable to assume that the mantra of Graham & Dodd (1934) to pick stocks that are "*selling well below the levels apparently justified by a careful analysis of the relevant facts*" continues to have merit.

Analysis of Additional Portfolio Statistics

The cheapness measure LTS/P can include various combinations of earnings behaviour and payout, implying different developments of book values over longer horizons. It may be helpful for the understanding of market expectations to dig deeper into the implied long-term speculation. That way, it is possible to elicit the expected evolution of value drivers within the RIM that is embedded in the current price, such as long-term earnings or ROEs. For the extraction of g_L based on Equation (1) and (2) however, assumptions regarding k_L are required. Appendix 2 shows the long-term behaviour of payout five years after portfolio formation within our forecast sample, similar to Nissim & Penman (2001). The analysis demonstrates that firms that overpay converge downwards from a full payout ratio to a value of ~ 70%, firms with healthy payout increase payout slightly from ~40% to a value of ~ 60% and firms that do not pay dividends typically start paying dividends within a few years, converging upwards to a value of ~ 30% after 5 years. In consequence, we set the payout ratio to values that correspond to that mean-reversion process, such that the long-term payout is 30% for firms with no current payout, 60% for firms with a healthy current payout ratio and 70% for firms that overpaid recently. Using that procedure, Panels A and B of Table 4 show additional portfolio statistics for g_L and k_L. As one would expect based on Table 3, returns behave disproportionally to implied earnings growth. Overvalued stocks exhibit much higher implied earnings growth; growth that investors pay for when buying these stocks.

The difference in earnings growth between the contrarian portfolios is roughly 10% (12.48% for low-high and 2.46% for high-low stocks). There are several factors that contribute to this discrepancy in earnings growth. First, as Panel D of Table 4 suggests, expensive junk stocks have a much higher beta than cheap quality stocks, with values of 1.40 and 0.69 (this also applies to overvalued stocks versus undervalued stocks). In the RIM, a higher beta leads to a higher discount rate and a smaller EROE, all else equal. It is thus natural that these firms need to achieve higher growth to justify a similar price. Second, Panel B shows that quality firms have a higher expected long-term payout ratio than junk firms (with values of 52.81% and 43.30%, respectively). Firms with a higher payout ratio need less earnings growth to justify the same price, because higher payout reduces book values and increases residual income (through ROE_L). Third, the lower quality goes hand in hand with lower earnings included in both B/P and STE/P. This shifts more earnings to the long-term future; earnings that are subject to risk and might not materialize.[46] It seems that markets underestimate the risk at times, at least for the stocks within the overvalued portfolios.

Asness et al. (2019) find evidence that quality stocks have higher prices, such that they have a lower overall B/P (higher P/B in their study). Since quality factors are documented to be rather persistent, quality firms justify a higher LTS/P than junk firms. Panel C of Table 4 reports the number of observations within each portfolio and the share of that portfolio in the entire sample. If quality firms justify a higher LTS/P, there should be a higher share of firms within the quality portfolio that is priced accordingly. Indeed, within the quality portfolio, 2,160 firms (35.17%) are classified as high LTS/P stocks, while only 676 (11.01%) are classified as low LTS/P. On the low-quality end, the shares are not as one-sided.

Within the junk portfolio, there are 1,270 firms (29.18%) that are classified as high LTS/P and 1,500 firms (34.46%) are classified as low LTS/P. There are more firms that have low expectations included in the junk portfolio as well, but the difference is not large. A potential explanation for this discrepancy in pricing is provided by Jiang et al. (2005). They demonstrate that information uncertainty has a role to play in the EiE-discussion. Junk firms exhibit a higher value ambiguity than quality firms, such that market participants are unsure as to what the right price for these firms could be – the difference in betas across portfolios makes that apparent. It is thus not surprising that the share of firms with incongruent expectations is larger for junk firms.[47]

[46] Penman & Reggiani (2013), Penman & Reggiani (2018) and Penman et al. (2019) discuss these dynamics of accounting in greater detail and provide additional insights. In a way, their study relates to our work since they also examine the degree of speculation embedded in prices and the relation between fundamentals (through E/P and B/P), risk and growth.

[47] We also investigated further characteristics for the portfolios, such as size, age and shorter horizon returns (monthly, quarterly, semi-annually). Size and age are tabulated in **Appendix 3**. Firstly, these findings revealed that junk firms are typically younger and smaller than quality firms. Secondly, the return-spread we document in the main hypothesis tests is present for shorter horizons, but to a lesser degree. It increases over time, especially for the returns longer than three months after portfolio formation. The standard deviation effect is also persistent for shorter horizons, as is the difference in Sharpe Ratios.

Table 4

Additional Portfolio Statistics on Long-Term Market Expectations, Portfolio Size and Risk (Forecast Sample)

Panel A: Implied Long-Term Earnings Growth

	g_L	Cheapness (LTS/P) Low (> 70%)	Middle (in between)	High (< 30%)	Mean	L-H
Quality (VScore)	Low (0 or 1)	12.48%	7.46%	5.00%	8.08%	-7.48%
	Middle (2 to 6)	6.89%	5.35%	3.64%	5.28%	-3.25%
	High (7 or 8)	4.89%	3.96%	2.46%	4.12%	-2.44%
	Mean	7.14%	5.28%	3.76%		
	H-L	-7.59%	-3.50%	-2.54%		

Panel B: Long-Term Payout Ratio

	k_L	Cheapness (LTS/P) Low	Middle	High	Mean	L-H
Quality (VScore)	Low	36.59%	46.34%	50.11%	44.80%	13.52%
	Middle	50.58%	53.76%	55.52%	53.41%	4.94%
	High	61.63%	60.79%	60.34%	61.04%	-1.28%
	Mean	51.14%	54.39%	55.04%		
	H-L	25.03%	14.45%	10.23%		

t-Test (One-sided for Undervalued < Overvalued)

µ (Undervalued)	**µ (Overvalued)**	**t-stat**	**p-Value**
3.66%	7.66%	-49.1344	< 0.0000

t-Test (One-sided for Junk < Quality)

µ (Junk)	**µ (Quality)**	**t-stat**	**p-Value**
51.06%	61.08%	15.9648	< 0.0000

Panel C: Share and Number of Observations per Portfolio

	N	Cheapness (LTS/P) Low	Middle	High	Sum	L-H
Quality (VScore)	Low	1,270 *(3.22%)*	1,583 (4.02%)	1,500 (3.81%)	4,353 (11.05%)	230 (0.58%)
	Middle	7,553 *(19.17%)*	13,462 (34.16%)	7,900 (20.38%)	28,915 (73.37%)	347 (0.88%)
	High	2,160 *(5.48%)*	3,306 (8.39%)	676 (1.69%)	6,141 (15.58%)	-1,484 (-3.77%)
	Sum	10,983 *(27.87%)*	18,622 *(46.56%)*	10,076 *(25.57%)*	39,410 (100%)	
	H-L	890 *(2.26%)*	1,723 *(4.37%)*	-824 *(-2.09%)*		

Panel D: Beta

	β	Cheapness (LTS/P) Low	Middle	High	Mean	L-H
Quality (VScore)	Low	1.40	1.56	1.48	1.48	0.08
	Middle	1.31	1.19	1.18	1.22	-0.13
	High	0.95	0.84	0.69	0.86	-0.26
	Mean	1.25	1.16	1.19		
	H-L	-0.45	-0.72	-0.79		

This table shows statistics for nine portfolios sorted based on quality (VScore) and cheapness (LTS/P) for 39,410 firm-years from 1998 to 2017. The outer rows and columns display statistics for portfolios sorted separately based on Quality and Cheapness and the difference between the high and low portfolios for each sort Portfolios are formed three months after fiscal-year end based on absolute boundaries for VScore and LTS/P. High (low) quality firms are all firms with a VScore greater than six (lower than two). Cheap (expensive) firms consist of firms with a LTS/P ratio smaller (greater) than 30% (70%). Middle VScore and LTS/P firms are all firms in between the outer portfolios. Implied long-term earnings growth is the perpetual growth rate in earnings, obtained by computing Equation (2). k_L is the long-term payout ratio, which we set to a fixed value depending on the interval of the payout ratio of the previous year (k_0): 30% for $k_0 = 0$, 60% for $0 < k_0 < 1$ and 70% for $k_0 \geq 1$. N is the total number of observations per portfolio, with the per-cent share of the respective portfolio in the sample in parentheses behind the absolute number. Beta is the five-year monthly beta with the S&P 500 used as the reference index. If the beta for a firm is negative, a median industry beta at the time is used. If the industry median beta is negative, a beta of one is used. The respective mean values for Panels A, B and D are weighted by N, which is the number of observations in each portfolio found in Panel C.

5.3. Controlling for Other Factors of Influence on Stock Returns

In summary, the portfolio analysis provided empirical evidence for a relation between returns and EiE in the last two decades; more importantly, a stock screen that tries to combine quality with cheapness – i.e. exploits these errors – can be a fruitful starting point for a fundamental analysis. In order to control for other factors of influence, we perform the following (nested) Fama & Macbeth (1973) regression, with standard errors adjusted for heteroscedascity and autocorrelation based on Newey & West (1987):

$$R_{it+1} = \underbrace{\beta_1 \mathrm{LikUnderV}_{it} + \beta_2 \mathrm{PotUnderV}_{it} + \beta_3 \mathrm{Aligned}_{it} + \beta_4 \mathrm{PotOverV}_{it} + \beta_5 \mathrm{LikUnderV}_{it}}_{\text{Classification Factors}} + \underbrace{\beta_6 \mathrm{MKT}_{it} + \beta_7 \mathrm{SMB}_{it} + \beta_8 \mathrm{UMD}_{it}}_{\text{Control Factors}} \tag{11}$$

The first five independent variables are binary factors set to one if the stock belongs to a specific set of portfolios based on Figure 4. The other three independent variables are equivalent to the respective Fama & French (1995) and Carhart (1997) factors and stem from the Kenneth French data library. In this case, the dependent variable is the buy-and-hold return instead of TSR to create consistency with the standard CAPM and multi-factor model regressions. The intercept is suppressed in the estimation to ensure non-collinearity among the classification variables. We do not control for the B/P anomaly and the PEAD because they are components of EiE; B/P mirrors LTS/P and PEAD is reflected in the quality index because we use long-term returns. Therefore, we expect the classification variables to capture the variation induced by their effects. The concern of the main test is whether the added anomaly factors affect the significance of the classification variables or add significant explanatory power over the EiE-factors. If they do, it might be that EiE is simply an outcome of other anomalies that have already been documented, either in isolation or when combined. Table 5 reports the results of the regression tests. Generally, the adjusted R^2 of the regressions is moderate, with values ranging from 17.84% in the regression with only classification variables to 20.74% in the regression with all control factors. As expected and consistent with Table 3, the coefficients of undervalued firms are consistently higher than the coefficients of overvalued firms across all regressions. The coefficients behave disproportionally to the predicted EiE, in line with the central empirical predictions. In fact, the only classification variables that produce a significance level larger than 0.1% are likely overvalued and potentially overvalued firms. Since the distribution of returns within these portfolios averages close to zero in our sample, that is not surprising. Also, it is noticeable that because of the positive average return in the sample, most of the coefficients have statistically significant positive signs, even the undervalued firms in several cases. These effects persist regardless of the included control factors. The inclusion of control factors has two visible (but minor) effects: first, the adjusted R^2 increases with the addition of the control factors, but just by a few percent. Second, especially when MKT is added, the coefficients adjust downwards (slightly), but remain significant and retain their respective sign. Some of the positive average return is explained by the bull market within the sample period, which also explains why the coefficients adjust downwards whenever MKT is included.

Table 5

Control Regressions – Results (Forecast Sample)

	(1)	(2)	(3)	(4)	(5)
LikUnderV	0.1437***	0.1068***	0.1486***	0.1326***	0.1156***
	(0.0147)	(0.0143)	(0.0182)	(0.0142)	(0.0142)
PotUnderV	0.1124***	0.0810***	0.1160***	0.1089***	0.0928***
	(0.0052)	(0.0051)	(0.0045)	(0.0049)	(0.0051)
Aligned	0.0575***	0.0334***	0.0621***	0.0615***	0.04784***
	(0.0039)	(0.0040)	(0.0036)	(0.0038)	(0.0040)
PotOverV	0.0329***	0.0112*	0.0367***	0.0401***	0.0269***
	(0.0054)	(0.0055)	(0.0050)	(0.0053)	(0.0055)
LikOverV	-0.0368*	-0.0599***	-0.0335*	-0.0376*	-0.0497**
	(0.0173)	(0.0173)	(0.0132)	(0.0173)	(0.0174)
MKT		0.0187***			0.0103***
		(0.0008)			(0.0008)
SMB			-0.0062***		-0.0009
			(0.0009)		(0.0010)
UMD				-0.0180***	-0.0136***
				(0.0008)	(0.0008)
Adj. R^2	0.1784	0.1930	0.1845	0.2027	0.2074

This table reports average Fama & Macbeth (1973)-regression coefficients and Newey & West (1987)-adjusted standard errors of 20 annual cross-sectional regressions from 1998 to 2017:

$$R_{it+1} = \underbrace{\beta_1 \mathrm{LikUnderV}_{it} + \beta_2 \mathrm{PotUnderV}_{it} + \beta_3 \mathrm{Aligned}_{it} + \beta_4 \mathrm{PotOverV}_{it} + \beta_5 \mathrm{LikOverV}_{it}}_{\text{Classification Factors}} + \underbrace{\beta_6 \mathrm{MKT}_{it} + \beta_7 \mathrm{SMB}_{it} + \beta_8 \mathrm{UMD}_{it}}_{\text{Control Factors}}.$$

R_{it+1} is the one-year ahead raw buy-and-hold return for the firm, starting three months after fiscal year-end. Delisting returns are set to -100%. The five classification variables are set to one if a firm belongs to a set of portfolios described in Figure 4; LikUnderV is set to one for all likely undervalued firms, PotUnderV for all potentially undervalued firms, Aligned for all firms with aligned expectations, PotOverV for all potentially overvalued firms and LikOverV for all firms that are classified as likely overvalued. These five variables are set to zero if the respective condition is not met. The control factors MKT, SMB and UMD stem from Kenneth French's website and correspond to the CAPM- and multifactor-model coefficients at the time as described by Fama & French (1992) and Carhart (1997). The intercept term is suppressed to rule out multicollinearity among the classification variables. ***, ** and * denote significance levels of 0.1%, 1% and 5% for the coefficients, respectively.

5.4. Decomposition of Ex-Post Total Stock Returns for Errors-in-Expectation-Portfolios

So far, the data supported the empirical predictions: even after controlling for empirically documented anomalies, the returns of portfolios that exploit EiE remain abnormally positive. This section takes a deeper dive into the patterns of ex-post stock returns and tries to explain why stock returns of junk firms are subject to higher risk compared to the returns of quality firms. Perhaps, a re-examination of Equation (5) can shed some light on the matter:

$$\mathrm{TSR}_t = \underbrace{\frac{D_t}{P_{t-1}} + \frac{\Delta B_t}{P_{t-1}} + \frac{\Delta \mathrm{STE}_t}{P_{t-1}}}_{(1)(\mathrm{QR})} + \underbrace{\frac{\Delta \mathrm{LTS}_t}{P_{t-1}}}_{(2)(\mathrm{CR})}. \tag{12}$$

Within this equation, we can identify two separate components of TSR. Term (1) includes all parts of the stock return that are due to changes in quality (QR), and (2) the part of returns that is due to changes in cheapness (CR). In turn, it may be insightful to examine the typical behaviour of these components. More crucially though, this allows the explanation of stock returns across different quality-cheapness portfolios with the EiE-hypothesis. Quality stocks will naturally have an easier time reaching a more positive outcome on (1), because they have a growing book value with high earnings yields and a healthy payout ratio with higher dividend yields. As long as the change in STE does not negate that positive effect, one would expect quality firms to have a more positive outcome on QR.[48] Junk stocks are expected to behave the opposite way. The expected outcome on (2) is more difficult to assess and requires a closer examination of EiE.

In this section, all returns are market-adjusted to control for macroeconomic effects over time.[49] For aligned firms, the expected market-adjusted return is indifferent from zero. If QR leads to values of TSR that would otherwise deviate too strongly from zero, CR compensates the positive (negative) effect on TSR for quality (junk) firms with congruent expectations, such that the signs of QR and CR are different. For undervalued (overvalued) portfolios, CR can be zero or positive (negative), since QR alone creates return patterns consistent with EiE. That is, we expect TSR to be dictated by QR, and CR to assume values that make TSR align with EiE. Table 6 shows the expected signs for QR and CR for all nine portfolios along with the mean market adjusted TSR, QR and CR. We see in Panel B of Table 6 that in most cases, the data on TSR matches the expected signs, with coefficients that are statistically and economically different from zero whenever it was expected. Panel C of Table 6 reveals that quality firms show a higher average QR than junk firms, with values of 2.95% and -6.37%, respectively. Interestingly, QR is larger for expensive stocks compared to cheap stocks, with values of 2.00% and -5.05%, respectively. This is largely due to the different distributions of junk and quality firms within the different levels of cheapness.

Referring back to Table 4, we see that the share of junk firms within the cheap (expensive) portfolios is larger (smaller), so it is not surprising that the overall QR for cheap (expensive) firms is negative (positive). While this explains the overall difference, it does not necessarily explain the disparity of QR between the middle-quality portfolios – the shares are rather evenly distributed there. These portfolios cover a wider range of scores (2 to 6) and therefore include considerable heterogeneity in terms of quality. The lower (higher) QR for cheap (expensive) middle-quality firms could be an outcome of lower (higher) quality. Indeed, the average VScore for expensive middle-quality firms is 4.33, while cheap middle-quality firms have an average score of 3.77.[50] It seems that within these broad portfolios, a larger share of firms with lower scores is concentrated in the cheaper portfolios, explaining the tendency for a lower QR.

Now, if EiE dictates return behaviour and QR does not align with EiE on its own, we expect CR to assume a correctional role for TSR. Since the values of the market-adjusted TSR were in line with EiE, it is a logical consequence that CR must follow this pattern – Equation (12) demonstrates this. Panel D of Table 6 shows the mean market-adjusted CR for the portfolios formed based on quality and cheapness. As expected, CR is negative (positive) for firms with QR that exceeds (falls short of) the returns implied by EiE. This evidence hints to a somewhat predictable pattern for both QR and CR that might aid investors in the search for lucrative investments. The patterns for QR and CR also explain why stock returns for junk stocks are riskier; as an outcome of the "fundamental" reasons mentioned earlier; they have a lower QR (particularly in the middle and cheap portfolios). In turn, CR must cover the difference to reach a similar return – and CR is subject to higher risk since it requires the estimation of the long-term behaviour of residual income.

[48] The effect of STE on TSR is rather small (often indistinguishable from zero), which is why we focus on the dividend yield and changes in book values in the analysis.

[49] The average market return within the timeframe is ~6.66%. As such, the expected returns on the nine portfolios evolve around rather positive values. Adjusting the returns for market movement results in predictable signs for TSR, QR and CR.

[50] The mean difference is statistically significant at the 0.1% level based on a two-sample t-test.

Table 6

Ex-Post Return Decomposition – Market-Adjusted Portfolio Returns due to Quality and Cheapness (Actuals Sample)

Panel A: Expected Signs of Market-Adjusted Returns due to Quality (QR) and Cheapness (CR)

		Cheapness (LTS/P)		
		Low (> 70%)	Middle (in between)	High (< 30%)
Quality (VScore)	Low (0 or 1)	$QR < 0 \mid CR \leq 0$	$QR < 0 \mid CR \leq 0$	$QR < 0 \mid CR > 0$
	Middle (2 to 6)	$QR \approx 0 \mid CR < 0$	$QR \approx 0 \mid CR \approx 0$	$QR \approx 0 \mid CR > 0$
	High (7 or 8)	$QR > 0 \mid CR < 0$	$QR > 0 \mid CR \geq 0$	$QR > 0 \mid CR \geq 0$

Panel B: Mean Market-Adjusted Total Stock Returns (TSR)

	$\mu(\mathbf{TSR_i - TSR_M})$	Cheapness (LTS/P) Low	Middle	High	Mean	L-H
Quality (VScore)	Low	-12.70%***	-6.73%***	2.21%	-5.38%	14.91%
	Middle	-2.56%***	-1.47%***	5.53%***	0.07%	8.09%
	High	0.64%	3.11%***	10.17%***	3.00%	9.53%
	Mean	-2.91%	-1.05%	5.40%	Market Average: 6.66%	
	H-L	13.34%	9.83%	7.96%		

Panel C: Mean Market-Adjusted Returns due to Quality (QR)

	$\mu(\mathbf{QR_i - QR_M})$	Cheapness (LTS/P) Low	Middle	High	Sum	L-H
Quality (VScore)	Low	-1.06%**	-3.69%***	-13.84%***	-6.37%	-12.78%
	Middle	2.29%***	1.50%***	-4.23%***	0.21%	-6.52%
	High	2.51%***	3.25%***	2.90%**	2.95%	0.39%
	Mean	2.00%	1.40%	-5.05%	Market Average: 1.61%	
	H-L	3.57%	6.94%	16.74%		

Panel D: Mean Market-Adjusted Returns due to Cheapness (CR)

	$\mu(\mathbf{CR_i - CR_M})$	Cheapness (LTS/P) Low	Middle	High	Mean	L-H
Quality (VScore)	Low	-11.65%***	-3.03%*	16.05%***	0.99%	27.70%
	Middle	-4.85%***	-2.97%***	9.76%***	-0.15%	14.61%
	High	-1.87%**	-0.14%	7.27%***	0.04%	9.15%
	Mean	-4.91%	-2.45%	10.45%	Market Average: 5.05%	
	H-L	9.77%	2.89%	-8.78%		

t-Test (One-sided for Junk < Quality)

μ (Junk)	**μ (Quality)**	**t-stat**	**p-Value**
-6.37%%	2.95%	-23.2882	< 0.0000

t-Test (One-sided for Undervalued > Overvalued))

μ (Undervalued)	**μ (Overvalued)**	**t-stat**	**p-Value**
6.67%	-5.31%	21.0907	< 0.0000

This table shows expected signs and ex-post return statistics for nine portfolios sorted based on quality (VScore) and cheapness (LTS/P) for 36,514 firm-years from 1998 to 2017. The outer rows and columns display statistics for portfolios sorted separately based on Quality and Cheapness and the difference between the high and low portfolios for each sort. Portfolios are formed three months after fiscal-year end based on absolute boundaries for VScore and LTS/P. High (low) quality firms are all firms with a VScore greater than six (lower than two). Cheap (expensive) firms consist of firms with a LTS/P ratio smaller (greater) than 30% (70%). Middle VScore and LTS/P firms include all firms in between the outer portfolios. $\mu(\mathrm{TSR_i} - \mathrm{TSR_M})$ is the mean total stock return less the respective market return for the portfolio, with $\mathrm{TSR_i}$ calculated as shown in Equation (12). $\mu(\mathrm{QR_i} - \mathrm{QR_M})$ is the mean stock return due to (changes in) quality less the market return due to quality for the portfolio, with $\mathrm{QR_i}$ calculated as term (1) in Equation (12). $\mu(\mathrm{CR_i} - \mathrm{CR_M})$ is the mean stock return due to (changes in) cheapness less the market return for the portfolio, with $\mathrm{CR_i}$ calculated as term (2) in Equation (12). ***, ** and * denote significance levels of 0.1%, 1% and 5% for t-tests of mean values to be different from zero, respectively.

6. *Concluding Remarks*

In this paper, we provide investors with a theoretically funded and practical tool for stock screening. Within our sample, our scoring index VScore identifies quality firms (firms that exhibit characteristics of profitability, growth, safety and payout). Furthermore, combining VScore with a measure for cheapness – operationalized by the part of the market price explained by long-term speculation – enables investors to screen markets for stocks that are lucrative candidates for a fundamental analysis. Using data on fundamentals, returns and analysts´ forecasts for the US-market, this paper demonstrates that the errors-in-expectations hypothesis (EiE) can explain return patterns when total (i.e. cum-dividend) stock returns are employed. Our main analysis of the risk-reward trade-off for portfolios – through Sharpe ratios sorted on quality and cheapness – shows that firms that are identified as undervalued yield higher returns, but also inherit lower fundamental risk. Consistent with other contributions to the field and absent of a risk-based explanation, this puts a rather big question mark on the hypothesis of consistently efficient markets. We further used the building blocks of valuation in accounting-based valuation models to show what drives returns; fundamental data and short-term expectations are rather persistent and create a sizeable and predictable share of returns; returns due to (changes in) quality (QR). Combining the assumption of EiE with the persistent development of QR results in predictable patterns for the share of returns that is due to (changes in) cheapness (CR) – at least on aggregate.

Obviously, VScore largely focuses on the relation of financial ratios to the prospective analysis and valuation of market participants to get a grasp on potentially lucrative investments. As is common in the field of research, strategies like the one we present are able to shift the distribution of returns to the right, but the data remains quite noisy. The share of positive returns for undervalued portfolios is significantly larger than for overvalued firms, but still includes a large share of returns that goes against the predicted direction. Ironically, the quality measure VScore does not include much qualitative information; it does not inform about the business strategy, the accounting (e.g. the conservatism) behind the numbers employed or management decisions that affect prices, but not necessarily historical fundamentals. As a result, the concluding point of this series is a reminder of a caveat we mentioned before: the approach provided in this paper is theoretically funded, parsimonious and empirically valid, so it provides a constructive starting point for a more comprehensive fundamental analysis; but it should in no way be seen as a valid stand-alone approach to fundamental analysis.

7. *Appendix*

Appendix 1

Detailed Formulae and Conditioning of the Scoring Binaries used for VScore:

Panel A: Historical Fundamentals

Binary Number	**Formula**	**Set to one if…**
(1)	$\mathrm{EROE_0} = \frac{\mathrm{E_0}}{\mathrm{B_{-1}}} - \mathrm{r_e}$	$\mathrm{EROE_0} > 0$
(2)	$\mathrm{EORBA_0} = \frac{\mathrm{NOPAT_0}}{\mathrm{TA_{-1}} - \mathrm{NFL_{-1}}} - \mathrm{r_e}$	$\mathrm{EORBA_0} > 0$
(3)	$\mathrm{IRBA_0} = \frac{\mathrm{NIPAT_0}}{\mathrm{TA_{-1}} - \mathrm{NFL_{-1}}}$	$\mathrm{IRBA_0} > 0$
(4)	$\mathrm{FLG_0} = \left(\frac{\mathrm{NOPAT_0} + \mathrm{NIPAT_0}}{\mathrm{TA_{-1}} - \mathrm{NFL_{-1}}} - \frac{\mathrm{NIEAT}}{\mathrm{TD_{-1}}}\right)\frac{\mathrm{TD_{-1}}}{\mathrm{B_{-1}}}$	$\mathrm{FLG_0} > 0$
(5)	$\mathrm{k_0} = \frac{\mathrm{D_0}}{\mathrm{E_0}}$	$0 < \mathrm{k_0} < 1$

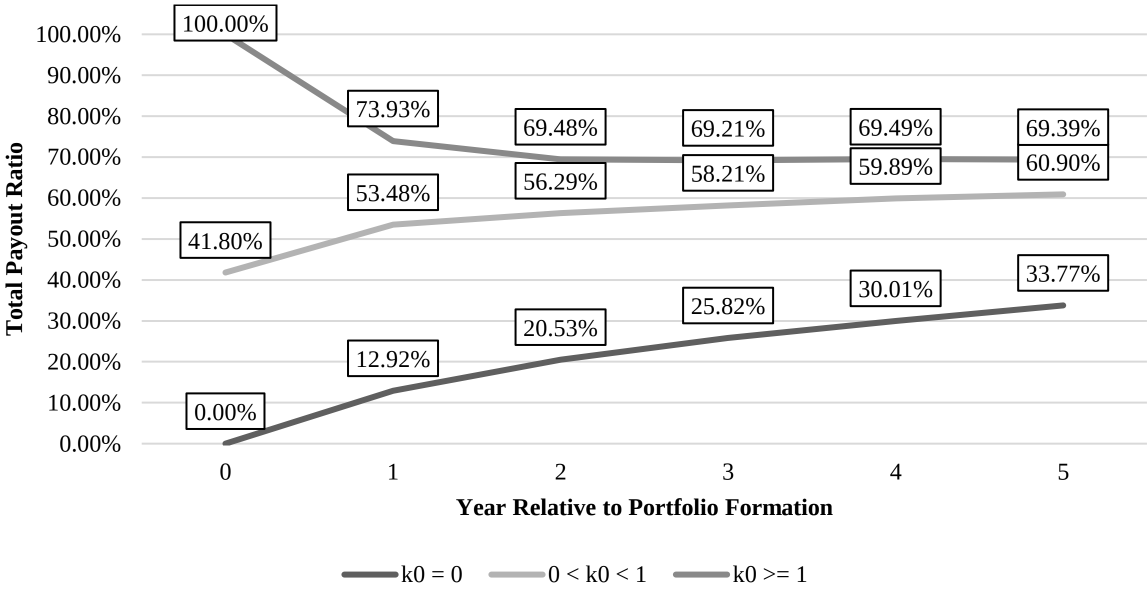
Panel B: Short-Term Expectations

Binary Number	**Formula**	**Set to one if…**
(6)	$\mathrm{EROE_1} = \frac{\mathrm{E_1}}{\mathrm{B_0}} - \mathrm{r_e}$	$\mathrm{EROE_1} > 0$
(7)	$\mathrm{k_1} = \frac{\mathrm{D_1}}{\mathrm{E_1}}$	$0 < \mathrm{k_1} < 1$
(8)	$\Delta\mathrm{RI_1} = \frac{\mathrm{EROE_1} * \mathrm{B_0} - \mathrm{EROE_0} * \mathrm{B_{-1}}}{\frac{(\lvert\mathrm{EROE_1} * \mathrm{B_0}\rvert + \lvert\mathrm{EROE_0} * \mathrm{B_{-1}}\rvert)}{2}}$	$\Delta\mathrm{RI_1} > 0$

This table shows the exact calculations of the binaries used for VScore and the condition for each binary to be set to one. Panel A includes the binaries for historical fundamentals, while Panel B includes binaries for short-term expectations. In case the condition is not fulfilled, the binary is set to zero. $\mathrm{B_t}$, $\mathrm{E_t}$ and $\mathrm{D_t}$ depict the ending book value of equity, earnings, and dividends for the respective fiscal year. $\mathrm{TA_t}$, $\mathrm{NFL_t}$ and $\mathrm{TD_t}$ are ending balances of total assets, non-financial liabilities, and total debt. $\mathrm{NOPAT_t}$, $\mathrm{NIPAT_t}$ and $\mathrm{NIEAT_t}$ are the net operating income, net investment income and net interest expense for the fiscal year, all after tax. For after tax calculations, the effective tax rate is used, obtained as the difference between net income before and after taxes divided by net income before taxes. $\mathrm{EROE_t}$ is the excess return on equity over the cost of equity capital ($\mathrm{r_e}$). $\mathrm{EORBA_t}$ and $\mathrm{IRBA_t}$ the excess operating return and investment return on business assets, with business assets defined as the difference between total assets and non-financial liabilities. $\mathrm{FLG_0}$ is the financial leverage gain to ROE, calculated as the spread between the return on business assets and the cost of debt capital, multiplied with financial leverage. k_t is the payout ratio for the fiscal year and $\Delta\mathrm{RI_1}$ depicts the short-term change in residual income, scaled by the average of sum of absolute values of residual income for the previous and ongoing fiscal year (similar to Penman & Reggiani (2018)). Note that this approach entails simplifying assumptions, namely that all items besides interest expense, investment income and interest income are considered as factors of operating income. It is possible to distinguish financial assets from operating assets to get further insights, but business assets form a more reliable basis since they remain positive when data is available (NFL is typically smaller than TA). We chose to apply these assumptions to make the approach more straightforward to execute in practice and to reduce specific data requirements.

Appendix 2

Portfolio Analysis of Long-Term Payout Behaviour



In this portfolio analysis, we clustered the sample firms based on the total payout behaviour in the respective year. For the total payout ratio ($\mathbf{k_0}$), we used the sum of cash dividends paid and share repurchases, divided by current earnings. We use a two-step classification approach to create the portfolios. As a first step, we classified firms into three categories: non-payers ($\mathbf{k_0 = 0}$%), payers ($\mathbf{k_0} \in (\mathbf{0}; \mathbf{100}$%)) and overpayers ($\mathbf{k_0} \geq \mathbf{100}$%). Overpayers are defined as firms that currently pay out more than they earn, so the category also includes loss firms that have non-zero payout ($\mathbf{k_0 < 0}$). In case a firm is classified as an overpayer, we set $\mathbf{k_0}$ to 100% (i.e. all overpayers start at 100% payout). This procedure reduces the impact of outliers, enables the inclusion of loss firms and restricts $\mathbf{k_0}$ to an economically plausible interval. After the classification into the three portfolios, we calculate means for the portfolios at the initial date and the five subsequent years after portfolio formation, for every year in the sample from 1998 to 2013. Finally, we calculate the mean values over the full sample period. The following figure includes the means for the three portfolios over the five years after portfolio formation. It is important to remark that the portfolio-based approach likely exhibits a stronger survivorship bias than ex-ante considerations. This has implications, especially for the overpayers and non-payers, because both of them may include loss firms that may have vanished from the market after the portfolio formation date.

Appendix 3

Supplementary Portfolio Statistics

Panel A: Forward Dividend Yield

	$\mathbf{DY_1}$	Cheapness (LTS/P) Low (> 70%)	Middle (in between)	High (< 30%)	Mean	L-H
Quality (VScore)	Low (0 or 1)	0.16%	1.09%	1.88%	1.09%	1.72%
	Middle (2 to 6)	0.92%	1.56%	2.50%	1.65%	1.58%
	High (7 or 8)	1.77%	2.26%	3.50%	2.22%	1.74%
	Mean	1.00%	1.64%	2.47%		
	H-L	1.60%	1.17%	1.62%		

t-Test for Undervalued > Overvalued

DY (Undervalued)	DY (Overvalued)	t-Stat	p-Value
2.64%	0.80%	7.88	< 0.0000

Panel B: Years Since Incorporation

	Age	**Cheapness** Low	Middle	High	Mean	L-H
Quality	Low	11.19	13.50	15.82	13.62	4.64
	Middle	17.09	20.92	19.43	19.51	2.33
	High	32.88	30.24	21.57	30.22	-11.32
	Mean	19.52	21.96	19.03		
	H-L	21.70	16.75	5.74		

t-Test for Junk < Quality

Age (Junk)	Age (Quality)	t-Stat	p-Value
13.62	30.22	3.05	< 0.0000

Panel C: Logarithmic Total Assets

	Size	**Cheapness** Low	Middle	High	Mean	L-H
Quality	Low	19.21	20.24	20.68	20.09	1.46
	Middle	20.51	21.03	21.56	21.04	1.05
	High	21.89	22.17	22.78	22.14	0.89
	Mean	20.63	21.17	21.51		
	H-L	2.68	1.93	2.10		

This table shows supplementary statistics for nine portfolios sorted based on (Quality) VScore and Cheapness (LTS/P) for 39,410 firm-years from 1998 to 2017. The outer rows and columns display statistics for portfolios sorted separately based on Quality and Cheapness and the difference between the high and low portfolios for each sort. Portfolios are formed three months after fiscal-year end. DY_1 is calculated as the sum of all dividends per share in the year after portfolio formation, scaled by the current price. Age is calculated as the number of (full) months since the incorporation of the firm, divided by twelve. Size is calculated as the natural logarithm of the most recently available total assets at portfolio formation date.

8. *References*

Abarbanell, J. & Bushee, B., 1998. Abnormal Returns to a Fundamental Analysis Strategy. *The Accounting Review*, 73, pp.19-45.

Ali, A., Hwang, L.-S. & Trombley, M.A., 2003. Residual-Income-Based Valuation Predicts Future Stock Returns: Evidence on Mispricing vs. Risk Explanations. *The Accounting Review*, 78(2), pp.377-96.

Arnott, R.D., Harvey, C.R., Kalesnik, V. & Linnainmaa, J.T., 2021. Reports of Value's Death May Be Greatly Exaggerated. *Financial Analysts Journal*, 77(1), pp.44-67.

Asness, C.S., Frazzini, A. & Pedersen, L.H., 2019. Quality minus junk. *Review of Accounting Studies*, 24, pp.34-112.

Balakrishnan, K., Bartov, E. & Faurel, L., 2010. Post Loss/Profit Announcement Drift. *Journal of Accounting and Economics*, 50, pp.20-41.

Ball, R., Kothari, S.P. & Shanken, J., 1995. Problems in measuring portfolio performance: An application to contrarian investment strategies. *Journal of Financial Economics*, 38, pp.79-107.

Bartov, E. & Kim, M., 2004. Risk, Mispricing, and Value Investing. *Review of Quantitative Finance and Accounting*, 23, pp.353-76.

Basu, S., 1977. Investment Performance of Common Stocks in Relation to Their Price-Earnings Ratios: A Test of the Efficient Market Hypothesis. *Journal of Finance*, 32(3), pp.129-56.

Beneish, M., Lee, C.M.C. & Tarpley, R., 2001. Contextual Financial Analysis Through the Prediction of Extreme Returns. *Review of Accounting Studies*, 6, pp.165-90.

Bradshaw, M., Drake, M.S., Myers, J.M. & Myers, L.A., 2012. A re-examination of analysts' superiority over time-series forecasts. *Review of Accounting Studies*, 17, pp.944-68.

Brav, A., Graham, J.R., Campbell, R.H. & Michaely, R., 2005. Payout policy in the 21st century. *Journal of Financial Economics*, 77, pp.483-527.

Campbell, J.Y., Hilscher, J. & Szilagi, J., 2008. In Search of Distress Risk. *Journal of Finance*, 63(6), pp.2899 - 2939.

Campbell, J.Y., Polk, C. & Vuolteenaho, T., 2010. Growth or Glamour? Fundamentals and Systematic Risk in Stock Returns. *The Review of Financial Studies*, 13(1), pp.305-44.

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

Cohen, R.B., Polk, C. & Vuolteenaho, T., 2009. The Price Is (Almost) Right. *Journal of Finance*, 64(6), pp.2739-82.

DeBondt, W.F.M. & Thaler, R., 1985. Does the Stock Market Overreact?. *Journal of Finance*, 40(3), pp.793-805.

Dichev, I. & Piotroski, J., 1999. The Performance of Long-run Stock Returns Following Issues of Public and Private Debt. *Journal of Business, Finance, and Accounting*, 26, pp.1103-32.

Dichev, I. & Piotroski, J., 2001. The Long-run Stock Returns Following Bond Rating Changes. *Journal of Finance*, 56, pp.173-203.

Doukas, J., Kim, C. & Pantzalis, C., 2002. A Test of the Errors-in-expectations Explanation for the Value/Glamour Stock Return Performance: Evidence from Analyst Forecasts. *Journal of Finance*, 57, pp.2143-65.

Doyle, J., Lundholm, R. & Soliman, M., 2003. The Predictive Value of Expenses Excluded from "Pro Forma" Earnings. *Review of Accounting Studies*, 8, pp.145-74.

Doyle, J., Lundholm, R. & Soliman, M., 2006. The Extreme Future Returns Following Extreme Earnings Surprises. *Journal of Accounting Research*, 44, pp.849-88.

Easton, P.D., Harris, T.S. & Ohlson, J.A., 1992. Aggregate accounting earnings can explain most security returns: The case of long return intervals*. *Journal of Accounting and Economics*, 15, pp.119-42.

Fairfield, P.M., 1994. P/E, P/B and the Present Value of Future Dividends. *Financial Analysts Journal*, 50(4), pp.23-31.

Fairfield, P.S. & Yohn, T.L., 2001. Using Asset Turnover and Profit Margin to Forecast Changes in Profitability. *Review of Accounting Studies*, 6, pp.371 - 385.

Fama, E.F., 1970. Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, 25(2), pp.383-417.

Fama, E.F. & French, K.R., 1992. The Cross-Section of Expected Stock Returns. *The Journal of Finance*, 47(2), pp.427-65.

Fama, E.F. & French, K.R., 1995. Size and Book-to-Market Factors in Earnings and Returns. *The Journal of Finance*, 50(1), pp.131-55.

Fama, E.F. & Macbeth, J.D., 1973. Risk, Return, and Equilibrium: Empirical Tests. *Journal of Political Economy*, 81(3), pp.6ß7-36.

Feltham, G.A. & Ohlson, J.A., 1995. Valuation and Clean Surplus Accounting for Operating and Financial Activities. *Contemporary Accounting Research*, 11(2), pp.689-731.

Frankel, R. & Lee, C.M.C., 1998. Accounting valuation, market expectation, and cross-sectional returns. *Journal of Accounting and Economics*, 25, pp.283-319.

Graham, B. & Dodd, D., 1934. *Security Analysis: The Classic 1934 Edition*. New York: McGrawHill.

Grossman, S.J. & Stiglitz, J.E., 1980. On the Impossibility of Informationally Efficient Markets. *The American Economic Review*, 70(3), pp.393-408.

Huefner, B. & Rueenaufer, M., 2020. A Fundamentalist Perspective on Reverse Engineering in Equity Valuation. *SSRN Working Paper*.

Ikenberry, D., Rankine, G. & Stice, E., 1996. What do Stock Splits Really Signal? *Journal of Financial and Quantitative Analysis*, 31, pp.357-77.

Israel, R., Laursen, K. & Richardson, S., 2020. Is (systematic) value investing dead? *SSRN Working Paper*.

Jegadeesh, N. & Livnat, J., 2006. Revenue Surprises and Stock Returns. *Journal of Accounting and Economics*, 41, pp.141-71.

Jegadeesh, N. & Titman, S., 1993. Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *Journal of Finance*, 48(1), pp.65-91.

Jiang, G., Lee, C.M.C. & Zhang, Y., 2005. Information Uncertainty and Expected Returns. *Review of Accounting Studies*, 10, pp.185-221.

La Porta, R., 1996. Expectations and the Cross-Section of Stock Returns. *Journal of Finance*, 51(5), pp.1715-42.

La Porta, R., Lakonishok, J., Shleifer, A. & Vishny, R., 1997. Good News for Value Stocks: Further Evidence on Market Efficiency. *Journal of Finance*, 52(2), pp.859-74.

Lakonishok, J., Shleifer, A. & Vishny, R., 1994. Contrarian Investment, Extrapolation, and Risk. *Journal of FInance*, 49(5), pp.1541-78.

Lee, C.M.C., 2015. Value Investing: Bridging Theory and Practice. *China Accounting and Finance Review*, 16(2), pp.10-38.

Li, K.K. & Mohanram, P., 2019. Fundamental Analysis: Combining the Search for Quality with the Search for Value. *Contemporary Accounting Research*, 0(0), pp.1-36.

Mohanram, P.S., 2005. Separating Winners from Losers among Low Book-to-Market Stocks using Financial Statement Analysis. *Review of Accounting Studies*, 10, pp.133-70.

Newey, W.K. & West, K.D., 1987. A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*, 55(3), pp.703-08.

Nissim, D. & Penman, S.H., 2001. Ratio Analysis and Equity Valuation: From Research to Practice. *Review of Accounting Studies*, 6, pp.109-54.

Novy-Marx, R., 2013. The other side of value: The gross profitability premium. *Journal of Financial Economics*, 108(1), pp.1-28.

Ohlson, J.A., 1995. Earnings, Book Values, and Dividends in Equity Valuation. *Contemporary Accounting Reseach*, 11(2), pp.661-87.

Ohlson, J.A. & Juettner-Nauroth, B., 2005. Expected EPS and EPS Growth as Determinants of Value. *Review of Accounting Studies*, 10, pp.349-65.

Ou, J. & Penman, S.H., 1989. Financial Statement Analysis and the Prediction of Stock Returns. *Journal of Accounting and Economics*, 11, pp.295-329.

Palepu, K.G., Healy, P.M. & Peek, E., 2019. *Business Analysis and Valuation: IFRS Edition*. 5th ed. Andover: Cengage Learning.

Penman, S.H., 1996. The Articulation of Price-Earnings Ratios and Market-to-Book Ratios and the Evaluation of Growth. *Journal of Accounting Research*, 34(2), pp.235-59.

Penman, S.H., 2011. *Accounting for Value*. New York: Columbia University Press.

Penman, S.H., 2016. Valuation: Accounting for Risk and the Expected Return. *ABACUS*, 52(1), pp.106-30.

Penman, S.H. & Reggiani, F., 2013. Returns to buying earnings and book value: accounting for growth and risk. *Review of Accounting Studies*, 18, pp.1021-49.

Penman, S.H. & Reggiani, F., 2013. Returns to buying earnings and book value: accounting for growth and risk. *Review of Accounting Studies*, 18, pp.1021-49.

Penman, S.H. & Reggiani, F., 2018. Fundamentals of Value vs. Growth Investing and an Explanation for the Value Trap. *Financial Analysts Journal*, 74(4), pp.103-19.

Penman, S.H. & Yehuda, N., 2019. A Matter of Principle: Accounting Reports Convey Both Cash-Flow News and Discount-Rate News. *Management Science*, 62(12), pp.5449-956.

Penman, S.H., Zhu, J. & Wang, H., 2019. The Implied Cost of Capital: Accounting for Growth. *SSRN Working Paper*.

Petkova, R. & Zhang, L., 2005. Is Value riskier than Growth? *Journal of Financial Economics*, 78(1), pp.187-202.

Piotroski, J.D., 2000. Value Investing: The Use of Historical Financial Statement Information to Seperate Winners from Losers. *Journal of Accounting Research*, 38, pp.1-41.

Piotroski, J.D. & So, E.C., 2012. Identifying Expectation Errors in Value/Glamour Strategies: A Fundamental Analysis Approach. *Review of Financial Studies*, 25(9), pp.2841-75.

Shroff, P.K., 1995. Determinants of the Returns-Earnings Correlation. *Contemporary Accounting Research*, 12(1), pp.41-55.

Skogsvik, K. & Juettner-Nauroth, B., 2013. A note on accounting conservatism in residual income and abnormal earnings growth equity valuation. *The British Accounting Review*, 45, pp.70-80.

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

Sloan, R.G., 2019. Fundamental Analysis Redux. *The Accounting Review*, 94(2), pp.363-77.

Spiess, D. & Affleck-Graves, J., 1995. Underperformance in Long-run Stock Returns Following Seasoned Equity Offerings. *Journal of Financial Economics*, 38, pp.243-67.

Turtle, H.J. & Wang, K., 2017. The Value in Fundamental Accounting Information. *Journal of Financial Research*, 40(1), pp.113-40.

Walkshäusl, C., 2020. Piotroski's FSCORE: international evidence. *Journal of Asset Management*, 21, pp.106-18.

Zhang, X.-J., 2000. Conservative accounting and equity valuation. *Journal of Accounting and Economics*, 29, pp.125-49.

Zhang, X.-F., 2006. Information Uncertainty and Stock Returns. *Journal of Finance*, 61(1), pp.105-37.