The Impact of Inter-Firm Heterogeneity on Regression-Based Valuation Models

A Comparative Study of GICS Segmentations and the SARD Approach in US Equity Markets

Master's Thesis

Master of Science in Finance and Strategic Management

Copenhagen Business School

Authors	**Supervisor**
Adam Erik Hjelmström (128859)	Jens Borges
Espen Ælmose Hessner (129357)	

No. of characters: 242,450

No. of pages:119

Hand-in date: 14.05.2023

Abstract

Within peer groups chosen for relative valuation, inter-firm heterogeneity is typically addressed by employing measures of central tendency, such as the average multiple. Limited research has been devoted to handling such heterogeneity through regression-based prediction models for valuation multiples – and the impact of heterogeneity on the prediction accuracy of linear models when compared to peer group averages. We explore how model characteristics change when varying peer groups from market to GICS sector groups, and GICS industry groups, and when applying the Sum of Absolute Rank Difference (SARD) approach for peer selection. Furthermore, we seek to narrow the research gap of the SARD approach within the context of linear regression models, and how data subsets sorted by the SARD approach handle the MLR.1-6 assumptions. In doing this, we will explore the degree to which variance in our selected underlying value drivers, growth, profitability, and risk, can significantly explain variances in EV/EBITDA multiples, across different peer segmentations.

Our dataset was comprised of observed EV/EBITDA, EBITDA CAGR, ROIC and WACC figures for a trimmed census of 929 companies from the S&P 1500 Composite Index, which throughout most of the segmentations adhered to the MLR.1-6 assumptions – but with indications of an omitted variable bias. Through running Simple Linear Regressions (SLR), we found there to be varying significant beta coefficients across the different segmentations; mostly positive for growth, close to zero for profitability, and surprisingly – positive for risk. We found the underlying value drivers to have joint significance in around half of the segmentations, with slightly better goodness-of-fit statistics for GICS segmentations than SARD groupings. However, we found that heteroskedasticity in error terms successively decreased, when moving from an aggregate market level, to sector, to industry, to SARD groupings. Relative prediction accuracy followed the same pattern, with SARD groupings having slightly smaller relative prediction errors. SARD groupings also had slightly fewer cases of non-normality and heteroskedasticity, we deemed that the SARD groupings were overall less susceptible to systemic bias and overfitting. Our models cannot fully explain the relationship between selected underlying value drivers and the EBITDA-multiple and are mostly not more accurate than peer group averages. We find that fewer heteroskedastic error terms seem to improve prediction accuracy and that SARD groupings can potentially handle MLR.5-6 assumptions better than GICS segmentations and can be tweaked through changing SARD selection criteria.

Contents

Chapter 1

Introduction

1.1 Motivation

Equity investments by rational investors are made with the expectation that the calculated "true" value of equity differs from the current market value to a degree where the payoff from dividends or a future divestment, sufficiently covers the risk assumed by the investor (Kinserdal, Petersen, & Plenborg, 2017). The total market capitalisation of only the 1,500 largest U.S. publicly listed companies as of 16.03.2023, was USD ~39t (Bloomberg Terminal, 2023). In this context, corporate valuation serves as a foundation for decisions on directing capital, which impacts economies and societies as such. Therefore, corporate valuation constitutes a need to understand firms' underlying value drivers.

While there are different approaches to corporate valuation, the most popular methods used by industry practitioners on a going concern basis[1] are absolute valuations in the form of discounted cash flow (DCF) analyses, and relative valuation in the form of comparable companies analyses and precedent transactions (M&A analysis) (Damodaran A., 2020). A study of 1,980 equity analyst members of the Chartered Financial Analysts (CFA) Institute found that 92.8% of finance professionals actively used the market multiples approach (Pinto, Robinson, & Stowe, 2015). In another study, Asquith et al. (2005) found that ~99% of published analyst reports include relative valuation, whereas only 12.8% use present value approaches. Relative valuation is a method of determining the value of a firm relative to similar firms (peers) operating in the same industry. The approach involves analysing a company's multiples and metrics relative to industry peers. It is popular among practitioners due to its low level of complexity and the speed by which a valuation can be performed (Kinserdal, Petersen, & Plenborg, 2017). A multiple is a numerical metric

[1] Going concern assumption in accounting assumes that a reporting entity continues operations for the foreseeable future, as opposed to liquidating assets (Kinserdal, Petersen, & Plenborg, 2017)

expressed as a ratio which can be conceptualised as a quotient of a firm's chosen value metric, such as the total enterprise value, and a relevant fundamental metric such as earnings or revenue. Accordingly, by comparing the chosen value metric of comparable firms or precedent corporate transactions as a benchmark, the value of a firm can be inferred (Kinserdal, Petersen, & Plenborg, 2017).

When applying the relative valuation approach, there are typically two assumptions that these predictions are based on. Firstly, value is linked to the fundamental metric used, such as earnings, revenue, cash flow, or book value. Secondly, similar proportionality applies to comparable companies, which are firms in the same industry or with similar characteristics. (Gupta, 2018)

Albeit the relative valuation approach is seemingly easy to understand and utilise, it is subject to pitfalls. One of the most salient concerns within the field of relative valuation is the question of how to apply and utilise multiples for valuation purposes in such a manner as to produce reliable, efficient, and accurate valuation outcomes. These issues related to implementation encompass but are not restricted to the selection of comparable firms, the utilisation of reported earnings versus projected earnings, and defining the most appropriate method for determining averages (Plenborg & Pimentel, 2016).

The fundamental elements of multiple valuation and the corresponding challenges related to its implementation constitute the primary areas of scholarly inquiry within the field of corporate valuation. Given there is limited academic literature on the topic, we are motivated to determine whether it is possible to develop an objective model with a high degree of valuation accuracy based on statistical regression, with fundamental value drivers as the determinants. The underlying value drivers, growth, profitability, and risk are deemed to be the fundamental determinant of firm value (Bhojraj & Lee, 2002; Plenborg & Pimentel, 2016; Overgaard Knudsen, Kold, & Plenborg, 2017). Hence this paper will investigate to what extent these three fundamental underlying value drivers can predict firm value by applying single and multi-linear regression models. Furthermore, the study aims to test the accuracy of the predictive models on different peer groups, including the market level, composed as a trimmed census of the S&P 1500 composite index, GICS sector level, GICS industry level, and the SARD approach with groupings similar sized as GICS industry level.

The aim of utilising the Sum of Absolute Rank Differences (SARD)[2] method for segmentation is to investigate whether augmented homogeneity in the data sample could result in amplified homoskedasticity in the residuals, and subsequently to determine if the augmented homoskedasticity in the residuals could enhance the prediction accuracy of the multivariate regression model.

1.2 Research Questions

Grounded in financial and econometric theory and literature, the research objective of this paper is to develop linear regression models with key value drivers as determinants (independent variables) of firm value, to produce accurate EV/EBITDA (dependent variable) valuation estimates, whilst exploring the impact of different peer groups. All this whilst minimising subjective adjustments.

The outlined research question formulations aim to systematically identify and investigate the strength of the relationship between fundamental value drivers and the studied multiple, evaluate prediction accuracy in isolation and relative to observed market multiples – and compare predictive power from derived models by regressing key value drivers across our chosen peer group segmentations. Find the following two research questions guiding our quantitative study.

- **Research Question 1**: *To which degree can variance in selected proxies significantly explain variances in EV/EBITDA multiple valuations for US publicly listed firms within the S&P 1500 composite index, when running linear regression models?*

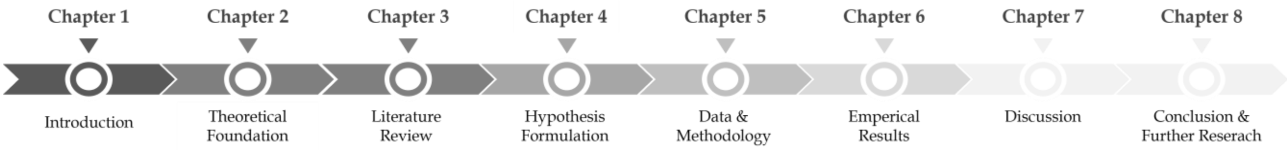
- **Research question 2**: *To what extent can OLS regression models utilising selected proxies, segmented by GICS codes and the SARD approach, accurately predict EV/EBITDA multiple valuations in congruence with observed multiple valuations?*

Guided by research questions 1 and 2, the paper is divided into 8 chapters which collectively form the elements of the deductive research approach followed in this study. **Chapter 1** addresses the motivation behind our quantitative study and research questions, with delimitations governing the scope. Chapters **2** and **3** outline the theoretical foundation and literature review that the paper draws

[2] The SARD approach is theoretically explained in **section 2.4** and emperically contextualised in **section 3.4**

upon when formulating the hypotheses in **Chapter 4**. **Chapter 5** presents the data & methodology applied in the study, including the research paradigm, mathematical derivation of the regression variables, research model, data collection and method of data analysis. **Chapter 6** presents the empirical results of the study, which works as the foundation for **Chapters 7** and **8** where novelty, limitations, generalisability, and further research will be discussed. See **Figure 1.1** for a graphical illustration of the paper structure.

Figure 1.1 – Research Approach & Structure of the Paper



1.3 Delimitations

The stated research questions of this study serve as the guide for the delimitation in terms of research focus, theoretical framework, literature review, and research methodology. The subsequent section will provide a detailed description of the specific areas that were excluded from the investigation of the research questions. The rationale for the delimitation made within the choice of sample size, data quality, model specification, and time period will also be outlined.

Starting with the literature review, the study is delimited to existing literature and empirical research in the overarching field of relative valuation, where the primary focus is on the accuracy of multiples in predicting implied firm value. Whilst the theoretical foundation of the study is also delimited to cover relative valuation, the study will still account for the theoretical relationship between relative- and absolute valuation. Furthermore, in line with the identification of comparable firms, the theoretical foundation is delimited to the GICS segmentation approach and the SARD segmentation approach.

In terms of research methodology, the constructed model is delimited regarding the choice of the dependent- and independent variable(s) and which type of regression models we use. Whilst there are several other enterprise- and equity-value multiples, which could have been analysed, the scope of this paper is delimited to EV/EBITDA, motivated by gaining an in-depth understanding of this multiple alone. Furthermore, the delimitation of testing EV/EBITDA as the dependent variable is theoretically and empirically grounded, due to its generalisable and comparable characteristics. EV/EBITDA is also less sensitive to differences in accounting policies, and capital structures, as compared to equity multiples[3].

Prior studies have supported the notion that predicting the value of a firm using fundamental value drivers such as growth, profitability, and risk is well-established, both in theoretical and empirical contexts. As such our study is delimited to only concern these value drivers as independent variables. Additional value drivers could have been included, such as operational efficiency and financial leverage. However, the scope of this study is, specifically, to test the relationship between EV/EBITDA and its fundamental value drivers. As for proxies of the value drivers, this analysis will be delimited to include one proxy for each value driver. More specifically, EBITDA CAGR will be used as a proxy for growth, ROIC will be used as a proxy for profitability, and WACC will be used as a proxy for risk. These proxies are supported by the theoretical derivation of EV/EBITDA as per Plenborg et al. (2017)[4], and an extensive amount of empirical research[5]. Whilst one might argue that including more independent variables should improve the accuracy of the models, this would also increase the complexity and the potential for error. E.g., if including additional proxies for the value drivers, these proxies could potentially possess a high degree of multicollinearity with the other fundamental value driver which subsequently could impact the interpretability of the result yielded by the model. This follows Damodaran's (2012) notion of parsimony with regard to variable(s) selection in regression models within a financial market's context.

Furthermore, the research methodology is delimited to the statistical model applied in the data methodology. More specifically the OLS regression will constitute the fundamental approach of data analysis, where the input is based on aggregated cross-sectional data. This delimitation is theoretically founded by a high degree of interpretability of the model's output. In addition, this delimitation is empirically substantiated in prior studies conducted on relative valuation[6]. Whilst other statistical methods could be applied in addition to the OLS regression model, it is argued that the scope of our study is not to compare statistical significance across models, but rather to test the

[3] See **section 2.1.2** & **3.2.1**

[4] See **section 2.2.2**

[5] See **section 3.1**

[6] See **section 3.2.3**

underlying relationship between EV/EBITDA and theoretically derived value drivers for different peer group segmentations.

The study is also delimited in terms of sample size, data quality and time period. The data sample is a trimmed census S&P 1500 composite index, consisting of the 1500 largest publicly listed firms in the United States. I.e. the sample is delimited to public firms in one single country. This is motivated by an increased degree of homogeneity achieved when gathering data from one single market. Due to different countries possessing different accounting policies and other market-specific differences, a cross-geography data sample would not be completely comparable. Furthermore, in terms of the time period, the analysis was initially delimited to generate output for 2022 based on 2020-2022 data. However, this delimitation was eased, and four additional years of data were gathered, for testing robustness in the EV/EBITDA-fundamental value driver relationship over time. Lastly, in terms of data quality, the data sample, including all proxies for the value driver, was gathered from a Bloomberg Terminal, which is arguably one of the largest and most reliable sources of financial data for finance professionals. See **Table 1.1** below for a summary of delimitations in the research methodology.

Table 1.1 – Delimitation Overview

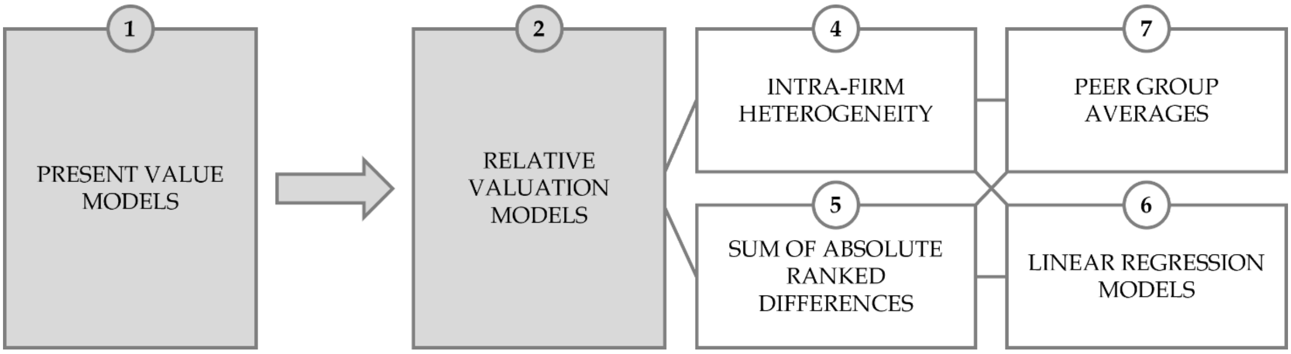
Delimiation 1: Model specification	Delimitation 2: Peer group selection	Delimitation 3: Data sample	Delimitation 4: Time period	Delimitation 5: Data source
Dependent variables:	**Peer group selection method:**	**Sample:**	**Time period:**	**Data Source:**
EV/EBITDA multiple **Independent variables:** 1. Growth (EBITDA CAGR); 2. Profitability (ROIC); 3. Risk (WACC) **Statistical model:** Simple linear regression (SLR) Multiple linear regression (MLR)	2- and 4-digit GICS segmentation Sum of Absolute Squared Differences (SARD) approach*	Trimmed census S&P 1500 Composite Index (US)	Full year (FY) 2020-2022	Bloomberg Terminal (2022)

*The mechanics of the SARD approach is elborated upon in section 2.3 and 3.4

Chapter 2
Theoretical Foundations

As this paper presents quantitative research with a deductive approach, it is essential to align the reader with the theoretical underpinnings of the research conducted. Hence, this chapter will in **section 2.1.1** describe the absolute valuation approach, to establish the foundation for how theoretical underlying value drivers affect firms' market capitalisations, as well as limitations to the approach. Successively, **section 2.1.2** will cover the relative valuation approach and its limitations. To further elaborate on the relationship between enterprise value and our suggested theoretical underlying value drivers. **Section 2.2.1** will cover derivations of the EBITDA multiple from said value drivers. **Section 2.3** touches upon inter-firm heterogeneity. Furthermore, **section 2.4** will outline the theoretical method underlying the SARD approach. Finally, **section 2.5** will outline the basic fundaments of linear regression.

Figure 2.1 – Structure of Theoretical Foundations



2.1 Corporate Valuation

Damodaran (2009) suggests that corporate valuation aims to estimate the value of a company that reflects its underlying fundamentals. There are different approaches to corporate valuations, where Plenborg et al. (2017) segment these models into 4 distinct categories:

- **Present value models**, which are based on the principle of the time value of money[7], sum the discounted future expected cash flows to arrive at a current intrinsic value.
- **Relative valuation models**, which are based on the principle of the law of one price[8], estimate firm value by comparing relevant performance metrics against a comparable peer group – followingly deriving a fair value against their market capitalisation.
- **Asset-based models**, calculate firm value as the sum of the value of assets and liabilities with different measurement bases, such as net asset value (NAV), sum-of-parts, and liquidation value.
- **Contingent claim valuation models** are relevant for companies with projects that can be seen as real options, where option pricing models can be applied to calculate firm value.

Adhering to the delimitations of the paper, the last two approaches to valuation will not be further elaborated upon.

2.1.1 Present Value Models

In the context of corporate valuation, present value models are designed to calculate firms' intrinsic enterprise value (EV), or intrinsic value of equity (Kinserdal, Petersen, & Plenborg, 2017). The market price differs from intrinsic value, in that it is simply the sum of outstanding shares times the price which stockholders are willing to pay for the security (Graham & Dodd, 1934).

Calculating free cash flow to the firm starts with operating profits after taxes (NOPAT), adjusts for non-cash items, changes in net working capital (NWC) and capital expenditure (CAPEX) (Kinserdal, Petersen, & Plenborg, 2017). In this way, the cash flow statement is linked to the income statement and balance sheet, and present value models must make assumptions on the future development of relevant financial items. Forecasting can either be approached top-down or bottom-up[9]. This illustrates how a corporate valuation can be derived from the underlying value-drivers profit,

[7] Time value of money suggests that money received today is worth more than money received in the future, as money can be invested

[8] Law of one price states that equivalent investment opportunities that trade simultaneously in different markets, must trade for same price in all markets (Berk & DeMarzo, 2019)

[9] Top-down approach entails looking at market-wide factors, and the firms performance relative to peers. Bottom-up starts by looking at internal factors such as historical development in financial items

growth, and risk. An absolute valuation, in its simplest form, can be expressed by the following equation.

Equation 2.1

$$Value_0 = \sum_{t=1}^{t=n} \frac{Net\ cash\ flow_t}{(1+r)^t}$$

Where…
$Net\ cash\ flow_t$ = Net difference between cash inflows and outflows in period t
r = risk-adjusted discount rate or cost of capital

The previous equation holds for both assets and corporations, alike. When calculating the intrinsic enterprise value of a firm, the following equation which sums together the discounted expected FCFF for different periods, can be used.

Equation 2.2

$$Enterprise\ value_0 = \sum_{t=1}^{t=n} \frac{FCFF_t}{(1+WACC)^t}$$

Where…
$FCFF_t$ = Free cash flow to the firm in period t
$WACC$= Weighted average cost of capital

FCFF is excess cash that can be distributed to investors and creditors (Kinserdal, Petersen, & Plenborg, 2017). Applying FCFF as a numerator constitutes using a discount rate that reflects the risk borne by both equity holders and debtholders, e.g., the WACC. The following equation can be used to calculate the intrinsic value of equity.

Equation 2.3

$$Value_0 = \sum_{t=1}^{t=n} \frac{FCFE_t}{(1+r_e)^t}$$

Where…

$FCFE_t$ = Free cash flow to equity holders in period t

r_e = required return on equity

FCFE differs from FCFF in that it deducts increases in net interest-bearing liabilities and net financial expenses after tax (Kinserdal, Petersen, & Plenborg, 2017). Hence, the risk factor should be adjusted accordingly. Both the discounted cash flow models (equations 2.1 and 2.2) assume that there is no future cash build-up in the firm and that the numerators grow from increases in income from operating assets (Damodaran A., 2012).

Shortcomings of the present value approach

As evident from the above DCF formulas, cash flows are discounted more the further away periods are from $t = 0$. This means the model is highly sensitive to changes in the discount rate, which should reflect the risk borne by residual claimants; hence requiring the availability of suitable proxies for a firm's idiosyncratic risk (Berk & DeMarzo, 2019).

Unstable- and/or currently negative cash flows complicate the process of forecasting financial items, given that subsequent forecasting years with decreased degrees of certainty, use current data as a starting point. Analysts often forecast financial items on a pro-forma basis[10] (Kinserdal, Petersen, & Plenborg, 2017), meaning assumptions needs to be made on e.g., a firm's top-line growth. In this process, projections are typically smoothed out without attempting to time recessions and recoveries - in which analysts' predictions may be biased by current economic standings (Damodaran A. , 2012). An example could be having overly negative predictions when a cyclical firm is struggling during a recession. Damodaran (2012) further points out the lower applicability of DCF models in certain situations, such as distressed firms requiring fire sales[11], firms undergoing restructurings (future uncertainty), and M&A requiring estimating (uncertain) synergy effects on cash flows.

The DCF analysis is viewed by many investors as the (theoretically) most correct valuation method, with (Buffet, 1997) arguing that a comprehensive DCF analysis will yield an intrinsic value less influenced by market sentiments than other valuation methods and is more accurate in that it

[10] Pro-forma forecasting entails forecasting financial items as %-of other items, typically revenue

[11] A fire sale involves selling off assets at discounts to repay claimants at a short notice

estimates the effect of taxes, financing decisions and capital expenditures. Furthermore (Koller, Goedhart, & Wessels, 2010) argues for increased validity when basing estimates on cash-in and -out, instead of accounting-based earnings. However, as argued by Damodaran (2012), the practical application may dwindle the more a firm's fundamentals vary from an "idealised" framework (e.g., unstable, and negative cash flows), as uncertainty increases, and more assumptions must be drawn.

DCF models' sensitivity to underlying assumptions has been well documented. (Steiger, 2010) performed a two-stage[12] DCF model sensitivity analysis with model inputs from Credit Suisse Equity Research, on a DAX 40 company. Findings showed that simultaneously decreasing the perpetual growth rate by 50bp and increasing the WACC by 100bp, decreased the calculated fair stock price by ~19%. Increasing the CAGR in the forecasting period by 25bp drove the fair share price up by 1.5%, with 3% by changing the perpetual growth rate. DCF models can be highly sensitive to even marginal changes from analysts in model inputs for estimates relatively far ahead in time, illustrating the potential for error. When performed by industry professionals, the fair value is typically expressed as a range and compared in a "football field chart" to the fair value derived from other valuation methods, such as the relative valuation (Pearl & Rosenbaum, 2009).

2.1.2 Relative Valuation Models

The theoretical focus of academic discussion on corporate valuation has predominantly been on absolute valuation approaches, however, in practice, industry professionals tend to use relative valuation (Kinserdal, Petersen, & Plenborg, 2017). Whilst the absolute valuation approach seeks to derive the intrinsic firm value based on individual firm fundamentals, without relative considerations of industry peers, the relative valuation approach derives firm value based on a comparison between firms (Damodaran A., 2012). Therefore, there is a clear relationship between the two approaches, where the relative valuation approach is drawn upon the absolute valuation approach (Bernström, 2014).

As stated by Sharma & Prashar (2013) relative valuation is the valuation of any assets based on how similar the same assets are priced in the market, by using indicators such as enterprise multiples or equity multiples. Furthermore, Sharma & Prashar (2013) provides a conceptual framework for

[12] Two-stage DCF analyses are comprised of a forecasting period (e.g. 5-10 yrs), and a subsequent perpetuity

relative valuation. They argue that the approach should be conducted in the following way: select the target company; create a list of comparable companies; obtain key financial information; find ratios, business statistics, and trading multiples; benchmark comparable companies; and determine valuation. Rosenbaum & Pearl (2009) suggests a similar approach whilst further emphasising the importance of finding comparable companies based on the business profile and the financial profile. They explain how valuation is driven based on historical performance and expected future performance.

Furthermore, when finding ratios, business statistics and trading multiples, it is important to allow for comparison across firms based on their performance and not only the relative size of the company (Sharma & Prashar, 2013). To determine the relative market performance, relative valuation takes the form of a multiple, which is a fractional expression of a firm's market value relative to its key financial statistics. There is an important connection between the numerator and the denominator, thus the denominator should be a determinant of the numerator. When this is the case, the multiple will be able to capture the main value drivers behind the firm valuation (Damodaran A., 2012). Subsequently, numerous studies have been conducted to evaluate the most suitable accounting variable to be used as a measure of scale.[13]

Valuation multiples for comparable firms serve as the central base for deriving an implied and appropriate valuation range of the target firm. The methodology of benchmarking entails conducting a thorough evaluation and comparison of comparable firms with a selected target, to determine the relative positioning of the target to frame the valuation accordingly (Pearl & Rosenbaum, 2009). Given, the methodology outlined above, one can, in a simplified way, calculate the value of a target firm as per below.

Equation 2.4

$$Implied\ Firm\ Value\ of\ Target\ = Peers'\ Mean\ Multiple * Target's\ Selected\ Financial\ Statistic$$

There are two basic types of multiples: enterprise value- and equity value multiples. Enterprise value multiples express the value of an entire enterprise relative to a financial statistic that relates to the entire enterprise value. Enterprise value multiples derive the value entitled by both debt- and

[13] The academic studies devoted to this area of research will be elaborated upon in **section 3**

equity-holders (Pearl & Rosenbaum, 2009). By contrast, equity multiples express the value to which only equity holders are entitled. Hence equity multiples derive the value of this claim relative to a financial statistic which applies to equity holders only. (UBS, 2001)

In terms of applicability, both groups of multiples have advantages and disadvantages. Equity multiples are deemed to be more familiar to investors and more reliable than enterprise multiples since estimating enterprise value involves subjectivity when pricing non-core assets. However, equity multiples are more sensitive to differences in accounting policies, compared to enterprise value multiples (UBS, 2001). More explicitly, equity multiples are derived from accrual-based value drivers, such as earnings, net income, and the book value of equity. The reliability of these value drivers is diminished due to the inconsistency of the allocation procedures and the selectivity of the accounting methods and estimations. Given that valuation multiples are comparative metrics, it is important to adjust for heterogeneity and increase the level of comparability (Koller, Goedhart, & Wessels, 2010). Furthermore, enterprise value multiples avoid the influence of capital structure, they are easier to apply to cash flow, and enable the user to exclude non-core assets (UBS, 2001).

While the relative valuation approach sidesteps the need for direct projections and calculations of present value, it is based on the same fundamental principles that guide the more thorough present value approach. Specifically, the value of an investment is proportional to its expected future returns and inversely proportional to the level of risk involved. Since relative valuation relies on the concept of firm comparability, it builds on the underlying assumption that markets must be efficient (Liu , Nissim, & Thomas, 2002). Meaning that securities markets are extremely efficient in reflecting information on individual stocks and the stock market as a whole (Maliel, 2003).

Shortcomings of the relative valuation approach

As described earlier in this section, the appeal of using relative valuation for valuing a firm lies in its straightforwardness and ease of comprehension. The relative valuation approach can be used to generate approximate valuations expeditiously for companies and assets and are especially beneficial when there are many comparable companies traded on the stock exchange, and the market is, typically, pricing these firms accurately.

However, it can be more of a challenge to use the relative valuation approach to assess distinct firms with no discernible analogues, with limited or no revenues, and with negative profits (Damodaran

A., 2012). The academic discourse on valuation unanimously stresses that recognising comparable companies – often referred to as peers – is a critical component of successful multiple valuations since divergent companies can produce biased and inaccurate valuation estimates (Plenborg & Pimentel, 2016). Therefore, an analyst with biases can select a set of similar companies to reinforce their preconceived notions about a firm's value (Damodaran A. , 2012). To identify firms as truly comparable, they must have identical financial and strategic profiles with identical cash flow streams. However, this requires the practitioner to conduct an analysis based on cash flow projections, which defeats the purpose of using relative valuation as a shortcut compared to an absolute valuation (Soffer & Soffer, 2003).

There seems to be an academic consensus that variation in firms' accounting policies is a factor influencing comparability (Beaver & Morse, 1978; Penman S. H., 2007). Young and Zeng (2015) discussed the implications of differences in accounting practices, and how they can make similar firms appear different and different firms appear similar. The result of different accounting practices is thus that the peer selection and valuation output can be biased. Therefore, it is essential to compare firms that adopt the same accounting practices, with recognition, measurement and classification of accounting items being done similarly across the comparable firms (Plenborg & Pimentel, 2016). Whilst there have been several studies conducted in the field of peer group selection (Alford, 1992; Bhojraj & Lee, 2002; Lee, Ma & Wang, 2015), there is lacking literature on the topic of how to deal with firm differences in regression-based valuation. One of the central objectives of this study is trying to fill this research gap via an empirical study and subsequent discussions on the potential implications of our findings.

2.2 Underlying Value Drivers

2.2.1 Mathematical Derivation of EV/EBITDA Multiple

In alignment with the designated emphasis on only utilising the enterprise multiple EV/EBITDA, this segment will demonstrate the intrinsic derivation of the EV/EBITDA multiple to exhibit the mathematical relationship between the dependent variable, firm value, to the underlying value drivers: growth, profitability and. Guided by the framework set up by Kinserdal et al. (2017), this section seeks to mathematically express, with a thorough decomposition, the relationship between absolute and relative valuation and how the chosen underlying value drivers are embedded in the EV/EBITDA multiple.

The enterprise value of a firm can be mathematically derived using the absolute valuation approach, specifically by employing a DCF model and assuming a constant growth rate, as presented in equation 2.5 below:

Equation 2.5

$$EV = \frac{FCFF}{(1 + WACC)}$$

Where,

$$WACC = \frac{E}{E + D} * R_E + \frac{D}{E + D} * R_D * (1 - T)$$

FCFF represents the cash flow available to all capital providers, both equity and debt, after accounting for all operating expenses, taxes, investments in working capital and fixed assets. FCFF can be decomposed, and the enterprise value can be expressed more comprehensively as below:

Equation 2.6

$$EV = \frac{NOPAT + D\&A - CAPEX - \Delta NWC}{(1 + WACC)}$$

Simply put, the free cash flow is determined by what the firm earns minus what the firm reinvests in the company. Hence, FCFF can also be expressed as NOPAT * (1-r), where NOPAT is the firm's Net Operating Profit After Tax, r is the firm's reinvestment rate and g is the growth rate. The expression for enterprise value can thus be expressed as follows:

Equation 2.7

$$EV = \frac{NOPAT * (1 - r)}{WACC - g}$$

NOPAT can be decomposed to Return on Invested Capital (ROIC) multiplied by Invested Capital, which gives us:

Equation 2.8

$$EV = \frac{(ROIC * IC) * (1 - r)}{(WACC - g)}$$

By rewriting r as $\frac{g}{ROIC}$ and dividing both sides with IC, the expression can be simplified to obtain the $\frac{EV}{IC}$ multiple:

Equation 2.9

$$\frac{EV}{IC} = \frac{ROIC - g}{WACC - g}$$

Given that $NOPAT = ROIC * IC$, one can multiply both sides with $\frac{1}{ROIC}$, to get the $\frac{EV}{NOPAT}$ multiple as below:

Equation 2.10:

$$\frac{EV}{NOPAT} = \frac{ROIC - g}{WACC - g} * \frac{1}{ROIC}$$

$NOPAT$ can be substituted with $EBIT * (1 - t)$, where t is the corporate tax rate. When multiplying both sides in the expression with $(1 - t)$, we get the $\frac{EV}{EBIT}$ multiple as below:

Equation 2.11

$$\frac{EV}{EBIT} = \frac{ROIC - g}{WACC - g} * \frac{1}{ROIC} * (1 - t)$$

To get to the desired $\frac{EV}{EBITDA}$ multiple, one can replace $EBIT$ with $EBITDA * (1 - D\&A)$ and multiply both the left-hand- and the right-hand side with $(1 - D\&A)$, where $D\&A$ is the depreciation- and amortisation rate measured as $\frac{Depreciation\ \&\ Amortisation}{EBITDA}$.

Equation 2.12

$$\frac{EV}{EBITDA} = \frac{ROIC - g}{WACC - g} * \frac{1}{ROIC} * (1 - t) * (1 - D\&A)$$

As per Kinserdal et al. (2017), the final expression is useful in determining the key performance factors that companies in a peer group must demonstrate identical performance in, for multiple valuations to be theoretically correct. This expression includes growth, profitability, and risk, which are represented through g, ROIC and WACC in the multiple. It is evident that achieving identical performance across comparable firms is impractical, and thus, the final expression derived from the multiples analysis should be used to identify the factors that analysts need to account for when analysing differences among comparable firms.

Moreover, the derivation can be utilised to explain why certain firms are traded at a multiple above or below their peers. The mathematical derivation thus supports the relevance of studying growth, profitability, and risk as fundamental drivers of Enterprise Value to Earnings Before Interest, Taxes, Depreciation, and Amortisation (EV/EBITDA).

2.3 Inter-firm Heterogeneity

As our quantitative research revolves around testing the statistical significance of predictive models of EBITDA multiples by defining and regressing underlying value drivers, based on a set of cross-sectional data, we need to control for differences across firms. As argued by Damodaran (2012), no peer group will be identical to the analysed firm across all variables - and there are 3 main ways of accounting for heterogeneity across the observed data points: regressions, subjective adjustments, and modified multiples. The foremost forms the foundation of our data analysis and will thus be expanded upon more in detail in the subsequent section.

Regression models enable us to make predictions, identify patterns in the data, and estimate the impact of one variable on another. They are particularly useful for identifying the fundamental value drivers of a company and how they affect its valuation. By understanding the causal relationships between these drivers and valuation multiples, analysts can more accurately predict the value of a company and make informed investment decisions. (Alexander, 2008)

In the context of our research, and according to Damodaran (2012), regressions for the purpose of relative valuation will be more robust for larger datasets and stable relationships between underlying value drivers and the EBITDA multiple, as sensitivity to outliers may decrease. Damodaran (2012) argues for basing regressions on one of two main types of datasets: sector and market. He notes that sourcing datasets by sector will help control for differences across variables, given similarities in characteristics. However, sectors may be small, and firms within a given sector may operate simultaneously within several sectors. He further adds that comparability between firms may not necessarily be due to operating within the same sector(s), but instead from similar dynamics in underlying value drivers, such as profitability and growth. Lastly, market regressions may be less sensitive to outliers due to the larger dataset (Damodaran A. , 2012).

As described, we will base our research on publicly listed companies within the S&P Composite 1500 index, on the whole market, GICS sector groups, GICS industry groups, as well as SARD groupings. Bernström (2014) notes that relative valuation should draw on data from publicly listed companies, given they are far more covered than private firms, meaning a higher degree of validity and reliability in data, as well as being more up-to-date. Furthermore, regressions allow testing for multiple independent variables' influences on the dependent variable. Research has shown that realised returns alone are a noisy[14] measure for price prediction, whereas allowing for more variables can potentially enhance predictability (Fama E., 1991; Fama & French, 1998; Poterba & Summers, 1988).

Modified multiples

When using modified multiples to account for heterogeneity across a peer group, one modifies the multiple by placing the most important variable (companion variable) that drives the multiple, as a denominator of the ratio (Damodaran A. , 2012). An example would be the price-to-earnings-growth (PEG) ratio, which divides a firm's P/E ratio by expected earnings-per-share (EPS) growth rate over a designated time period – used by analysts to control for variations in growth rates across firms (Koller, Goedhart, & Wessels, 2010; Easton, 2004). Damodaran (2012) argues the largest drawback of this approach is that one assumes a linear relationship between the added variable (value driver) and the multiple, as well as the assumption that there is one key value driver – excluding the potential autocorrelation with other variables (assumed to be uniform across peer group). In contrast to regressions, where you can test for multiple variables at once.

[14] Noisy data is data from the independent variable that does not help explain the dependent variable

Subjective adjustments

In relative valuation, multiples are calculated on the individual company as well as peer group, followed by computations of arithmetic means and/or medians of said multiples. When significant differences arise (e.g., a higher EBITDA-multiple than peers), subjective reasoning is typically used to judge whether a firm's characteristics like growth, risk, or cash flow, can justify the difference (Damodaran A. , 2012). Hence, a subjective adjustment may be used to justify a revised multiple. As mentioned, corporate valuation by industry professionals is typically derived from using several valuation techniques. Subjective weightings between different methods are often applied in practice, whereas the calculated value will be prone to manipulation by adjusting the weightings (Kasperzak, Janke, & Erkilet, 2021; Asquith, Mikhail, & Au, 2005). Subjective adjustments in either weightings or in relative valuation, can both produce inaccurate price estimates and affect buy/sell volumes in the market, with research suggesting a positive correlation between consensus price recommendations (or revisions) and following negative or positive abnormal returns (Bonini et. al, 2010; Lloyd-Davies and Canes, 1978; Bjerring et. al., 1983; Elton et. al., 1986; Liu et. al, 1990; Beneish, 1991; Stickel, 1994; Womack, 1996). Subjective adjustment implies defining a peer group for the firm in question. Conflicting incentives for financial analysts and valuation experts constitute a need for objectivity (Chan, Karceski, & Lakonishok, 2007), and with many practitioners considering the choice of comparable firms more of "an art form" (Bhojraj & Lee, 2002), we look towards potentially more objective and robust methods to account for heterogeneity across firms.

2.4 Sum of Absolute Rank Differences (SARD)

This section will outline the theoretical method underlying the SARD analysis conducted by Knudsen, Kold and Plenborg (2017). The method selects comparable companies on the bases of the Sum of Absolute Rank Differences (SARD) across a range of chosen variables which are expected to affect the multiple analysed. In mathematical terms, the SARD equation can be expressed as below:

Equation 2.12

$$SARD_{i,j} = |r_{X,i} - r_{X,j}| + |r_{Y,i} - r_{Y,j}| + \cdots + |r_{Z,i} - r_{Z,j}|$$

The SARD is the sum of the absolute differences in rank between two companies, *i* and *j*, with ranks determined by variables x, y, z and so on. A low SARD value for a potential peer implies that the target company and the potential peer share similarities in the chosen variables. If these variables accurately reflect the underlying drivers of the multiple, then the identified peer and the target company should be priced similarly. (Overgaard Knudsen, Kold, & Plenborg, 2017)

2.5 Basics of Linear Regression Models

A regression model is a statistical model which depicts the association between one (or more) independent random variables and a dependent random variable. Usually, a linear relationship of a simplistic nature is assumed between the variables, which takes the form of a simple linear relationship: (Alexander, 2008)

Equation 2.13

$$Y = \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k$$

The left-hand side in the above equation is the dependent variable, often denoted Y. On the right-hand side of the equation, we have the following:

- A set of k independent variables, denoted as X1, X2,…,Xk, are typically referred to as explanatory variables. They are included in the model to explain the behaviour of the dependent variable.
- The model consists of k coefficients, represented as β_1, β_2,…, β_k. These coefficients are typically treated as constants, rather than random variables. The value of the coefficients is estimated based on the data[15] collected for the dependent and independent variables. Each coefficient quantifies the impact of a change in its corresponding independent variable on the value of Y. One should note that in situations where an estimated coefficient does not differ significantly from zero, the related explanatory variable may be removed from the regression model.

2.5.1 Simple Linear Regression

The simplest form of a regression model is, ironically called, the simple linear regression. The model makes the simple assumption that the dependent variable has a linear relationship with just one

[15] The date must constitute an equal amount of observations on each variable, which may be in the form of time series (indexed by the subscript *t*), cross sectional (indexed by the subscript *i*), or panel data (a mixture of time series and cross section, indexed by subscript, *i,t*)

explanatory variable. Simple linear regression relates the response variable, y, to an input variable, x, by the following equation (Ross M., 2017):

Equation 2.14

$$y_t = \alpha + \beta x_t + \varepsilon_t$$

The equation above holds the characteristics of a straight line, where the quantity α is the regression constant which represents the intercept of the line with the vertical axis, and the quantity β is the regression coefficient which represents the slope of the line. The quantity ε is an error term. When there is a low correlation between x and y, the statistics suggest that the error process has a relatively high variance, while a strong correlation between x and y implies a relatively low variance in the error term (Ross M., 2017). The characteristics of the error term determine the most suitable approach for calculating the equation of the line that provides the best fit to the scatter plot. One can adopt the (^) notation to denote an estimator to find the best-fitting line through the scatterplot. We denote equation 2.15 to the following:

Equation 2.15

$$\hat{y_t} = \hat{\alpha} + \hat{\beta} x_t$$

Where, $\hat{\alpha}$ and $\hat{\beta}$ denote the estimates of the regression line's intercept α and slope β. At a given time t, the difference between the factual value y_t and the fitted value $\hat{y_t}$ is denoted ε_t, which is the residual value at time t. This can be expressed as: $\varepsilon_t = y_t - \hat{y_t}$. Given this definition of the residual at time t, the data point for Y at time t is the fitted model value plus the residual. Hence the formula for y_t is (Alexander, 2008):

Equation 2.16

$$y_t = \hat{\alpha} + \hat{\beta} X_t + \varepsilon_i$$

When comparing the theoretical model with the estimated model, firstly, one observes that we can regard the residuals as observations on the error term ε_t. Therefore, when testing the properties of the residuals one can test the assumptions about the behaviour of the error term. Secondly, one observes that the residuals will depend on the values obtained for the coefficient estimates $\hat{\alpha}$ and $\hat{\beta}$. Thus, if we apply two different types of estimators to estimate the coefficients based on the same set of data, we will have two pairs of estimates $(\hat{\alpha}_1, \hat{\beta}_1)$ and $(\hat{\alpha}_2, \hat{\beta}_2)$. This would also give you two series of residuals $[\varepsilon_{1t}]$ and $[\varepsilon_{2t}]$, from which we can compare the properties to decide which method of estimation is best.

2.5.2 Ordinary Least Squares

According to Alexander (2008), the most logical approach within linear regression is to select a method that minimises the residuals, thus ensuring that the predicted values of the dependent variable are as accurate and as close to the observed value as possible. A simple way to obtain estimators that have simple mathematical properties is to minimise the variance of the residuals or to minimise the sum of the squared residuals. This is the criterion of optimisation for Ordinary Least Squares (OLS) (Alexander, 2008). Using the OLS models, one can analyse the connection between one or more explanatory variables and e.g., continuous or interval outcome variables. This approach minimises the sum of squared residuals, which are the discrepancies between the actual and predicted values of the outcome variable. The most common analytical method utilising OLS models is linear regression. This can be done with a single or multiple predictor variable. (Zdaniuk, 2014)

The OLS regression method is commonly used in a variety of scientific disciplines such as physics, economics, and psychology, and numerous textbooks have been written explaining this statistical method and its application in different research areas (Cohen, Cohen, West, & Aiken, 2003; Kleinbaum, Kupper, & Muller, 1988; Montgomery, Peck, & Vining, 2012). OLS regression relies on several assumptions that, if violated, can render the results unreliable. Some of these assumptions are those of homoskedasticity, independence, and the normality of the residuals. For the assumption of homoskedasticity to hold, the residual variance must be consistent across all predictor values. If this is violated, then heteroscedasticity is present, which can be identified by plotting the residuals against the predicted values or by using White's test. The independence of the residuals assumes that the residual of one observation is independent of the residual of another observation. A violation of the independence assumption occurs when some unmeasured variables are systematically similar between some groups of observations. For the normality of residuals, it must be assumed that the residuals follow a normal distribution. This can be evaluated by plotting the residuals or applying the Shapiro-Wilk test. In addition to meeting the assumptions stated above, an important concern in regression is the presence of outliers, which can have an excessive influence on the results (Zdaniuk, 2014).

Chapter 3

Literature Review

Having established the theoretical underpinnings, this section aims to delineate the literature and prior empirical research which have contributed to the field of firm multiple valuation. Whilst the theoretical framework has provided a structure for testing the hypotheses and to identify key variables and relationships which need to be studied, the literature review seeks to identify and synthesise the existing research and knowledge related to multiple valuation. Hence, the hypothesis will later be formulated based on the theoretical framework and the identified gaps in the current literature.

Current academic studies concerning multiple valuation are mainly focusing on which multiple to use for firm valuation, as well as the accuracy of multiples in predicting firm value. According to the scholarly discourse, the literature pertaining to multiple accuracy can be categorised into two subcategories: those that focus on multiple construct and those that examine the selection of comparable companies. The differentiation facilitates a more comprehensive understanding of the central themes in this field of study. See **Appendix 1** for a table depicting a summary of the literature review in this paper.

3.1 Valuation Multiples

This section aims to map the landscape of existing literature on the topic of valuation multiples, which lays the bed upon which our paper is nested. Through contextualising our paper against preceding research, we can better understand where it departs from existing literature and potentially contributes with novelty.

As previously touched upon, multiples are applied extensively by industry professionals and private investors alike, when performing corporate valuations. These lay the foundation for buy/sell decisions in public markets, affecting company share prices and market capitalisations, which in turn significantly impacts the attractiveness of equities and companies' ability to raise capital in equity markets (Pearl & Rosenbaum, 2009). Therefore, it might be unsurprising that prior literature has predominantly been centred around the accuracy of valuation multiples and relevant financial metrics, both in terms of their effectiveness in representing firm performance, as well as comparability across defined peer groups. We will further elaborate on comparability across defined peer groups in **section 3.3**.

3.1.1 Multiples

There is some documentation of the use of valuation multiples in US courts already during the early 1900s, with courts in certain states using earnings multiples (PE ratios) to value utility companies in rate-setting cases (Goddard, 1924; Bonbright, 1927). Following the 1929 stock market crash, one could imagine an increase in demand for knowledge on analytical approaches to investing. A body of literature on multiples was introduced during the 30s, such as Benjamin Graham's book "Security Analysis" from 1934, discussing the application of P/E and P/B ratios. However, a maturing of corporate finance theory and popularisation of multiples in corporate valuation by industry professionals and retail investors was first seen in the 60s (Jensen & Smith, 1984). As financial markets have been successively deregulated and equity markets have grown in size (Helleiner, 1995), the amount of academic research on the topic of multiples has gradually followed. Given that the theoretical foundations of valuation multiples have been previously outlined, this particular section will concentrate on the current empirical evidence within the existing literature.

Kaplan and Ruback (1995) produced one of the earliest widely cited research papers on the accuracy of valuation multiples. In their empirical research on a dataset of highly leveraged transactions, they applied rigorous DCF models alongside valuation multiples – and found that simple EBITDA multiples yielded comparable valuation accuracy to the DCF models (Kaplan & Ruback, 1995). Liu et al. (2002) arrived at a similar conclusion, using forward multiples. Further highly cited research has demonstrated on different datasets that valuation multiples can yield similar accuracy as DCF models, but that they tend to be slightly less accurate (especially for small-cap companies) and are best used in conjunction with a DCF analysis (Damodaran A., 2009; Fernandez P., 2013).

Kim & Ritter (1999) used a sample of 7,470 completed IPOs to study the accuracy of prices predicted by multiples using trailing, current, and forward-looking multiples, against actual prices realised.

They found that forward-looking multiples on average were the most accurate, which in later times generally has been the consensus. Liu, Nissim & Thomas (2002) also support this view stating that forward earnings multiples on average produced lower pricing errors. Furthermore, research by Lie & Lie (2002), Schreiner & Spremann (2007), and Plenborg & Pimentel (2016) arrive at similar conclusions, with Lie & Lie (2002) estimating a mean absolute percentage error of 10-15%.

Abukari et al. (2000) found book value and earnings-based ratios to be the most significant. Liu, Nissim & Thomas (2002) found in their study that earnings multiples were more accurate than cashflow-based multiples. Lie & Lie (2002) studied the accuracy and potential limitation of valuation multiples, by comparing implied value from estimated valuation multiples, and actual transaction prices for 1,000 firms across different industries. Amongst their findings, they point out that EBITDA multiples on average were more accurate than EBIT multiples and that asset multiples are more accurate than earnings multiples, which opposes the two preceding views of Abukari et al. (2000) and Liu, Nissim & Thomas (2002). Harbula (2009) applied a similar approach to test valuation multiple accuracy, but on European firms, and found cashflow-based multiples were more accurate than earnings-based multiples. This was further supported by Koller, Goedhart and Wessels (2010), who argues that cashflow-based multiples are more reliable than accrual-based multiples, as accrual-based multiples are susceptible to subjectivity in accounting methods and arbitrary allocation procedures.

Baker & Ruback (1999) also pointed out that their methodology and evidence suggested that harmonic means should be a preferred method of aggregating valuation multiples across a chosen peer group. A statistical analysis conducted by Liu et al. (2002) supported this view, adding that it worked especially well for skewed data with outliers, characteristic of stock market returns[16]. They also coined the term "trimmed harmonic mean", a slightly adjusted harmonic mean with a lower- and higher quantiles specified cut-off, which their findings suggested being an improved method (Liu, Nissim, & Thomas, 2002). Findings from both Herrmann & Richter (2003) and Schreiner & Spremann (2007) suggested that harmonic means generally understated firm equity values. Plenborg and Pimentel (2016) suggested a lack of evidence as to whether harmonic means was the superior method, and suggested that medians also tackled the issue of outliers in data – and further

[16] Stock returns in markets are generally characterized by so-called "fat tails" – leptokurtic distributions of varying degrees of returns (Munk, 2021)

compared these methods against geometric means, size-adjusted weighted averages and simple arithmetic means (Plenborg & Pimentel, 2016).

EV/EBITDA multiple

Rossi and Forte (2016) argued in their study that EV/EBITDA holds major theoretical advantages as it is independent of differences in taxation and accounting policies concerning depreciation and amortisation across firms. This is in line with Rosenbaum & Perl (2009) which argues for practitioners to use the EV/EBITDA due to EBITDA being independent of capital structure and tax regimes. Prior studies have demonstrated that divergence in taxation and accounting policies between comparable firms leads to impaired comparability and biased valuation estimates (Plenborg & Pimentel, 2016). Furthermore, using EBITDA rather than sales as the denominator in the EV/EBITDA ratio is supported by the extensive use of EV/EBITDA by practitioners for enterprise valuation purposes. (Credit Suisse, 2016; UBS, 2001). In an empirical study conducted by Kim & Ritter (1999), they used several measures for valuing IPO companies. The study analysed Price/Equity (P/E), Price/Book (P/B), Price/Sales (P/S), EV/Sales, and EV/EBITDA, and found that all multiples yielded positively biased estimates, but EV/EBITDA yielded the most precise valuation.

To identify comparable companies, analysts commonly rely on industry classifications, as businesses operating within the same industry typically exhibit similar economic traits. Nonetheless, the efficacy of such classifications as a tool for identifying comparable firms has been brought into question, as Lee, Ma, and Wang (2015) have noted that these classifications are only rough approximations. Therefore, while the EV/EBITDA multiple is a valid predictor of firm value across firms, its appropriateness as a valuation measure depends on the industry-specific financial items that are relevant to each sector. Optimal multiple constructs vary across sectors and industries, and practitioners should consider these factors when selecting an appropriate multiple for valuation purposes.

Baker & Ruback (1999) used Gibbs sampling on 22 S&P industries to estimate minimum variance multiples and error structure and found that the best measure of financial performance (in this case, EBIT, EBITDA and revenue multiples) varied across industries. This was further supported by Coppola et al. (2000) who illustrated that the accuracy of EBITDA multiples depended on a multitude of factors, such as asset lifespans. Fernandez (2001) supported the notion of differences in applicability across industries, by quantifying the average multiple dispersion across 175 companies within different industries. Furthermore, Plenborg & Pimentel (2016) argues that the optimal construct generally depends on the relative importance of certain financial items, such as depreciation and amortisation, leverage, CAPEX, or earnings growth in specific industries.

For instance, Damodaran (2012) argues that the EV/EBITDA multiple is particularly useful for firms that operate in sectors requiring large investments in infrastructure with long formation periods. Empirical evidence supports this view, as Gupta (2018) found that EV/EBITDA multiples provide optimal estimates for companies operating within the capital-intensive steel sector in India, whereas e.g. P/BV multiples were found to be more appropriate for the banking sector. Similarly, Rossi and Forte (2016) concluded in their study that EV/EBITDA multiples should preferably be applied when estimating the value of firms operating in infrastructure or manufacturing industries where CAPEX and depreciation hold significant importance. Meanwhile, Harbula's (2009) research revealed that EV/EBITDA was particularly suboptimal for specific industries, including banking and insurance, life sciences and healthcare, and real estate.

3.1.2 Econometrics & Valuation Multiples

Attempts to explain economic phenomena in statistical terms can be dated back to Sir William Petty's "Political Arithmetic" in the 17th century (Cox, 1962). However, modern econometrics was introduced by the likes of R. Frisch, J. Tindbergen, T. Haavelmo and H. Moore (Cox, 1962; Heckman, 1992). In attempts to understand, model and predict the economy, early econometricians such as C. Juglar and W. Persons sought empirical regularities in data that could help model the economy, with U. Yule, E. Slutsky and R. Frisch introducing the concept of random shocks into time series modelling (Heckman, 1992). Heckman (1992) further states that early economical statisticians sought to find "causal" relationships to the economy by studying potential deterministic data, whilst the idea of random shocks moved academic consensus towards the economy being innately cyclical. Haavelmo (1944) introduced probabilistic frameworks based on stochastic random variables and illustrated the importance of testing the predictive power (model fit) of said statistical models against historical market developments.

Markowitz (1952) pioneered the use of mathematical frameworks in equity markets, through his proposal of the mean-variance frameworks, which arrived at "optimal" portfolio allocations through studying the covariance structure of assets under management. This laid the ground for the widely covered portfolio diversification theory, on which a large bulk of fund managers and retail investors base their investment philosophies to this day (Munk, 2021). With the popularisation of portfolio theory, academia, and investors alike, started drawing on theory from early econometricians such as Haavelmo and his linear regression models. One of its most popular manifestations is the capital asset pricing model (CAPM), which uses linear regression to model the relationship between an asset's expected return and its beta (Sharpe, 1964). Fama & French (1993) proposed a three-factor model applying linear regression to model the relationship between asset returns and three different proposed factors. They later added to their model with the Fama and French five-factor model (Fama & French, 2015), which has gained widespread acceptance within financial academia and practices (Munk, 2021).

Whilst regression methods are widely used in financial markets, differing opinions exist on the explanatory power of such models, as well as which ones are most appropriate. For example, several studies suggest non-linear relationships between underlying value drivers and valuation multiples. Basu (1977) demonstrated a non-linear relationship between stock returns and price-earnings ratios. Ang & Chen (2007) studied the CAPM performance from 1926-2001 and found market beta values and stock returns vary non-linearly. Fama and French (1993) showed that there is a concave relationship between stock returns and book-to-market ratios. Damodaran (2012) illustrated a non-linear relationship between PEG ratios and growth rates in cross-section samples.

Lastly, research points towards the potential of biased outputs in regression models from intertemporal differences. Damodaran (2012) argues that when regression models attempt to model and explain differences over time, they may be skewed due to changes in market conditions. Damodaran (2012) illustrated this point by showing varying R-squared statistics for valuation multiples on a dataset from 1987-1991, with a reduction from 90% to 30%. Further illustrating this point, Harbula (2009) showed that valuation errors peaked around economic recessions.

3.2 Underlying Value Drivers

In much of the empirical research conducted in the field of relative valuation, the underlying value drivers and the multiple constructs are assumed given, with little regard to the inputs utilised in the analysis. This approach often overlooks critical factors that significantly impact the result, which may lead to inaccurate conclusions. Consequently, it is imperative to prioritise careful considerations and deliberate selection of inputs in relative valuation analyses.

Section 2.2.2 of the study outlined a theoretical basis for the value driver of the EV/EBITDA multiple. The derivation, based on Kinserdal et al's (2017) theoretical foundation, revealed that the driver that underlies the EV/EBITDA multiple primarily includes growth, profitability, and risk. The subsequent section of this study aims to offer a comprehensive overview of the essential value drivers and associated considerations.

3.2.1 Growth

This section aims to review previous literature conducted on the relationship between growth and multiples. The impact of growth on firm value has been extensively studied in the literature. It can be argued that the concept of growth and classical growth theory dates back to the classical economist Adam Smith (Ucak, 2015). Adam Smith (1776) contributes to economic growth theory with his ideas of competitive behaviour and dynamics in equilibrium, the role of diminishing returns, as well as the interaction between the growth rate of population and per capita income.

The impact of growth on valuation multiples is widely regarded as a significant value driver by analysts (UBS, 2001). However, it is important to note that the correlation between growth and value multiples may also vary depending on the type and origin of growth. An instance of this phenomenon can be observed when an increase in the prices of goods and services leads to a growth in the top-line revenue of a company. Theoretically, this growth may lower the multiples by increasing the expenses incurred in investments. Conversely, growth resulting from the implementation of efficient processes and practices unequivocally contributes to the overall value of the firm UBS, 2001). Furthermore, Damodaran (2006) found in his study that a firm's expected growth rate is one of the most important value drivers in any firm. Lastly, Yin et al. (2018) discovered in their study that firms with higher future expected growth, as compared to their peers, are assigned premiums to their trading multiples.

However, there are some empirical findings that dispute the relationship between growth and firm value. Gupta (2018) found growth to be a significant explanatory variable for EV/EBITDA for several sectors, but not for the steel- and banking sectors. Furthermore, according to Credit Suisse

(2016), growth is one of the key determinants of multiples, but its relative importance has decreased over time. The study found that in the pre-financial crisis period, growth was considered the most significant determinant of firm value, while in the post-financial crisis period, profitability surpassed growth's significance. This suggests that while growth remains an important factor, other metrics, such as profitability, are also critical in determining the value of a firm. Overall, most of the literature suggests a strong and positive relationship between growth and firm value.

Proxy for growth

Empirical evidence from prior studies has concluded that growth is an important determinant of firm value. However, the proxy that can be used for growth takes different shapes. Zarowin (1990), Kakita (2005) and Damodaran (2006) all found in their own studies that earnings growth (such as in EBIT and EBITDA) has a positive and significant impact on the valuation of a firm. Zarowin (1990) studied the relationship between growth and PE ratios. Whereas Kakita (2005) studied the relationship between growth in sales, among other value drivers, and stock price. Achleitner (2011) argues in his study that EBITDA growth is a good measure of operating performance improvements, and cash flow improvements, which has a direct impact on firm value. Similarly, Hammer et al. (2023) argue in their study that EBITDA growth, especially EBITDA CAGR, is a suitable proxy for growth, due to its relationship with firm value. Lastly, Damodaran (2012) outlines that companies with elevated EBITDA tend to produce more substantial cash flows, resulting in a higher enterprise value.

3.2.2 Profitability

In theory, when a firm's profitability increases, its expected future cash will likely increase, which is a crucial determinant of the firm's intrinsic value. Essentially, a more profitable firm is expected to generate more cash in the future, leading to a higher valuation of the firm.

Empirical evidence from prior studies has concluded that profitability is a significant determinant of firm value. Bernard (1994) found that variations in Price / Book Value (P/BV) can be explained by a firm's profitability together with its risk and growth. Nel (2009; 2010) conducted a critical examination of the disparity between the theoretical frameworks proposed by academic scholars and the actual practices employed by investment bankers concerning firm valuations. Their study concluded that both academia and investment practitioners agree on the suitability of earnings as a value driver. Koller, Goedhart & Wessels (2010) found in their study that there is a strong correlation between increased profitability and increased firm value. According to Credit Suisse (2016), there is evidence to suggest that an increase in profitability has a more significant impact on multiple expansions than a comparable increase in growth. Specifically, they found evidence supporting that firms which perform above the median level in terms of profitability experience a two folded increase in their valuation multiples as compared to a firm that performs above the median level in terms of growth. Furthermore, Gupta (2018) evaluated the prediction accuracy of four valuation multiples across three sectors for all Indian-listed companies, to identify the fundamental value drivers for these multiples, using a regression-based approach. His study concluded that for EV/EBITDA, Return on Employed Capital (ROC) and Price to Sales (P/S), profitability is the key driver.

However, Credit Suisse (2016) emphasised the fact that sustainability in profitability has a significant effect. Meaning that the nature of the profitability, whether it is short-term or long-term, influences the value drivers' significance. This is somewhat aligned with UBS’s (2001) statement on a multiple’s significance is dependent on whether the profit used to calculate it can provide insight into future profit potential. Hence, in cases where this is not true, UBS (2001) argues there be two options: exclude extraordinary items when utilising historical profits; or rely on forecasted profits instead of historical ones. Conclusively, there is an academic and professional consensus which suggests that profitability has a significant role in determining firm value.

Proxy for profitability

Whilst being an academic and professional consensus suggesting that profitability is a fundamental underlying value driver of firm value, profitability is a wide term and can be interpreted in relation to several different accounting elements, such as sales (return on sales), asset (return on asset), investment (return on invested capital) or equity (return on equity). Hence, which measure to use as a proxy for profitability has been a fruitful area of research, which has been highly discussed and empirically tested among researchers.

Berk & DeMarzo (2019) state in their book that the most commonly used ratios related to profitability are gross margin, operating margin, EBIT margin and net profit margin. The gross margin reflects a firm’s ability to sell a product for more than the cost of producing it. However, a firm is not only subject to the cost of goods sold but also operating expenses. Hence another commonly used profitability ratio is the operating margin. Furthermore, Berk & DeMarzo (2019) argues for the net profit margin as a proxy for profitability. In continuation, compared to the operating margin, one should be cautious when using the net profit margin as a relative measure, since differences in the metric can be due to efficiency- and leverage differences as well as accounting practices.

Freeman et al (1982) created a probabilistic model of earnings changes and successfully managed to reject the hypothesis that firm profitability follows a "random walk" and concluded that ROE is a useful proxy for predicting firm profitability. Similarly, Ohlson (1995) and Feltham & Ohlson (1995) have emphasised the theoretical significance of ROE in the application of valuation models in general, and in the residual income model specifically. This underscores the pivotal role played by ROE as a metric for evaluating a firm's financial performance and its potential to generate residual income, which is the difference between the firm's actual income and its required rate of return. Nissim and Penman (2001) criticised ROE as a measure of profitability since it can be affected by the firm's choice of capital structure. They argued in their study that return on net operating asset (RNOA) is a better ratio thus it captures the firm's operating profitability without financial leverage. Since Nissim and Penman's (2001) study RNOA has become commonly used in valuation research. (Nissmin & Penman, 2001; Fairfield, Sweeney, & Yohn, 1996; Penman & Zhang, 2003; Fairfield, Whisenant, & Yohn, 2003a; Richardson, Sloan, & Tuna, 2005)

Yousaf & Dey (2022) tested three different proxies of firm performance, return on asset (ROA), return on equity (ROE) and Return on capital employed (RoCE) as dependent variables. The study deployed a Chi-squared Automatic Interaction Detector (CHAID) to examine the best proxy of firm profitability. The study was conducted using a sample of 287 firms taken from the automobile, construction, and manufacturing sectors, and concluded ROA to be the best proxy for profitability.

Opposingly, Koller et al (2010) and Kinserdal et al. (2017), provides different theoretical reasoning on the fundamentals of profitability. They argue that ROIC is a crucial metric in financial analysis since it incorporates a firm's invested capital, which is not accounted for in nominal operating profit ratios such as EBIT, NOPAT or NOPLAT. Due to its consideration of invested capital, the ROIC ratio is more appropriate for assessing the profitability of a firm's operations and determining whether the actual return is acceptable in comparison to the required return of investors. Therefore, the ROIC ratio serves as a more reliable indicator of a firm's profitability than other operating profit ratios. Furthermore, empirical research has shown that a higher ROIC tend to improve a firm's credit ratings and increase firm valuation, which also stresses the importance of ROIC. (Koller, Goedhart, & Wessels, 2010; Kinserdal, Petersen, & Plenborg, 2017).

Kinserdal et al. (2017) further discuss ROIC in relation to Economic Value Added (EVA), which is expressed in the following way:

Equation 3.1

$$EVA = \left(ROIC_{after\ tax} - WACC\right) * Invested\ capital$$

The above expression shows that a firm will create an excess return when ROIC exceeds WACC and, opposingly, destroy value when WACC exceeds ROIC.

3.2.3 Risk

The purpose of this section is to investigate the fundamental determinant of risk and how it configures firm value. However, compared to growth and profitability, risk is a much more complex value driver due to its underlying sentiments. According to Chandra (2014), this is due to its various components which may have conflicting effects. These components include but are not limited to, capital structure, market sentiment, performance volatility, opinion on management performance, and the firm's business portfolio attractiveness. (Chandra, 2014)

In the field of corporate finance, a pivotal inquiry concerns the optimal composition of securities a firm should offer to the public to procure funding from investors. This decision plays a crucial role in determining the capital structure of the firm, which represent the aggregate value of debt, equity, and other types of securities that the firm currently has outstanding. (Berk & Demarzo, 2020)

Since the publication of the Modigliani and Miller's (1958) "irrelevance theory of capital structure", corporate capital structure has been a well-discussed topic among finance academics and practitioners (Luigi & Sorin, 2009). Modigliani and Miller (1958), hereinafter M&M, have contributed significantly to the field of corporate finance, since before their theory, there was no theory of capital structure which was generally accepted.

M&M's theory assumes that a firm has a predetermined set of expected cash flows. They argue that when a firm chooses to finance its assets using a specific mix of debt and equity, all it does is allocate the cash flows among its investors. The theory assumes that both investors and firms have equal access to financial markets, allowing for homemade leverage. More specifically, this means that investors can create leverage if it is not provided by the firm or eliminate leverage if the firm has taken on more than desired. The work M&M (1958) conducted established sound evidence that, when assuming perfect capital markets with no arbitrage opportunities, no bankruptcy risk, and no corporate taxes, the capital structure of the firm has no impact on its market value. Many adjacent authors concluded the same results in their empirical studies, under different and more general assumptions. (Fama & Miller, 1972; Hirshleifer, 1966; Stiglitz, 1969; Stiglitz, 1974).

However, when easing the assumptions required by M&M, there are some empirical findings that dispute the fact that capital structure has no effect on firm value. Baumol and Malkiel (1967) argue in their paper that capital structure does influence firm value if investors are faced with transaction costs when engaging in arbitrage activities. Rubinstein (1973) found in his research that if security markets are segmented to some degree, and if debt is traded in a separate market where more risk aversion is present compared to investors in the firm's equity, then an increase in debt can decrease the total firm value. Following the same notion, Stiglitz (1972) demonstrated in his study that if debt and equity are traded in separate markets, and debt holders are more pessimistic about the firm than equity holders, then a large increase in debt can lower the total firm value.

Furthermore, in the presence of bankruptcy and reorganisation costs, Baxter (1967) Bierman & Thomas (1972), Kraus & Litzenberger (1973), and Robichek and Myers (1966) concluded in their respective studies that debt policy significantly affects firm value and that an internal optimal capital structure can be established. Finally, Scott Jr (1976) argues in his framework that debt is valuable, primarily because interest payments are tax deductible, but that an increase in the level of debt increases the probability of bankruptcy leading to the firm incurring bankruptcy cost. Furthermore, when analysing a firm's financial ratio such as the Return on Equity (ROE), M&M's proposition may value the firm's capital structure as not relevant. This may seem odd, since when assuming a positive spread, ROE can be mechanically increased with the use of leverage. However, it is argued that the increase in leverage would also increase the discount rate, which theoretically should result in no change in equity value (Nissmin & Penman, 2001).

Another well-established theory discussing a firm's capital structure is the Trade-off theory. According to this theory, the optimal capital structure represents a balance between the advantages of the interest tax shield and the cost of financial distress. In the absence of debt, the value of the firm can be assessed by estimating the cost of equity and an unlevered beta. As the level of debt increases, the tax shield benefit initially increases at a faster rate than the cost of financial distress. (Mauboussin & Callahan, 2023)

The conventional perspective which considers cost of capital as the most appropriate approach to gauge risk in company valuation implies that unfavourable risk due to ambiguity surrounding future cash flows must be evaluated separately rather than comprehensively. This approach is useful when conducting relative valuations, as the underlying value drivers such as growth, and profitability should reap the positive effect of risk on expected cash flows. (Mauboussin & Callahan, 2023)

In conclusion, empirical evidence suggests that risk and firm value have an inverse relationship, which may seem strange to investors thinking risk-taking is rewarded with higher returns. Risky assets are expected to have higher returns than less risky ones, but they also have more uncertain cash flows, which means their value is discounted more (Berk & Demarzo, 2020).

Proxy for Risk

A firm is exposed to various types of risks, including but not limited to, market risk, credit risk, operational risk, and liquidity risk. However, on the back of prior empirical research, this section seeks to define a proxy for risk which covers most risk a firm is exposed to.

The theoretical framework of capital markets posits the Capital Asset Pricing Model (CAPM) which was postulated by Sharpe (1964), Lintner (1965), and Mossin (1966). CAPM provides an explicit characterisation of the appropriate risk metric for a security issued by a firm, specifically in systematic or non-diversifiable risk. Furthermore, the CAPM suggests that the systematic risk (β) of a security is an appropriate measure of the sensitivity of its returns to changes in the market portfolio, and only systematic risk bears any material impact on an individual security's expected return.

However, CAPM only measures the cost of equity and thus only the risk entitled to the equity holder. Hence, to account for the total firm risk one needs to account for the firm's selected capital structure, configuring both assets and liabilities (Abid & Mseddi, 2010). Mauboussin & Callahan (2023) argues that whenever a firm's capital structure is comprised of both debt and equity, the measure of risk needs to be weighed on both the cost of equity and the cost of debt. They further argue that the firm's weighted cost of capital (WACC) is the most appropriate measure of risk, thus it works as the investor's hurdle rate which the investor requires for bearing the associated risk (Mauboussin & Callahan, 2023). Similarly, Higgins (2005) argues there to be an inverse relationship between a firm's enterprise value and WACC, where a higher WACC results in a higher discount rate and lower profitability for the firm.

Empirical evidence from prior studies has concluded the same. Hussain and Chakraborty's (Hussain & Chakraborty, 2010) conducted a study on 24 listed commercial banks in the Dhaka Stock Exchange and concluded that there is a strong negative correlation between the commercial banks' cost of capital and their respective returns. Ross (2007) discusses in his study that WACC often serves as the discount rate of the projects undertaken by the firm and concluded that a higher discount rate often results in less cash flow which implies low NPV projects, and ultimately lower profitability of the firm. In a study conducted by Loughran & Wellman (2011) during the period of 1963 to 2009, they discovered comparable outcomes on how the cost of capital affects stock returns. Their research found that firms with lower discount rates generally had higher justified enterprise multiples. Lastly, Damodaran (2012) argues in his research that there is an inverse relationship between risk and firm value. I.e., firms experiencing an increased level of risk should also experience a decrease in firm value.

Following Kinserdal et al. (2017) derivation of EV/EBITDA, depicted in equation 2.12, one can note that WACC is present in the denominator on the right-hand side of the equation. This means that an increase in WACC would increase the denominator. Hence, when holding all else equal, an increase in WACC would decrease the EV/EBITDA multiple.

3.3 Peer Group Selection

Plenborg & Pimentel (2016) segments implementation issues, applicable to valuation multiples, into 8 groups: 1) Peer group selection, 2) Cashflow based vs accrual value drivers, 3) Reported earnings versus expected earnings, 4) Aggregating methods, 5) Accounting differences, 6) Normalisation of earnings, 7) impact of size, and 8) illiquidity discount and control premium.

There is a scholarly consensus on the importance of identifying comparable firms when performing relative valuation, as benchmarking a firm against dissimilar firms will likely yield biased and inaccurate estimates (Plenborg & Pimentel, 2016). As such, a body of literature has accumulated over time on the topic of peer group selection, with different schools of thought on appropriate methods. Two main schools of thought have emerged, as well as new methods gaining traction. The first school of thought, which is the most used, approaches the peer group selection by basing on industry affiliation (Berk & DeMarzo, 2019). The second school of thought argues for higher accuracy through peer group selection of firms with similar dynamics in underlying value drivers, such as growth and profitability (Pearl & Rosenbaum, 2009). Newer methods have surfaced and gained traction in financial academia in recent years – online traffic data and the SARD approach (Lee et al., 2015; Plenborg et al., 2017). The SARD approach will be elaborated upon in section 3.4.

Alford (1992) conducted a cross-sectional analysis of US publicly listed firms, comparing the accuracy of peer group selection by basing peer groups on Standard Industrial Classification (SIC) codes, and by basing peer groups on proxies for earnings and risk, with accuracy expressed as the absolute percentage deviation from observable market capitalisations. Findings by Alford (1992) suggested that peer group selection increased in accuracy when increasing SIC digits (up to three digits). The notion of peer group selection through industry affiliation was further supported by Fama & French (1997), who found that using Global Industry Classification Standard (GICS) reduced variation in valuation multiples. Increased accuracy in valuation through using industry classification was also suggested in findings by Eberhart (2004). Bhojraj, Lee & Oler (2003) found that GICS codes yielded more accurate valuations than other industry classifications, as well as better explaining differences in e.g., growth rates. Cheng & McNamara (2000) suggested that a combination of using industry classification as well as P/E and P/B ratios leads to the most accurate valuations. Young & Zeng (2015) suggests increased accuracy in peer group selection through industry affiliation, due to more similar accounting methods.

Bhojraj & Lee (2002) regressed forward values of observed multiples (P/B and EV/Sales), to measure cross-sectional valuation, and found that selecting peer groups based on similarities in the 'warranted' multiples achieved higher accuracy. This revised the view of Alford (1992) and others, Bhorjaj et al. (2003) later found that combining this method with sampling on CIGS codes, could better explain cross-country variation. Findings by Dittman & Wiener (2005) showed the highest statistical significance when basing peer groups on similarities in ROA. Henscke & Homburg (2009) found that P/B showed the highest statistical significance and that choosing peer groups based on multiples yielded the most accurate estimates. However, as pointed out by Damodaran (2012), relative corporate valuation against industry peers may be subject to under- or overvaluation of the sector itself.

Other researchers have devoted different methods of weighing the multiples. Cheng & McNamara (2000), Yee (2004) and Yoo (2006) suggested linear combinations by averaging different value estimates – whilst Yee (Yee K., 2008) suggested applying a Bayesian framework. Findings by Bhorjaj (2002) and Damodaran (2002) suggested regression approaches, arguing this would best account for interfirm differences.

Alternative methods have been proposed when it comes to peer group selection, with one of them being the selection of peers based on the search traffic pattern of websites (Plenborg & Pimentel, 2016). In a study conducted by Lee, Ma and Wang (2015), search traffic patterns on the U.S. Securities and Exchange Commission database EDGAR (Data-Gathering, Analysis and Retrieval), were analysed to find comparable companies. They found that firms that are frequently co-searched (search-based peers) in several tests outperformed GICS6 industry classification peers.

As mentioned by Young and Zeng (2015), accounting practices can make different firms appear similar, and similar firms appear different. Lie & Lie (2002) found a disparity between depreciation schedules and actual deterioration of asset values, and Coppola et. al. (2000) found that EBITDA could be easily manipulated by applying aggressive accounting policies for revenue and expense recognition, concomitant adjustments, asset write-downs, depreciation schedules, as well as arbitrary adjustments in EBITDA for non-recurring items. Valuations, in theory, should normalise earnings for non-recurring items not affecting future cash flows, such as restructuring costs, legal settlements and impairment charges (Penman S. H., 2007; Plenborg & Pimentel, 2016). Studies show multiple valuation using earnings excluding extraordinary items and/or before special items, outperform non-adjusted earnings (Liu, Nissim, & Thomas, 2002; Nissim, 2013). In general, academic consensus supports earnings multiples with non-recurring items removed (Plenborg & Pimentel, 2016).

Several studies find significantly lower prediction errors for both equity and enterprise multiples and higher valuation accuracy for larger and older firms, compared to smaller firms (Alford, 1992; Kim and Ritter, 1999; Cheng & McNamara, 2000; Lie & Lie, 2002). Pratt et al. (2008) and Comment (2012) suggest that the most common valuation discounts and premiums are the illiquidity discount and the control premium. Further studies suggest the same, and find different average ranges of discounts, typically between 15% to around 30% (Silber, 1991; Bajaj et al. 2001; Emory, Dengel and Emory, 2002; Officer, 2007). For control premiums, theory suggests that it should increase the further the status quo value of a firm is from its optimal value Plenborg & Pimentel, 2016), and control premium averages for U.S. public targets vary between studies, typically between ~25% to ~45% (Pratt, Reilly, & Schweihs, 2008; Betton, Eckbo, & Thorburn, 2009). For non-US firms, Petersen et al. (2006) found an average control premium of ~30%. Financial academia suggests that illiquidity discounts should apply when securities are less marketable, i.e., selling/buying stocks takes longer time and effort, at higher transaction costs (Silber, 1991; Damodaran, 2012; Bernström, 2014). Control premia are generally paid by acquiring companies due to increases in bargaining power, as controlling interests enable decision-making (Dyck & Zingales, 2005).

3.4 Sum of Absolute Rank Differences (SARD)

In a more recent study, Knudsen, Kold and Plenborg (2017) proposed an alternative method. Their method was named the 'Sum of Absolute Rank Differences' (SARD) and follows the second school of thought to identify peers. Compared to peer selection through industry affiliation, the SARD approach selects peer groups based on similar dynamics in underlying value drivers. Specifically, the SARD approach ranks each company within a sample based on a set of selection variables relative to the remaining companies in the data sample. The peer group, to the chosen target company, is later defined as the companies with the smallest sum of absolute rank differences across the target company's variables of interest.

Knudsen et al. (2017) argue that the SARD approach offers notable benefits, compared to other peer group selection approaches. Firstly, it does not face limitations in terms of the number of variables that can be employed for identifying peers or the number of observations that are accessible, whilst being unaffected by industry classification. Secondly, the approach offers flexibility by allowing integration with other methods, such as the industry classification approach. Thirdly, the selection variables in the SARD approach can be customised to fit the requirements of any desired multiple, which arguably should lead to a more precise valuation estimate. Finally, the SARD approach is not only intuitive but also straightforward to apply. Knudsen et al. (2017) applied their SARD approach to a sample consisting of companies that compose the S&P Composite 1500. The research concluded that selecting companies by using the SARD approach within an industry yields more accurate valuation estimates than selecting peers on GICS industry affiliation.

Potential disadvantages of using the SARD approach compared to GICS industry affiliation are outlined by Rossi & Forte (2016). The authors argue that the utilisation of pre-defined industry affiliations, such as GICS, in the process of categorising comparable firms is believed to reduce the potential selection bias inherent in subjective methodologies, thereby enhancing the likelihood of obtaining efficient and unbiased estimations.

Chapter 4

Hypothesis Formulation

This section deductively derives a set of hypotheses based on a comprehensive review of relevant literature and empirical research. These hypotheses will subsequently undergo rigorous testing whilst drawing on insights and outcomes from prior research conducted in the realm of multiple valuation. The purpose of this section is to showcase the soundness and significance of the chosen areas of investigation.

The theoretical framework provides convincing evidence that a few core value drivers significantly influence multiples. Specifically, some scholars have widely recognised growth, profitability, and risk as the three most significant drivers of fundamental value (Bhojraj & Lee, 2002; Plenborg & Pimentel, 2016; Knudsen et al., 2017).

Based on previous empirical research, growth[17] and profitability[18] are expected to have a positive association with the analysed multiple. In contrast, risk is expected to have a negative association with the studied multiple[19]. Additionally, earlier studies indicate that slight differences in fundamental value drivers are largely responsible for explaining why some firms trade at a multiple above or below their peers. Comparing valuation premiums and discounts to relative performance in fundamental value drivers provides a more accurate depiction of firms' under- and overvaluation. Furthermore, GICS industry affiliation for selecting peer groups has been widely discussed among researchers and industry practitioners. Lastly, even though SARD grouping is a relatively new

---

[17] As argued by UBS (2001), Zarowin (1990), Kakita (2005), Damodaran (2006) in **section 3.1**

[18] As argued by Bernard (1994), Nel (2009; 2010), Koller, et al. (2010), Credit Suisse (2016), and Gupta (2018) in **section 3.2**

[19] As argued by Malkiel (1967), Rubinstein (1973), Baxter (1967), Bierman & Thomas (1972), Kraus & Litzenberger (1973), Robichek & Myers (1966), and (Berk & Demarzo, 2020) in **section 3.3**

method for peer group segmentations, it has garnered interest within financial academia and industry professionals. Based on these reasons, the present study formulates the following hypotheses concerning the two research questions in aim.

Research Question 1: To which degree can variance in selected proxies significantly explain variances in EV/EBITDA multiple valuations for US publicly listed firms within the S&P 1500 composite index, when running linear regression models?

- Hypothesis 1: *In isolation, growth as a dependent variable has a positive and significant impact on EV/EBITDA, with a t-statistic ≠ 0*
- Hypothesis 2: *In isolation, profitability as a dependent variable has a positive and significant impact on EV/EBITDA, with a t-statistic ≠ 0*
- Hypothesis 3: *In isolation, risk as a dependent variable has a negative and significant impact on* EV/EBITDA, *with a t-statistic ≠ 0*
- Hypothesis 4: *When accounting for differences amongst independent variables, relative performance in growth, profitability and risk jointly contributes with significant explanatory power in EV/EBITDA, with an F-statistic ≠ 0*

Research question 2: To what extent can OLS regression models utilising selected proxies, segmented by GICS codes and the SARD approach, accurately predict EV/EBITDA multiple valuations in congruence with observed multiple valuations?

- Hypothesis 5: *Estimated homoskedasticity will successively increase when moving from the market to GICS sectors, to GICS industries, to SARD groupings*
- Hypothesis 6: *Constructed multiple linear regression models, benchmarked against arithmetic averages of observed peer group multiples, will yield smaller and significant prediction errors in EV/EBITDA multiples against observed EV/EBITDA multiples*
- Hypothesis 7: *The accuracy of the predicted EV/EBITDA multiples derived from a regression analysis of fundamental value drivers will show a reduction in significant prediction error as the segmentation progresses from the market level to GICS sectors, to GICS industries, to SARD grouping*

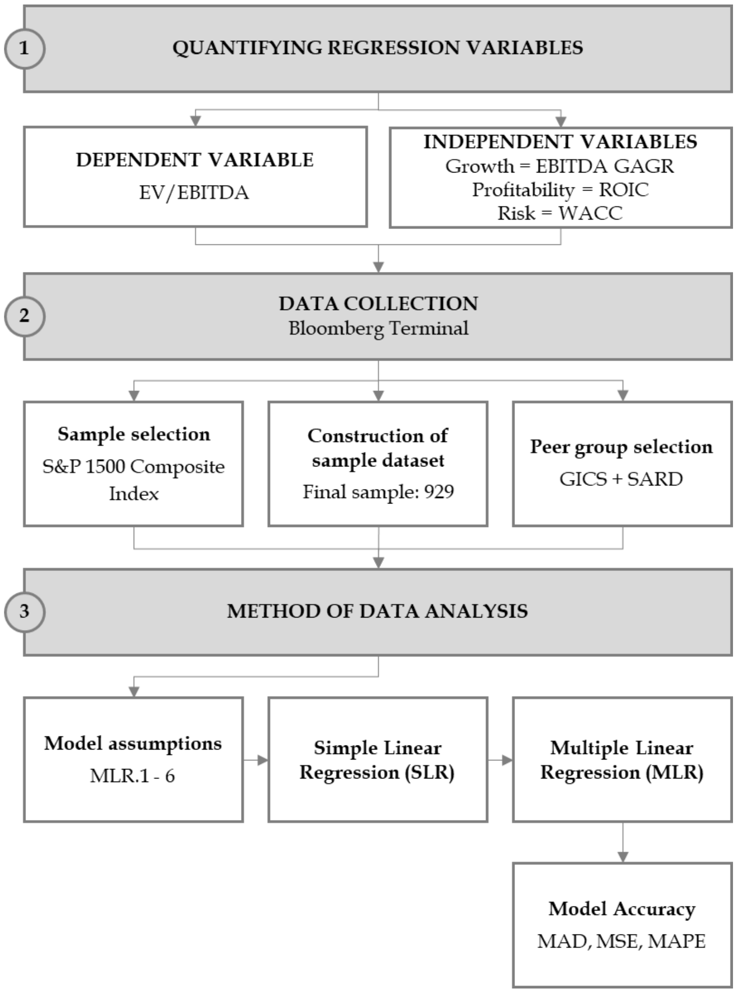
Chapter 5

Data & Methodology

The empirical study presented herein employs a research methodology that comprises five key sub-sections, guided by the workflow as per **Table 5.1**, with the objective of comprehensively addressing the research questions and hypotheses formulated. These sections are designed to incorporate the critical aspects associated with conducting accurate relative valuation studies.

Section 5.1 outlines the research paradigm of our paper. **Section 5.2** describes the variable definition and measurement of the dependent and independent variables included in the study. **Section 5.3** cover the data collection process, including sample selection, construction of the sample, analysis of outliers, and the peer group selection process. Subsequently, **section 5.4** outlines the research model specification, including the mathematical construct of the univariate regression model and the multivariate regression model. **Section 5.5** outlines the method of data analysis, including the underlying assumption of the multilinear regression model together with the accuracy tests applied. Please find **Table 5.1** on the next page for an overview.

Table 5.1 – Research design



5.1 Research Paradigm

The philosophy of science behind our quantitative study pertains to contextualising our research within the way scientific knowledge is produced, substantiated, and used in society (Holm, 2013). Through mapping out and reflecting on our data collection and knowledge generation, we can provide insight into the intrinsic, and more holistic, approaches of our paper when we seek to explore the relationship between EBITDA multiples and suggested key underlying value drivers.

We argue our research paradigm can most aptly be defined by positivism, with an inherent objectivist ontology and empiricist epistemology. Easterby-Smith et al. (2002) notes ontology as concerned with the nature of reality, which includes aspects such as what exists and what is real. Objectivism can be seen as the view that categories used in knowledge generation have an existence that is independent or separate from actors (Bahari, 2010). Davis et al. (1993) note that classic objectivists view science as accumulating knowledge in a progressive manner, that gets ever closer to the correct description of reality.

An empiricist epistemology argues that knowledge must be based on empirical evidence gathered through observation, experimentation, and testing, and not only through reasoning (Marczyk, DeMatteo, & Festinger, 2005)]. We draw key metrics from Bloomberg on S&P 1500 Composite Index constituents, which we in turn feed our Single Linear Regression (SLR) and Multiple Linear Regression (MLR) models and reflect on the implications of our observations.

Further, in line with a positivist approach, we apply a deductive methodology in that we aim to test existing theoretical propositions through observation (Holm, 2013). We draw on cross-sectional data whilst assuming efficient markets, which shows an element of naturalism, as we concern ourselves with identifying causality and that different phenomena can be explained by "natural" laws (Holm, 2013). Naturalism is characteristic of positivism.

5.2 Quantifying Regression Variables

The subsequent section presents the chosen dependent and independent variables for the applied research models, along with a detailed discussion of the mathematical derivation. Prior to elaborating on the specifics, it is worth noting that standard practice for most variables included in the study involved normalising their values. The normalisation is done to address and minimise potential biases resulting from collecting data from a single period. More specifically we have utilised a three-year period (2020-2022) to aggregate the EV/EBITDA, CAGR, and ROIC. WACC was obtained from the most recent fiscal year (LFY), due to limitations in data availability. Our objective is to employ a three-year average to account for cyclical fluctuations and to provide a more accurate representation of the relationship between the valuation multiple and the identified value drivers. Despite variations in the methods of normalising financial data, averaging values over time remains the most utilised approach in similar studies[20]. **Table 5.2** the variables included in the study, as well as their respective derivation.

[20] As outliend by Baker & Ruback (1999), Liu et al. (2002), and Plenborg and Pimentel (2016) in **section 3.1**

Table 5.2: Quantifying regression variables

	EV/EBITDA	Growth	Profitability	Risk
Financial metrics	Enterpise Value (FY) divded by Best EBITDA (FY+2), aggregated Hmean between '20-'22	Median EBITDA CAGR (FY1-FY3) between '20-'22	Median ROIC (FY) between 2020-2022	WACC
Timeline	FY + 2 (forward multiple)	FY + 3 (forward metric)	FY (current)	LFY (trailling)
Variable	Dependent	Independent	Independent	Independent

The independent variables utilised in this paper serve as proxies for the fundamental value driver of EV/EBITDA multiples. The analysis covers three independent value drivers, including growth, profitability, and risk. Hence, in line with the formulated research questions, the aim is to test the degree to which variance in these theoretically derived value drivers can explain the variance in EV/EBITDA multiple valuation. Furthermore, this paper seeks to compare the predictive power from the model regressed on the GICS industry affiliation approach with the model segmented using SARD grouping.

5.2.1 Dependent Variable

As explicated in section 1.4 and empirically substantiated in section 3.2.1 of this paper, the primary variable of interest is EV/EBITDA. To obtain the multiple estimations, the Bloomberg Terminal (2023) database has been used, which defines EV as the current market capitalisation of a firm's equity plus the market value of a firm's net interest-bearing debt. Furthermore, in this paper, EBITDA was found and defined as the two-year forward-looking BEst consensus estimates (Bloomberg Terminal, 2023). The formula for the forward-looking EV/EBITDA (EVE) is found in equation 5.1.

Equation 5.1

$$EVE_i = HM(EVE_{2020,i}, EVE_{2021,i}, EVE_{2022,i}) = HM\left(\frac{EV_{2020}}{EBITDA_{2022}}, \frac{EV_{2021}}{EBITDA_{2023}}, \frac{EV_{2022}}{EBITDA_{2024}}\right)$$

$$= \frac{3}{\frac{1}{\frac{EV_{2020}}{EBITDA_{2022}}} + \frac{1}{\frac{EV_{2021}}{EBITDA_{2023}}} + \frac{1}{\frac{EV_{2022}}{EBITDA_{2024}}}}$$

Where…

$HM = Harmonic\ Mean$

There are several other enterprise and equity multiples which could have been analysed, **section 3.1.1** in this study outlined various theoretical and empirical advantages that substantiate the decision to concentrate on the EV/EBITDA multiple. Whilst there is empirical evidence suggesting that EV/EBITDA is suboptimal for specific sectors and industries, this study will treat EV/EBITDA as a valid valuation metric for all industries but will examine the concept of industry best multiples by comparing and discussing the accuracy estimates across sectors and across industries.

Lastly, this study employs the EV/EBITDA multiple based on forward-looking earnings through BEst consensus estimates, which is an average of broker estimates of future earnings (Bloomberg Terminal, 2023). While forward-looking estimates have inherent biases and uncertainties, research has shown that forward-looking multiples, particularly those that consider estimates for two years into the future, generally provide more accurate estimates than trailing multiples (Yee K., 2004; Liu, Nissim, & Thomas, 2002; Begley & Feltham, 2002). Yee (2004) explains that historical earnings are typically transient and may lead to inaccurate predictions of a firm's future earnings. Therefore, using forward-looking estimates may yield better valuation estimates.

5.2.2 Independent Variables

Growth

Consistent with prior research examining the relationship between EV/EBITDA and its underlying value drivers, this study employs growth in EBITDA as the proxy for growth (Achleitner, Braun, & Engel, 2011; Hammer, Matter, Scheizer, & Wunsche, 2023; Damodaran A., 2012; Damodaran A., 2006). Specifically, the proxy is calculated as the compounded annual growth rate (CAGR) of Bloomberg's BEst EBITDA consensus estimates between the fiscal years FY+1 and FY+3, using the timing methodology proposed by Bernström (2014). To address cyclicality and time dependence, the estimate will be aggregated by taking the median between the base years 2020 and 2022 for the dependent variable. See equations 5.2-5.3 below.

Equation 5.2

$$Growth_i = M(CAGR_{i,2020}, CAGR_{i,2021}, CAGR_{i,2022})$$

Equation 5.3

$$Growth_i = M\left(\left(\frac{EBITDA_{2023}}{EBITDA_{2021}}\right)^{\frac{1}{2}} - 1, \left(\frac{EBITDA_{2024}}{EBITDA_{2022}}\right)^{\frac{1}{2}} - 1, \left(\frac{EBITDA_{2025}}{EBITDA_{2023}}\right)^{\frac{1}{2}}\right)$$

Where,

$M = Median$

$i = Firm\ i$

The compounded annual growth rate (CAGR) metric is widely adopted by scholars and practitioners. Much due to simplicity in calculating periodic averages (considering the influence of compounding effects), whilst minimising the repercussions of volatility (Pearl & Rosenbaum, 2009).

Profitability

Following the theoretical derivation of EV/EBITDA in **section 2.2.2**, and the theoretical logic outlined by Koller et al (2010) and Kinserdal et al. (2017) in **section 3.2.2**, this paper will use return on invested capital (ROIC) as a proxy for profitability. Also, to mitigate the effect of cyclicality, the proxy will be calculated as the median ROIC for 2020, 2021 and 2022. The rationale behind taking the median instead of the harmonic mean is substantiated by several firms having negative ROIC multiples, which could potentially distort the stability of the profitability ratio if taking the harmonic mean. See equation 5.4 below.

Equation 5.4

$$Profitability_i = M(ROIC_{i,2020}, ROIC_{i,2021}, ROIC_{i,2022})$$

Where,

$M = Median$

$i = Firm\ i$

The ROIC metric is extracted from Bloomberg, where it is calculated by dividing the trailing 12-month net operating profit after tax (NOPAT) by the average total invested capital. Derived from the last fiscal year data minus YTD data minus prior YTD Data. As stated in **section 3.1.1**, forward-looking estimates are considered to generally provide more accurate estimates than trailing multiples. To construct a forward-looking measure of ROIC, various manual approaches could have been made to estimate the constituent variables such as net operating profit after tax (NOPAT), net-interest-bearing debt (NIBD), and market value of equity (MVE). However, since Bloomberg does not provide forward-looking estimates for these variables, it is argued that the ROIC estimation based on current data provides a more dependable proxy than other alternatives. Although, using current data rather than forward-looking estimates for ROIC may cause inconsistency, the negative impact is mitigated to some extent by taking the 3-year historical median to approximate a more stable profitability ratio. Therefore, using this ROIC estimation, based on current data, as a proxy for profitability is considered more reliable and practical in the absence of forward-looking estimates of the relevant variables.

Risk

As concluded in **section 3.2.3** there is an academic and professional consensus which suggests that risk has a significant and inverse relationship to firm value (Berk & Demarzo, 2020). Hence, when following the theoretical derivation of EV/EBITDA in **section 2.2.2** and the empirical consensus outlined in **section 3.2.3**, this paper will use WACC as a proxy for firm risk. The firm's WACC is obtained from Blomberg for the most recent fiscal year (LFY) and is calculated as the weighted average cost of equity and the cost of debt, considering the tax shield.

Following the logic that WACC serves as a discount rate in the present value approach for determining the enterprise value of a firm, it is deemed highly relevant for our study, since we employ an enterprise-value-based multiple. Risk as a proxy is calculated by the below formula:

Equation 5.5

$$Risk_i = WACC_{i,LFY} = \frac{E}{E+D} * R_E + \frac{D}{E+D} * R_D * (1 - T)$$

5.3 Data Collection & Quality of Underlying Data

Bloomberg terminal was used as the data source for constructing the data sample. The Bloomberg terminal, which is also known as the Bloomberg Professional Services, is a software system which covers over 5 million bonds, equities, commodities, and currencies. The terminal provides services in Research, News, Collaboration Tools, Charts, Monitors, and Alerts, and has over 325,000 subscribers, worldwide (Bloomberg Terminal, 2023). The Bloomberg platform is widely recognised as one of the largest and most reliable sources of financial data for business- and investment-professionals. Furthermore, Bloomberg provides comprehensive forecast estimates on a wide range of variables, which are crucial for the variable constructs of this study. More specifically, the Bloomberg consensus estimates cover all key financial statements, including income statements, balance sheets, and cash flow measures, which are comparable to firms' reported results. (Bloomberg Terminal, 2023)

However, it is important to note that all estimates, including the dependent variable EV/EBITDA and the independent variable ROIC used in this study, are subject to measurement error due to their subjective nature. Although consensus estimates are used, the approximation of EV/EBITDA may be affected by noise, as the calculation of a firm's enterprise value[21] involves approximating the market value of debt using the book value of debt. Rossi & Forte (2016) argues that this approximation can result in inaccurate estimates due to variations in the composition, recognition, and accounting of debt on a firm's balance sheet. Additionally, the calculation of NOPAT and average invested capital, which are both used to approximate ROIC, are subject to subjective approximations. These discrepancies may introduce biases in the operationalisation of variables.

5.3.1 Sample Selection

The sample for our study comprises companies that constitute the S&P 1500, which is composed of 500 large-cap companies represented by the S&P 500, 400 mid-cap companies represented by the S&P MidCap 400, and 600 small-cap companies represented by the S&P SmallCap 600 (S&P, 2023).

Our sample selection draws from the S&P 1500 composite index, which includes 1505 publicly traded firms and has been widely used in accuracy research. This approach offers several benefits, including easy access to reliable financial data which all constituents must disclose as required by US regulatory mandates. Additionally, the large sample size enhances the efficiency of estimations and allows for sufficiently sized sub-samples to examine different research models at market, sector, and industry levels. Drawing from a single market, such as the US, mitigates increased heterogeneity from potential cross-geographical differences, thereby improving comparability and statistical inferences among companies. Adherence to GAAP fundamental accounting principles ensures that the measurement, recognition, and classification of accounting items are uniform across all firms in the sample. Further, selecting a sample from a single market also reduces heterogeneity among firms

[21] Enterprise Value = Market Capitalization + preferred stock + market value of debt + Minority interest – Cash and cash equivalents

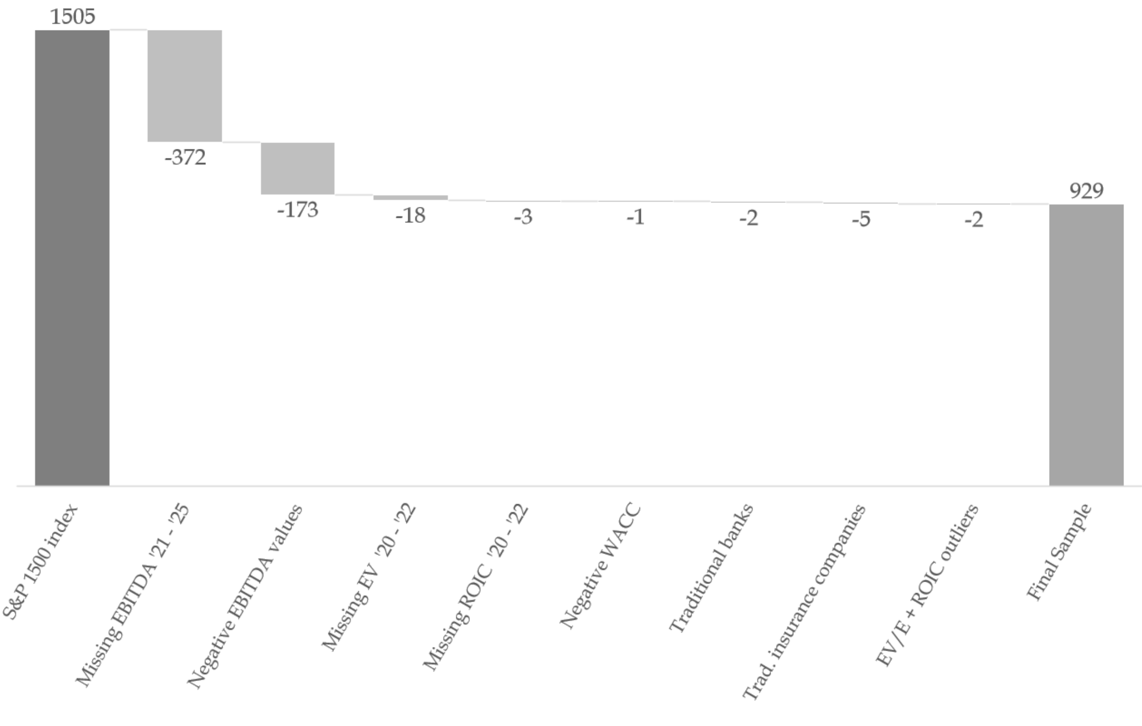
with respect to market factors, such as interest rates, inflation, and tax rates, which may distort the indicative value of valuation estimates over time.

Although our sample mitigates heterogeneity to some degree in several aspects, it is not entirely homogeneous, given the composite nature of the S&P 1500, which includes three different indices ranging from small cap to large cap, and substantial differences in relevant fundamentals. We will implicitly evaluate systematic variations among different sectors and industries, as well as firm heterogeneity and its potential impact on the predictive power of multiples, at different levels of analysis.

5.3.2 Construction of Sample & Dataset

The accounting data on each constituent of the S&P 1500 composite index was obtained from Bloomberg. The process of constructing the final data is outlined below and graphically visualised in **Figure 5.1**.

Figure 5.1 – Data filtering process



Firstly, we excluded observations of companies lacking- and or not possessing positive BEst EBITDA values from 2021-2025, to avoid quantitative errors in the proxy for growth. Then, we excluded companies lacking EV values between 2020-2022, thus this metric is necessary for all analyses made in this report. Furthermore, we removed firms lacking ROIC values between 2020-2022, thus these values serve as the fundament of the proxy for profitability. Also, we removed observations which showed a negative WACC for LFY 2022, thus a negative WACC would distort the assumed negative relationship between risk and EV. Thereafter, in line with prior studies (Harbula, 2009) we excluded both banking and insurance industries from our analysis since the EV/EBITDA metric may not accurately reflect the valuation of these types of firms, which have distinct accounting practices and capital structures. It is worth noting that while we removed traditional banks and insurance companies from the sample, we retained Diversified Financials Industry with GICS code 4020. This was substantiated by these firms differing from other financial companies in terms of balance sheet recognition of debt. Finally, to minimise the bias in our data, we excluded 2 firms which were identified as extreme outliers. These are presented in **Table 5.3**.

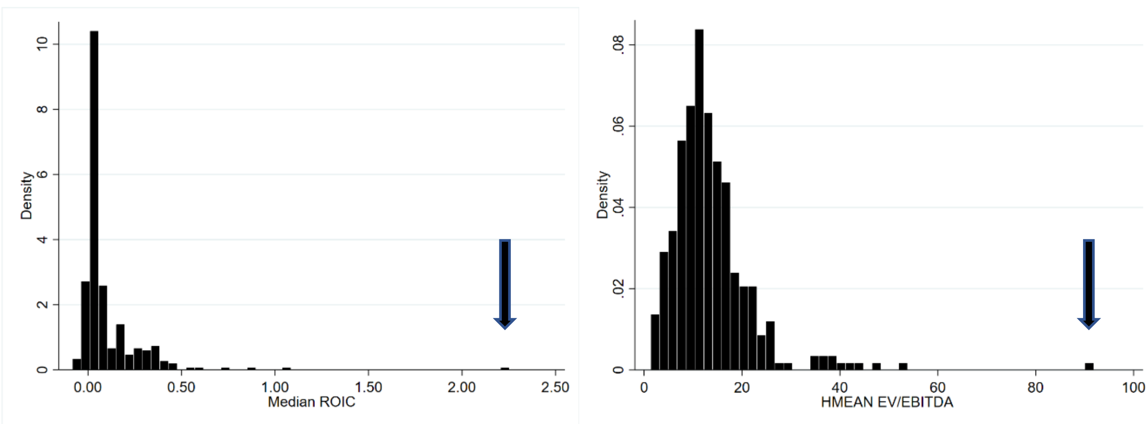
Table 5.3 – Overview of outliers excluded

Overview of outliers excluded

	GICS code	EV / EBITDA	EBITDA CAGR	ROIC	WACC
Insulet Corporation	3510	91,8x	32,4%	2,7%	9,3%
Verosign Inc.	4510	22,9x	9,1%	225,3%	9,6%

We excluded Insulet Corporation from the data sample due to its unusually high EV/EBITDA. Similarly, we removed Verisign Inc. from the sample as it was a potential outlier in terms of ROIC. Please refer to **Figures 5.2** and **5.3** for a visual representation of this.

Figure 5.2 & 5.3 – ROIC- and EV/EBITDA outliers



The exclusion of outliers follows the logic of Soliman (2008), on how OLS regression is highly sensitive to outliers because outliers could violate the assumption of normal distribution and constant variance of the errors, distort the calculation of residuals, and destabilise the estimate of the coefficients. While some scholars would argue that more than 2 observations in individual sectors and industry-sub samples could be considered outliers, we took a conservative approach. A decision motivated by the fact of other scholars have highlighted the statistical limitation of using smaller sample sizes.

5.3.3 Peer Group Selection

A pivotal facet of this research paper, in accordance with preceding studies, concerns the approach used to choose comparable firms, both on a methodological and a conceptual level. In this study, the selection criteria for different sets of comparable companies will be done using two different approaches, following the two schools of thought regarding peer group selection methodology. The first method is to select different peer groups according to the Global Industry Classification Standard (GICS), which identifies sectors with a 2-digit code and industries with a 4-digit code. As concluded in **section 3.3** by Fama & French (1997) the GICS industry affiliation reduced variation in valuation multiples and provided higher accuracy of valuation multiples. Eberhart (2004) and Bhojraj, Lee & Oler (2003) concluded similar findings, and suggest in their studies that GICS codes yield more accurate multiple valuations compared to other industry classifications.

Furthermore, in the domain of relative valuation, it is commonly held that larger peer groups yield more precise valuations due to the greater likelihood of accounting for the peculiarities among firms (Schreiner & Spremann, 2007). However, this advantage is only applicable when a target firm conforms entirely to the average sample performance, which is rarely the case in real practical scenarios. As such, a more refined industry categorisation with 4-digit codes, where firms share greater similarities in operating characteristics, would imply enhanced comparability and consequently more accurate prognostications (Schreiner & Spremann, 2007).

Despite the general assertion that more specific industry classifications lead to greater homogeneity among peer groups, it is widely recognised that heterogeneity will inevitably persist among firms operating in the same industry. Knudsen et al. (2017) argue that the use of GICS industry affiliation as a basis for comparison presupposes that firms operating within the same industry share identical economic attributes, namely, profitability, risk, and growth. However, it should be noted that firms within the same industry do not necessarily exhibit equivalent levels of their fundamental value drivers, and as such, should not be expected to trade at similar multiples.

Hence, the second method for peer selection is the Sum of absolute rank differences (SARD) approach, postulated by Knudsen, Kold and Plenborg (2017). The SARD approach deals precisely with the issue of firms within the same industry group not exhibiting equivalent levels of fundamental value drivers. Knudsen et al. (2017) concluded in their study that the approach presents significant advantages over the other methods of peer group selection. Firstly, the SARD approach is not constrained by the number of variables used to identify peers or the number of observations available, nor is it influenced by industry classification. Also, the selection variable in the SARD approach can be tailored to suit the requirements of any desired multiple, resulting in a valuation estimate that is arguably more precise.

It is important to note that the SARD approach is relatively new and has not been extensively tested. However, this factor adds to our curiosity to compare the statistical significance resulting from regression analyses conducted on the SARD-segmented peer groups against that of the GICS industry-segmented peer groups. While the principal objective remains to assess the feasibility of predicting the EV/EBITDA multiple through regression analysis of the underlying value drivers, namely profitability, risk, and growth., it will be of interest to observe whether the novel SARD approach can augment homogeneity among the chosen peer groups.

To facilitate the comparison of results between the SARD-segmented peer groups and the GICS industry-segmented peer groups, the SARD segmentation will comprise the same number of firms, namely 929, distributed across 22 peer groups, corresponding to the number of GICS industries. Each SARD peer group will include an equal number of firms corresponding to the specific GICS industry. For example, the GICS Energy Industry include 26 firms and therefore the SARD "energy grouping" will also include 26 firms.

The target company within each peer group will be selected by determining the median EV/EBITDA of each GICS industry and then identifying the firm closest to the median of the 3-year harmonic mean EV/EBITDA of the respective industry group. The firm having the median EV/EBITDA is deemed representative of that particular industry in terms of EV/EBITDA. Once the target firm is identified for each industry group, the SARD analysis is performed by ranking the full data sample of 929 firms across industries based on profitability, growth, and risk as variables. The target company's peer group is defined as the companies with the smallest sum of rank differences across the various proxies, including ROIC, EBITDA CAGR, and WACC. The number of peers in each SARD group is determined by the number of peers in the industry from which the target company was selected. This approach selects firms that are most similar to the target based on the chosen variables.

5.4 Research Model Specification

5.4.1 Simple Linear Regression (SLR) model

We seek to investigate the relationship between EV/EBITDA and the three fundamental value drivers, namely growth, profitability, and risk. In reference to **section 4**, Hypothesis Formulation, we wish to validate **hypotheses 1-7** by employing both univariate and multivariate regression models. Equations 5.6 through 5.8 represent the univariate regression models employed to analyse the respective association between the valuation multiple and the traditional value drivers.

Equation 5.6

$$\left(\frac{EV}{EBITDA}\right)_i = \alpha + \beta_i(Growth_i) + \varepsilon_i$$

Equation 5.7

$$\left(\frac{EV}{EBITDA}\right)_i = \alpha + \beta_i(Profitability_i) + \varepsilon_i$$

Equation 5.8

$$\left(\frac{EV}{EBITDA}\right)_i = \alpha + \beta_i(Risk) + \varepsilon_i$$

Where,

$i = Firm\ i$

$t = Time\ t$

The construction of the regression model depicted in equations 5.6-5.8 and its proxies need further quantification. As concluded in **section 5.2**, the chosen proxies rely on previous empirical research within the field of firm valuation. The 2-year compounded annual growth rate (CAGR) in EBITDA has been utilised as a proxy for growth. The return on invested capital (ROIC) has been examined as a proxy for profitability, while the firm's weighted average cost of capital (WACC) is utilised as a proxy for risk. By substituting the respective proxies for growth, profitability, and risk, the model illustrated in equations 5.9-5.11 are derived.

Equation 5.9

$$\left(\frac{EV}{EBITDA}\right)_i = \alpha + \beta_i(CAGR_i) + \varepsilon_i$$

Equation 5.10

$$\left(\frac{EV}{EBITDA}\right)_i = \alpha + \beta_i(ROIC_i) + \varepsilon_i$$

Equation 5.11

$$\left(\frac{EV}{EBITDA}\right)_i = \alpha + \beta_i(WACC_i) + \varepsilon_i$$

5.4.2 Multiple Linear Regression (MLR) Model

When testing for the joint significance of the traditional value drivers, we shall run the multivariate regression model illustrated in equation 5.12.

Equation 5.12

$$\left(\frac{EV}{EBITDA}\right)_i = \alpha + \beta_1(Growth)_i + \beta_2(Profitability)_i + \beta_3(Risk)_i + \varepsilon_i$$

Following the same methodology as in equations 5.9-5.11, we substitute Growth, Profitability, and risk with the same proxies.

Equation 5.13

$$\left(\frac{EV}{EBITDA}\right)_i = \alpha + \beta_1(CAGR)_i + \beta_2(ROIC)_i + \beta_3(WACC)_i + \varepsilon_i$$

5.5 Method of Data Analysis

The objective of this study is to assess the predictive ability of specific accounting information in determining a firm's value using multiple valuations, which relies on the EV/EBITDA multiples. To analyse the accuracy of these estimates, the study will compare them to observed market values of the company and relative to peer group multiples. A strong relationship between accounting information and market values indicates a level of coherence, while a weak relationship indicates the opposite. In summary, this research aims to determine to what extent certain accounting information can be used to predict a company's value, assess the accuracy of the valuation method relative to peer group averages, and finally evaluate if increased data sample homogeneity through SARD can lead to more statistically significant output.

The following sections will start by introducing the key concepts of the two primary regression models that have been developed, along with the methodology that has been used to assess the accuracy of these models. Next, we will discuss various statistical factors and potential limitations associated with the techniques that have been employed. Finally, we will provide a detailed explanation of the research model specification, highlighting the key elements of the developed models, and the methodology that has been utilised for testing their accuracy, with the aim of ensuring reproducibility.

5.5.1 Ordinary Least Squares (OLS) Regression

The application of univariate and multivariate OLS regression models is supported by the models' inherent interpretability. Model structures and parameters have been thoroughly covered in academia and are widely applied in a professional context – whereby linear regression models are considered to be interpretable. This stands in contrast to more comprehensive models such as machine learning models, which can potentially better explain complex underlying relationships, but lack rigorous definitions of interpretability. Univariate regression is applied to consider the relationship between EBITDA-multiple valuations and the underlying predictor variables growth, profitability, and risk, in isolation. Subsequently, we expand on the model to allow for all three predictor variables to be tested jointly, in a multivariate regression. Both regressions are applied across the market, GICS sector groups, GICS industry groups and groups compiled through applying the SARD approach with a basis in GICS industry groups. As mentioned, we have hypothesised the three predictor variables to be significant with T-statistics $\neq 0$, with a positive relationship to growth and profitability, and negative to risk. Furthermore, we hypothesise that the modelled underlying value drivers hold explanatory power when predicting EBITDA multiples. Data analysis, hereunder regressions, statistical tests and graphics, have been conducted in the statistical software STATA.

Standard error

Examining the standard error (SE) of the regression can provide additional insights into the performance of the model at the sector and industry level, relative to the market level. Ideally, we would expect the SE to decrease as we move from the market-wide regression to sector-wide, to the industry level. The SE for each regression model is calculated in matrix form using:

Equation 5.14

$$SE = \sqrt{\frac{\sum_{i=1}^{n} e_i^2}{n - 1 - k}}$$

Equation 5.14 shows that the SE is positively related to the sum of squared error ($\sum_{i=1}^{n} e_i^2$), and negatively related to the number of observations (n) and the number of independent variables (k). Therefore, when moving from the market level regression to the sector-, and industry level, the standard error may decrease. However, this is only true if the reduction in the sum of squared errors outweighs the negative effect of losing observations. I.e., the decrease in standard error depends on how much the model improves in terms of its explanatory power at the sector-, and industry-level. The average distance between the observed data points and the regression line is lower for sectors with a more homogenous sample and a narrower distribution of the EV/EBITDA multiples. Hence, it is expected that the model will exhibit superior accuracy performance for the SARD groupings.

Moving from market level to sector, industry, or SARD industry, will affect the standard errors of the regression coefficients. The standard error of the regression coefficients is calculated as the square root of the estimated variance of the error term (residual variance) times the variance-covariance matrix of the estimated coefficients. Where, the variance-covariance matrix is a square matrix that depicts the variances of the estimated regression coefficients on the diagonal, and the covariances between the estimated coefficients on the off-diagonal. See equations 5.15-5.16 below:

Equation 5.15

$$Var(\hat{\beta}) = S^2(X'X)^{-1}$$

Equation 5.16

$$SE(\hat{\beta}) = \sqrt{Var(\hat{\beta})}$$

Therefore, the standard errors of the estimated regression coefficients are adjusted on the degree of variation in the residuals of the regression model. I.e. when the standard error of the regression is large, the estimated standard errors of the regression coefficients will also be large, which exhibits greater uncertainty in the estimated coefficients.

R-squared

R^2, the coefficient of determination, is a statistical output measure from a regression model, which represents the variance in the dependent variable, which is explained by the independent variable. The performance of our regression model will be evaluated on several different metrics, where, R^2 is one of them. The regression R^2 takes a value between 0 and 1, and a large value indicates a good fit. However, when evaluating model performance, it is important not to do this solely based on the R-squared, but further contextualising with output and other test statistics. See formula 5.17 and the discussion below for the theoretical arguments of its applicability and limitations.

Equation 5.17

$$R^2 = \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n}(y_I - \bar{y})^2} = \frac{SSR}{SST} = 1 - \frac{SSE}{SST}$$

The regression R^2 is the square of the correlation between the fitted value $\hat{y}$ and y. A statistical test of the significance of R^2 can be performed using the F statistic. From the equation, R^2 represents the proportion of total variation in the Sum of Squared Total (SST) that is accounted for by the variation in the data that is explained by the regression model - Sum of Squared Residuals (SSR). When comparing the accuracy of the different regression models that will be conducted (Market, sector, industry & SARD grouping), it is important to be aware that the sample sizes of the dependent variables vary across the models being compared. When sample sizes vary in models being compared, the R^2 value may not be directly comparable. This is due to the range of values for the dependent variable can affect the R-squared value (Daoud, 2017). Hence, the performance of the various models will be evaluated on the combination of the coefficient of determination and prediction accuracy tests. The accuracy tests applied will be presented in **section 5.5.3**.

5.5.2 MLR.1–6, Classical Linear Model Assumptions

We will follow the methodology of Woolridge, J.M. (2019) when testing the inherent assumptions of both our simple linear regressions (SLR) and multiple linear regressions (MLR). The first 4 assumptions are shared by both SLR and MLR[22], whilst the added complexity of MLR models involves 2 additional assumptions. The assumptions can be summarised as follows:

Equation 5.18

$$y|\boldsymbol{x} \sim Normal(\beta_0 + \beta_! x_1 + \beta_2 x_2 + \cdots + \beta_k x_k, \sigma^2)$$

We need to understand the assumptions of our models (and whether the model deviates from them) to correctly comment on findings, remediate or adjust for shortcomings in the model, and to understand if further testing is needed. If the assumptions are satisfied, we should have unbiased coefficient estimates with model parameters and test statistics with higher validity and reliability (Woolridge, 2019), and we can effectively test our hypotheses on the statistical inference of the underlying value drivers against the EBITDA multiples.

MLR.1 – Linear in parameters

The first assumption of the model is that of linearity in the population model between the beta parameters, ε, denotes an unobserved random error):

Equation 5.19

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon$$

We apply the logarithm of the response variable to further linearise the relationship between the EBITDA multiples and the underlying value drivers. Fitting a linear model on highly non-linear

[22] And the first 5 assumptions are generally known as the Gauss-Markov assumptions for cross-sectional regressions

data can lead to predicted values that consistently deviate from observed values, which can affect significance tests and our ability to infer conclusions from our findings. We explore the linearity through scatterplots[23] of the response variable against the predictor variables and find there to be a seemingly linear relationship.

MLR.2 – Random sampling

The random sampling assumption can be violated in cases of missing data, non-random samples and outliers and influential observations (Woolridge, 2019). Our final dataset (see **Figure 5.1**) contains no missing data for the given parameters in the model, which has been drawn from publicly listed companies which are audited and/or highly covered in financial markets. As our scope pertains to firms in the S&P 1500, our sampling technique can most aptly be characterised as trimmed census sampling, where the data-filtering process has not unjustly excluded an unproportionate number of certain parts of the population, hence we do not have an issue of non-random samples.

Furthermore, our scope also implies that we do not explore the model's generalisability towards other financial markets, public or private. Lastly, as OLS regression minimises the sum of squared residuals, they can be sensitive to extreme observations, which especially holds true for smaller datasets (Woolridge, 2019). We have removed extreme outliers from the model, as illustrated in **Figure 5.1**.

MLR.3 – No perfect collinearity

Some correlation between the explanatory variables is to be expected. However, if any of the explanatory variables are exact linear combinations of the other explanatory variables, our model will suffer from perfect collinearity – hence cannot be estimated by OLS (Woolridge, 2019). Woolridge (2019) explains that this typically happens when one explanatory variable is a constant multiple of another, or if it can be expressed as an exact linear function of two or more of the other explanatory variables. Multicollinearity manifests itself in one or two ways: data-based multicollinearity from poorly designed experiments or purely observational data collection, or

[23] Please find scatterplot diagrams in **appendix 2**

structural multicollinearity from generating new independent variables from existing variables (Daoud, 2017).

Daoud (2017) further explains two signs of collinearity: high correlation between explanatory variables, and when t-tests are not significant, but the F-test for the whole model is significant. In **Appendix 3** we have listed correlation matrices for the market, GICS sector groups, GICS industries and SARD groupings. We argue neither of the above-proposed diagnostics holds for our data, and that the assumption of no perfect collinearity holds. In **Table 5.5** the degree of multicollinearity has been assessed by computing the Variance Inflation Factor (VIF) for our different segmentations, with SARD output marked in orange. The VIF formula is structured as per equation 5.20 and represents the marginal increase of variance in an explanatory variable when correlated with another variable. We draw on the R^2 from auxiliary regressions, whereby 3 MLRs are computed, each holding one different respective underlying value driver as the dependent variable, whilst holding the other underlying value drivers as independent variables (EVE is excluded).

Equation 5.20

$$VIF = \frac{1}{1 - R^2}, \qquad where\ R^2 = \frac{SSR}{SST} = \frac{\sum_{i=1}^{n} \hat{\varepsilon}_1^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}$$

…where SSR is the sum of squares for the residuals ($\hat{\varepsilon}$), and SST for the sum of squares total (y_i) is observed values, $\bar{y}$ is mean of observed values). Daoud (2017) defines the range for a degree of correlation as per below, where we see that all except one of our variables are moderately correlated:

Table 5.4 – Variance inflation factors

VIF-value	Correlation
VIF = 1	*No correlation*
1 < VIF ≤ 5	*Moderately correlated*
VIF > 5	*Highly correlated*

Table 5.5 – Variance inflation factors output

Auxiliary MLR Regressions - Testing Variance Inflation Factors (VIF) between the Different Independent Variables

		Dependent variable: Growth				**Dependent variable: Profitabilit**				**Dependent variable: Risk**			
	N	R^2		VIF		R^2		VIF		R^2		VIF	
Market	929	0.1358		1.1571		0.1823		1.2229		0.1926		1.2385	
Energy Sector	26	0.3761		1.6028		0.3708		1.5893		0.0150		1.0152	
Energy	26	0.3761	0.4184	1.6028	1.7194	0.3708	0.3709	1.5893	1.5896	0.0150	0.0979	1.0152	1.1085
Materials Sector	56	0.4540		1.8315		0.4765		1.9102		0.2156		1.2749	
Materials	56	0.4540	0.7163	1.8315	3.5249	0.4765	0.7737	1.9102	4.4189	0.2156	0.4312	1.2749	1.7581
Industrials Sector	159	0.2817		1.3922		0.3275		1.4870		0.1956		1.2432	
Capital Goods	109	0.3051	0.4774	1.4391	1.9135	0.3097	0.5403	1.4486	2.1753	0.1118	0.1706	1.1259	1.2057
Commercial & Professional Services	32	0.1279	0.4320	1.1467	1.7606	0.2963	0.7269	1.4211	3.6617	0.2109	0.6064	1.2673	2.5407
Transportation	18	0.2114	0.4067	1.2681	1.6855	0.5070	0.7579	2.0284	4.1305	0.4892	0.6922	1.9577	3.2489
Consumer Discretionary Sector	144	0.0834		1.0910		0.1033		1.1152		0.1549		1.1833	
Automobiles & Components	18	0.5875	0.5914	2.4242	2.4474	0.5968	0.7062	2.4802	3.4037	0.6879	0.6684	3.2041	3.0157
Consumer Durables & Apparel	37	0.0876	0.4851	1.0960	1.9421	0.2220	0.2691	1.2853	1.3682	0.2799	0.2459	1.3887	1.3261
Consumer Services	39	0.0543	0.2991	1.0574	1.4267	0.0107	0.2374	1.0108	1.3113	0.0583	0.4371	1.0619	1.7765
Retailing	50	0.1074	0.0065	1.1203	1.0065	0.0372	0.2661	1.0386	1.3626	0.1373	0.2628	1.1592	1.3565
Consumer Staples Sector	65	0.1042		1.1163		0.1404		1.1633		0.1144		1.1292	
Food & Staples Retailing	13	0.3763	0.6022	1.6033	2.5138	0.5275	0.6542	2.1164	2.8918	0.5990	0.7776	2.4938	4.4964
Food, Beverage & Tobacco	39	0.1342	0.1938	1.1550	1.2404	0.2097	0.5595	1.2653	2.2701	0.1533	0.5003	1.1811	2.0012
Household & Personal Products	13	0.1850	0.3807	1.2270	1.6147	0.0823	0.5165	1.0897	2.0683	0.1212	0.4075	1.1379	1.6878
Health Care Sector	123	0.0760		1.0823		0.0405		1.0422		0.0717		1.0772	
Health Care Equipment & Services	76	0.2277	0.3964	1.2948	1.6567	0.0901	0.4016	1.0990	1.6711	0.2625	0.3533	1.3559	1.5463
Pharmaceuticals, Biotechnology & Life Sciences	47	0.0545	0.3375	1.0576	1.5094	0.0527	0.4619	1.0556	1.8584	0.0084	0.2626	1.0085	1.3561
Financials Sector	40	0.1404		1.1633		0.2279		1.2952		0.3052		1.4393	
Diversified Financials	40	0.1404	0.2463	1.1633	1.3268	0.2279	0.5821	1.2952	2.3929	0.3052	0.5937	1.4393	2.4612
Information Technology Sector	138	0.1568		1.1860		0.1064		1.1191		0.1130		1.1274	
Semiconductors & Semiconductor Equipment	37	0.1050	0.2860	1.1173	1.4006	0.2469	0.3178	1.3278	1.4658	0.2041	0.1019	1.2564	1.1135
Software & Services	56	0.2501	0.4472	1.3335	1.8090	0.0977	0.5747	1.1083	2.3513	0.1801	0.5339	1.2197	2.1455
Technology Hardware & Equipment	45	0.2398	0.2420	1.3154	1.3193	0.1110	0.1633	1.1249	1.1952	0.2062	0.1452	1.2598	1.1699
Communication Services Sector	40	0.3055		1.4399		0.2643		1.3592		0.4538		1.8308	
Media & Entertainment	30	0.2946	0.4091	1.4176	1.6923	0.3066	0.6112	1.4422	2.5720	0.4418	0.6906	1.7915	3.2321
Telecommunication Services	10	0.7521	0.7632	4.0339	4.2230	0.0832	0.0382	1.0908	1.0397	0.7578	0.8252	4.1288	5.7208
Utilities Sector	52	0.4958		1.9833		0.5039		2.0157		0.2687		1.3674	
Utilities	52	0.4958	0.5730	1.9833	2.3419	0.5039	0.5340	2.0157	2.1459	0.2687	0.2070	1.3674	1.2610
Real Estate Sector	86	0.1043		1.1164		0.2128		1.2703		0.1614		1.1925	
Real Estate	86	0.1043	0.1806	1.1164	1.2204	0.2128	0.1761	1.2703	1.2137	0.1614	0.0645	1.1925	1.0689

Blue colour indicates VIF > 5 and high correlation

One may argue that some degree of correlation should be expected, given their relationship in financial theory. Using the CAPM formula[24], increasing risk will increase the beta of the investments, which increases the expected return of the investment and marginally increases a firm's growth, which will lead to higher profits if outgrowing expenses (Sharpe, 1964; Lintner, 1965; Mossin, 1966).

MLR.4 – Zero conditional mean

The 4th assumption holds that the error term ε has an expected value of 0 for any values of the independent values:

Equation 5.21

$$E(\varepsilon|x_1, x_2, \ldots, x_k) = 0$$

[24] $E[r_i] = r_f + \beta_{project}(E[r_M] - r_f)$

In other words, the error term should be uncorrelated with each of our explanatory variables, and we have exogenous explanatory variables, instead of correlated endogenous explanatory variables (Woolridge, 2019). Typical ways the assumption is violated is by omitting important factors or misspecification of the population model (see equation 5.18). Correlation between error terms and explanatory variables can lead to upwards bias inflating the model's goodness-of-fit, or downwards bias understating the model's goodness-of-fit. In the below figures, we find the residual plots of the three univariate models and the multivariate model, where we see that the residuals are apparently randomly distributed around the fitted line, suggesting that the assumption is met. However, we will never know for sure whether the average value of u is not related to our explanatory variables, but it remains a critical assumption (Woolridge, 2019).

Figure 5.2 – Scatter plot (EV/EBITDA)

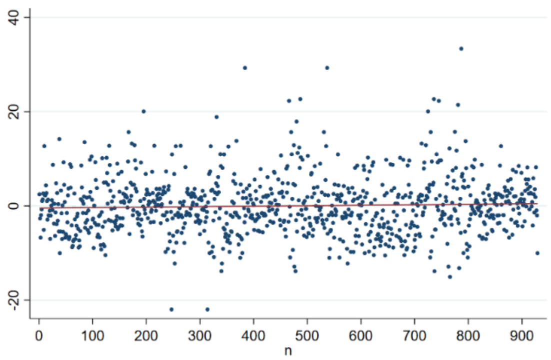


Figure 5.3 – Scatter plot (growth)

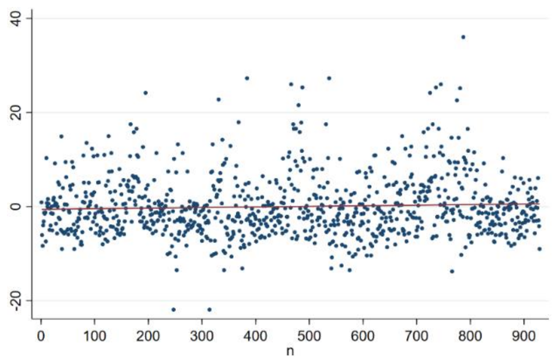


Figure 5.4 – Scatter plot (Profitability)

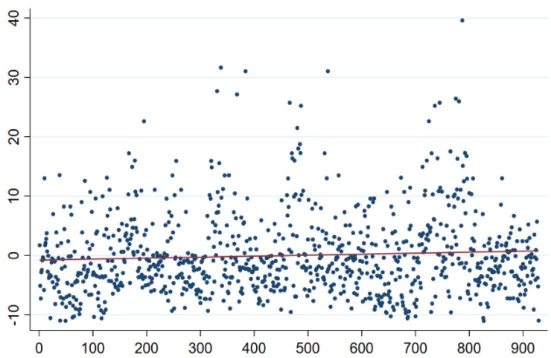
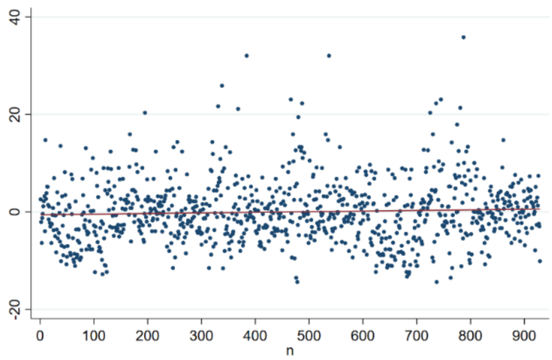


Figure 5.5 – Scatter plot (risk)



As the assumptions of MLR. 1-4 seems to be met, our OLS estimators should presumably not suffer from downward or upward bias.

MLR.5 – Homoskedasticity

Homoskedasticity is the assumption that the error terms, ε, have the same variance for all combinations of outcomes of the explanatory variables (Woolridge, 2019).

Equation 5.22

$$Var(\varepsilon|x_1, \dots, x_k) = \sigma^2$$

Heteroskedasticity may indicate that residual variances increase or decrease with changing estimated explanatory variables (Rousseuuw, 1987), whereby such non-constant variance can lead to faulty error terms non-representative of the model's uncertainty. Should we assume homoskedasticity when error terms are heteroskedastic, our statistical inference may be incorrect, given that we likely do not have the minimum variance unbiased estimators (Woolridge, 2019). To assess whether our error terms are homoskedastic or heteroskedastic, we have conducted White's test. White's test has a null hypothesis, H_0, of the error terms being homoskedastic, and an alternative hypothesis, H_1, that the error terms have unrestricted heteroskedasticity in an arbitrary manner. Findings in **Appendix 5** indicate homoskedastic error terms, with few exemptions, and the assumption of homoskedasticity is considered met.

MLR.6 – Normality

Normality is the assumption that the error terms must be independent of the explanatory variables, with a normal distribution with zero mean and variance σ^2(Woolridge, 2019):

Equation 5.23

$$\sigma^2 : \varepsilon \sim Normal(0, \sigma^2)$$

We have transformed our dataset through natural logarithms to potentially yield residual distributions closer to normality. Without normality, we cannot accurately construct confidence intervals and test hypotheses, and furthermore, it may be a sign that other MLR assumptions are violated (Woolridge, 2019). As per best practice, we conduct skewness and kurtosis normality tests, as well as the Shapiro-Wilks test. As per findings in **Appendices 6.1-6.2,** we see that the assumption of normality holds in 75% of all groupings (hereunder GICS sector groups, industry groups and SARD groupings), and we consider the assumption to hold for our dataset.

5.5.3 Model Accuracy

The forthcoming section will assess the efficacy of the developed regression model(s) in predicting the implied EV/EBITDA multiple across various market levels in S&P 1500. The different market levels are defined by GICS codes, where the whole S&P 1500 is defined as the market, followed by

GICS sectors, GICS industries, and lastly SARD groupings. Our objective is to evaluate the performance of the models in terms of their ability to predict EBITDA multiples accurately across different sectors and industries relative to peer group averages. To test the model's accuracy, we will employ various prediction methods. Our first objective of the conducted accuracy tests is to determine whether any particular sector or industry exhibits more significant performance in predicting the EV/EBITDA multiple. Our second objective is to determine whether the SARD industry groups will more accurately predict the EV/EBITDA multiple than compared to GICS industry groups. The prediction accuracy is analysed on a relative basis to peer group averages.

As starting point, the EV/EBITDA multiple will be predicted with the multivariate OLS regression as a base, across the market, GICS sectors, GICS industries, and SARD groups for each firm:

Equation 5.24

$$\widehat{EVE}_i = \hat{\alpha}_i + \hat{\beta}_1(Growth_i) + \hat{\beta}_2(Proftiability_i) + \hat{\beta}_3(Risk_i)$$

Since different coefficients will be generated for the proxies in the regression analysis for the various market segments, the predicted multiples will vary on market-, sector-, industry- and SARD grouping levels. To assess the accuracy of the predictions, three tests will be conducted. The objective of these tests is to measure the difference between the predicted multiples and the observed peer group multiples, at different levels of market segmentation. The three tests that will be performed are Mean Squared Error (MSE), Mean Absolute Deviation (MAD), and Mean Absolute Percentage Error (MAPE), as shown in equations 5.25-5.27.

Equations 5.25 *Equations 5.26* *Equations 5.27*

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2, \qquad MAD = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|, \qquad MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right|$$

The construction of the peer group (i) average (PGA) involves taking the harmonic mean of the observed median EV/EBITDA multiples of the different segmentations, across constituents x, y…

Equation 5.28

$$PGA_i = H\ Mean_i\left(Median(EVE_{2020,x}, EVE_{2021,x}, EVE_{2022,x}) + \cdots + Median(EVE_{2020,y}, EVE_{2021,y}, EVE_{2022,y})\right)$$

Chapter 6

Empirical Results

This chapter will present the statistical findings of the empirical research conducted and will be structured as follows: **Section 6.1** reports the overall summary statistics of the data, together with the correlation among the included variables. This section will also thoroughly depict the distribution of the variables between market, sectors, industries and SARD groupings. **Section 6.2.1** seeks to outline the empirical findings in relation to **Hypothesis 1-3**, i.e., to identify the degree to which variance in EV/EBITDA multiple can be explained by the variance in its theoretically derived value drivers: growth, profitability, and risk, in isolation. **Section 6.2.2** seeks to outline the empirical findings in relation to **Hypothesis 4** i.e., whether the relative performance in the theoretically derived underlying value drivers jointly have a significant impact on EV/EBITDA. **Section 6.3** evaluates the output from White's test and outlines the degree of heteroscedastic in the error terms across our different groupings. Furthermore, **section 6.4** will evaluate the performance of the different regression models, in terms of how accurately they predict the observed EV/EBITDA multiples relative to peer group averages, for the market, sectors, industries and SARD groupings. This section seeks to outline the findings in relation to **Hypothesis 6** and **Hypothesis 7**. Moreover, in **section 6.5** we run a regression diagnostic to test for potential sporadic curvature of WACC, through a shifted power transformation. In **Section 6.6** we re-conducted the multivariate analysis on a market level for 4 additional years, to test for robustness in the models. Finally, **section 6.7** examines the results presented in the previous sections and summarises which hypotheses that are confirmed or rejected. This chapter serves as a foundation for the discussion in **chapter 7**.

6.1 Summary Statistics

Table 6.1 – Summary statistics and correlation matrix

Market

Sample: Trimmed census S&P 1500
Number of firms: 929

A) Summary statistics						B) Correlation matrix			
Variable	Mean	Median	Std. dev.	10P	90P	LN(EVE)	CAGR	ROIC	WACC
EVE	12.4583	11.2109	6.8386	5.4897	20.8065	1.0000			
CAGR	0.0902	0.0694	0.1706	-0.0573	0.2358	0.3562****	1.0000		
ROIC	0.1115	0.0875	0.1011	0.0226	0.2321	0.0300	-0.1974****	1.0000	
WACC	0.0810	0.0795	0.0155	0.0630	0.1020	0.3672***	0.2292****	0.3196****	1.0000

**** p<0.01
*** p<0.025
** p<0.05
* p<0.1

Table 6.1 shows the cross-sectional information about the statistical distribution of the analysed variables, together with the correlation among them. Subset A) depicts the summary statistics for the dependent variable EV/EBITDA, and the independent variables growth, profitability, and risk. Subset B) displays a correlation matrix between the EV/EBITDA multiple and the independent variables. It is worth noting that the dependent variable LN(EVE) in subset B), is the natural logarithm of EV/EBITDA. As outlined in **section 5.5.2** "MLR-6", this transformation is conducted to yield residual distribution closer to normality (Woolridge, 2019).

As described in **section 5.3** the S&P 1500 composite index is composed of S&P 500, S&P MidCap 400, and S&P SmallCap 600. As of 16th of March 2023 S&P 500 index consists of the 500 largest companies which capture 92.3% of the total S&P 1500 market capitalisation. The large constituents cause a right skewness of the distribution in metrics such as sales, EBITDA, NOPAT, invested capital, book value of equity, EV, and net debt. A similar distribution was found by Plenborg et al. (2017). Our study ought to increase the normality by only working with relative metrics. Nevertheless, as seen in **Table 6.1**, subset A), the means of the dependent and the independent variables are greater than the medians, which suggests a positively skewed distribution. For multiples, this distribution is consistent with the finding of Lie and Lie (2002) and Plenborg et al (2017). Moreover, when looking at the interquartile range, one can note that it ranges from 5.5x to 20.8x. Given a median of 11.2x, the longer distance to the 90% quartile versus the 10% quartile further indicates that the distribution is rightly skewed.

For the explanatory variables, it can be seen in subset A that EBITDA CAGR has a mean of 9%, with the largest standard deviation of 0.1706, and the largest interquartile range from -5.7% to 23.6%. The distribution for ROIC is fairly similar with a mean of 11%, and standard deviation of 0.1011, and a similar interquartile range from 2% to 23%. The distribution of WACC has a mean of 8%, a standard deviation of 0.0155 and an interquartile range that varies from 6% to 10%. It can be concluded that firms in the trimmed census S&P 1500 have a high variation in terms of growth and profitability and a relatively lower variation in terms of WACC.

Table 6.2 and **Table 6.3** provide a visual representation of the distribution of variables across market segments, including sectors, industries, and SARD groupings. **Table 6.2** shows the mean EV/EBITDA multiples for each GICS sector, indicating that the Health Care, Information Technology, and Real Estate sectors have the highest mean EV/EBITDA multiples, with multiples of 15.8x, 14.6x, and 16.0x, respectively. The three GICS sectors with the lowest EV/EBITDA multiples are Energy, Materials, and Consumer Discretionary, showing multiples of 4.4x, 9.3x and 9.9x, respectively.

Upon examining **Table 6.2**, on a sector level, it can be seen that the Information and Technology sector exhibits the highest median EBITDA CAGR, while the Consumer Discretionary sector exhibits the lowest, with 11.2% and 2.8% respectively, which supports hypothesis 1. Additionally, the Information Technology sector had the highest median ROIC, while the Real Estate sector had the lowest, with 13.3% and 3.8% respectively. When considering the median WACC, it can be observed that the Information Technology sector had the highest level of risk, while the Utilities sector had the lowest, with a WACC of 9.2% and 6.25% respectively.

From examining the correlation matrix in **Table 6.1**, subset B, it is evident that there is a strong correlation among all three independent variables on a 1% significance level. On a market level, we note that CAGR and ROIC have a significant and positive correlation with risk. This follows the theory of how firms with higher growth rates and higher returns are likely to be riskier. It can further be noted that profitability and growth have a negative and significant relationship, which is interesting given the literature suggesting the opposite.

Moving from the market level to the sector- and industry level, the summary statistics can be found in **Appendix 3**. The summary statistics, and the correlation between the dependent and independent variables varies largely across all sectors and industries. 6 out of the 11 sectors exhibit a negative correlation between LN EV/EBITDA and ROIC but only 3 of these on a 1% significance level. For the materials sector, the LN EV/EBITDA and ROIC correlation was -0.4325 on a 1% significance level. This is an interesting finding since theory suggests that profitability has a positive and significant relationship with firm value. For WACC, 9 out of the 11 sectors exhibit positive correlation, with 6 of the sectors being significant on a 1% level, one sector significant on a 5% level, and two sectors significant on a 10% level. This also disputes the theory that WACC has a negative correlation with firm value. EBITDA CAGR, on the other hand, exhibits a positive correlation with EV/EBITDA for all sectors, with 7 of them being significant on a 1% level. This is arguably in line with theory and confirms **hypothesis 2**, that growth has a positive and significant impact on EV/EBITDA.

At first glance the initial examination of the summary statistics and variable distribution across market segments, it is difficult to draw conclusions regarding the confirmation or rejection of **hypotheses 1-3**. However, when analysing the data at a more granular level by grouping companies into sectors and industry groups according to GICS classification codes provides further insights. **Appendix 3** shows that the direction and significance of the relationships vary greatly across sectors and industries.

Appendix 4 depicts the summary statistics, including the statistical distribution of the analysed variables and the correlation among them for the SARD grouping. On a general level, the output from the statistical analysis done on the SARD industry segmentation differs quite significantly from the GICS industry affiliation output (see **Table 6.3**).

We find in **Table 6.3** that the three highest EV/EBITDA multiples for the SARD groupings are Media & Entertainment (20.43x), Technology Hardware & Equipment (18.65x), and Software & Services (17.08x). The three groupings showing the lowest EV/EBITDA are Food, Beverage & Tobacco (8.75x), Transportation (9.29x), and Materials (9.69x). We find the highest CAGR for groupings the following SARD groupings: Energy (26.52%) and Health Care Equipment & Services (25.49%), and the lowest for groupings based on Transportation (-4.00%) and Semiconductors & Semiconductor Equipment (-2.72%). The highest median ROIC can be seen for Transportation (33.62%), Semiconductors & Semiconductor Equipment (25.94%), and Consumer Durables & Apparel

(23.66%), which also holds one of the highest median EV/EBITDA multiples. The spread for median WACC for SARD groupings is lower than the GICS industries, with the highest median WACC seen for Media & Entertainment (10.58%) which also has the highest median EV/EBITDA multiple of the different groupings, and the lowest median WACC seen for Telecommunication Services (5.99%). For the different correlation matrices, we see more positive correlations between the underlying value drivers and the EBITDA multiple, with varying significance levels.

Table 6.2 – Variable distribution for GICS segmentation

Central tendency, standard deviation, and coefficient of variance, across market and GICS segmentations

		EV/EBITDA		Growth (%)		Profitability (%)		Risk (%)		CoV*
	Firm count	Harm. Mean	Std.Dev.	Median	Std.Dev.	Median	Std.Dev.	Median	Std.Dev.	
Market	929	12.46	5.87	7.29	15.15	9.48	8.66	8.00	1.07	0.617072
Energy Sector	26	4.35	2.59	8.85	24.93	5.28	6.37	7.71	0.83	1.993555
Energy	26	4.35	2.59	8.85	24.93	5.28	6.37	7.71	0.83	1.993555
Materials Sector	56	9.30	5.10	7.35	16.55	10.19	8.82	8.18	1.29	0.853641
Materials	56	9.30	5.10	7.35	16.55	10.19	8.82	8.18	1.29	0.853641
Industrials Sector	159	11.57	5.40	6.55	12.79	9.89	8.52	8.29	1.02	0.59945
Capital Goods	109	11.68	5.05	7.42	15.18	10.08	8.21	8.46	0.91	0.627988
Commercial & Professional Services	32	12.51	8.13	6.06	6.59	8.12	8.67	7.86	1.25	0.492381
Transportation	18	9.02	2.48	1.91	9.22	11.97	10.23	7.98	1.25	0.641955
Consumer Discretionary Sector	144	9.86	5.05	2.80	11.92	11.40	9.87	8.13	1.36	0.714492
Automobiles & Components	18	8.38	5.33	12.25	14.80	6.52	6.23	7.89	1.78	0.839838
Consumer Durables & Apparel	37	8.96	4.22	1.68	11.22	13.54	12.12	8.59	1.31	0.80576
Consumer Services	39	12.71	6.62	6.53	13.26	10.28	11.17	7.72	1.27	0.635315
Retailing	50	8.89	4.35	-2.72	10.35	12.43	8.49	8.19	1.32	0.68945
Consumer Staples Sector	65	12.78	5.04	6.14	11.34	9.63	8.03	6.96	0.97	0.496251
Food & Staples Retailing	13	9.91	3.80	4.60	9.78	8.50	4.64	7.38	0.84	0.480653
Food, Beverage & Tobacco	39	13.24	5.38	6.93	11.43	8.93	9.43	6.80	0.98	0.514042
Household & Personal Products	13	14.28	5.25	5.33	12.64	12.86	7.19	7.03	1.07	0.457595
Health Care Sector	123	15.82	8.81	10.90	20.00	8.98	8.93	7.77	1.61	0.621931
Health Care Equipment & Services	76	16.38	9.49	11.88	18.55	7.41	8.74	7.81	1.60	0.585535
Pharmaceuticals, Biotechnology & Life Sciences	47	14.89	7.69	9.28	22.41	11.57	9.25	7.70	1.63	0.688115
Financials Sector	40	11.06	6.72	5.08	14.79	10.70	8.58	8.54	1.77	0.720328
Diversified Financials	40	11.06	6.72	5.08	14.79	10.70	8.58	8.54	1.77	0.720328
Information Technology Sector	138	14.63	7.86	11.22	22.58	13.33	14.48	9.22	1.62	0.795443
Semiconductors & Semiconductor Equipment	37	12.64	7.05	7.90	22.79	14.40	11.67	10.47	1.60	0.852491
Software & Services	56	17.89	9.09	14.81	25.65	13.13	16.43	9.00	1.92	0.741909
Technology Hardware & Equipment	45	12.20	7.01	9.47	18.59	12.71	14.35	8.48	1.28	0.844504
Communication Services Sector	40	11.37	4.58	7.52	13.52	7.64	6.48	7.32	1.59	0.575307
Media & Entertainment	30	11.63	4.91	7.78	15.46	9.39	5.91	7.56	1.68	0.601166
Telecommunication Services	10	10.61	3.61	6.75	7.70	2.42	8.19	6.62	1.29	0.490248
Utilities Sector	52	11.65	2.88	7.66	6.42	4.98	2.30	6.25	0.55	0.260704
Utilities	52	11.65	2.88	7.66	6.42	4.98	2.30	6.25	0.55	0.260704
Real Estate Sector	86	16.02	4.89	5.64	11.18	3.83	3.23	7.47	0.90	0.315333
Real Estate	86	16.02	4.89	5.64	11.18	3.83	3.23	7.47	0.90	0.315333

* CoV = Coefficient of variance, calculated as (average std. dev. all proxies)/(Hmean EV/EBITDA). Gives a comparable view of std. dev. across segmentations
- Trimmed census S&P 1500 Composite Index, final sample: 929
- 2020-22 (estimation from 2020-2024)
- GICS sector = 2-digits, GICS industry = 4 digits (e.g., energy sector = 10, energy industry = 1010)

Table 6.3 – Variable distribution SARD groupings

Central tendency, standard deviation, and coefficient of variance, across SARD segmentations

		EV/EBITDA		**Growth (%)**		**Profitability (%)**		**Risk (%)**		**CoV***
	Firm count	Harm. Mean	Std.Dev.	Median	Std.Dev.	Median	Std.Dev.	Median	Std.Dev.	
Energy	26	12.93	6.95	26.52	25.91	1.10	2.87	8.11	0.45	0.699634
Materials	56	9.69	3.91	-0.07	7.09	6.81	1.74	6.77	0.52	0.342132
Capital Goods	109	10.61	5.88	-0.72	8.47	19.23	11.13	8.45	0.99	0.623426
Commercial & Professional Services	32	11.46	4.53	8.62	2.20	8.76	1.67	6.93	0.41	0.192263
Transportation	18	9.29	4.60	-4.00	8.66	33.62	15.42	8.47	0.39	0.782655
Automobiles & Components	18	11.61	4.53	14.59	6.49	2.18	1.07	6.49	0.56	0.272368
Consumer Durables & Apparel	37	16.98	6.63	8.83	2.52	23.66	7.44	9.51	0.65	0.253961
Consumer Services	39	10.49	3.36	4.42	1.79	7.48	1.76	6.20	0.55	0.178095
Retailing	50	10.17	5.63	-2.55	8.32	18.82	6.46	9.30	0.61	0.516463
Food & Staples Retailing	13	10.66	5.40	6.00	1.03	8.59	1.66	7.47	0.30	0.196631
Food, Beverage & Tobacco	39	8.75	2.47	2.54	3.84	9.00	1.96	6.27	0.69	0.25571
Household & Personal Products	13	9.96	2.36	15.88	3.58	3.44	0.95	6.04	0.34	0.181504
Health Care Equipment & Services	76	14.93	8.91	25.49	26.59	2.35	2.99	8.55	1.11	0.66291
Pharmaceuticals, Biotechnology & Life Sciences	47	13.05	5.41	11.96	6.22	18.71	6.92	7.72	0.61	0.366912
Diversified Financials	40	10.14	3.81	0.71	7.41	4.18	1.98	6.00	0.67	0.341759
Semiconductors & Semiconductor Equipment	37	10.89	6.45	-2.72	11.19	25.94	8.89	9.78	0.93	0.630133
Software & Services	56	17.08	8.02	8.85	3.45	23.11	9.85	9.65	1.10	0.328344
Technology Hardware & Equipment	45	18.65	10.04	21.44	15.46	10.93	3.34	10.21	1.21	0.402845
Media & Entertainment	30	20.43	9.95	15.95	9.63	22.62	11.18	10.58	1.24	0.391724
Telecommunication Services	10	10.69	4.67	1.53	1.34	1.93	1.67	5.99	0.76	0.197271
Utilities	52	11.36	3.92	14.20	8.10	3.83	1.61	6.36	0.55	0.312094
Real Estate	86	11.72	5.95	17.89	17.98	2.54	2.67	7.00	0.73	0.583006

* CoV = Coefficient of variance, calculated as (average std. dev. all proxies)/(Hmean EV/EBITDA). Gives a comparable view of std. dev. across segmentations
- Trimmed census S&P 1500 Composite Index, final sample: 929
- 2020-22 (estimation from 2020-2024)
- SARD groupings based on lowest sum of absolute ranked differences in growth, profitability and risk

Furthermore**, figures 6.1-6.4** confirm that the distribution of EV/EBITDA, growth, profitability, and risk is positively skewed, meaning that the values are concentrated towards the lower end of the distribution and have a longer tail towards the higher end. This is consistent with the fact that the mean value is higher than the median value for these variables. Moreover, the figures show that EV/EBITDA, ROIC, and CAGR have more outliers than WACC. Additionally, the visual representations demonstrate that the distribution spread of EBITDA CAGR and ROIC is much wider than the distribution spread for risk.

Figures 6.1-6.4 – Variable distribution histograms

Variables distribution histograms

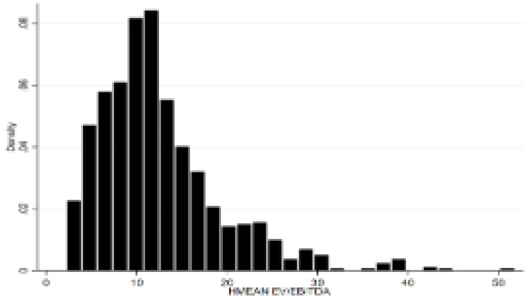


Figure 6.1 - MTPL

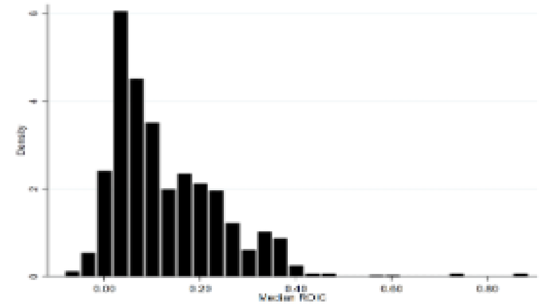


Figure 6.2 - Profitability

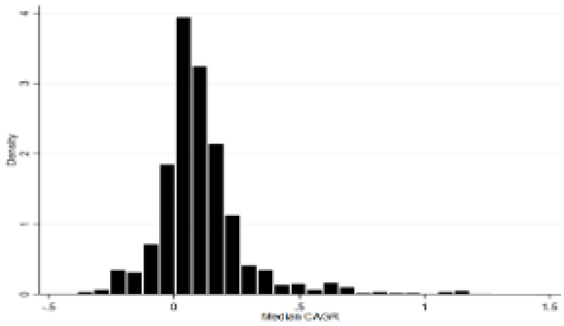


Figure 6.3 - Growth

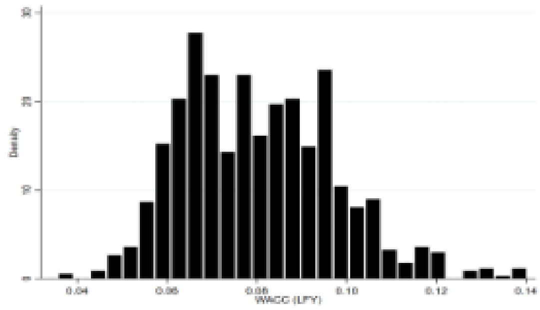


Figure 6.4 - Risk

6.2 Linear regression output

6.2.1 Simple Linear Regression (SLR)

We have laid out the theoretical underpinnings, methodology and assumptions for our quantitative analysis of the trimmed census sample of the S&P 1500 Composite Index, as well as key summary statistics. This section will present the findings of our SLR exploring the relationship between the EBITDA-multiple and the proxies for growth, profitability, and risk, with output as per **Table 6.4**. The table shows relevant intercepts, coefficients, and t-statistics by market, GICS sector group, GICS industry group as well as SARD groupings[25].

Rooted in our theoretical underpinnings, we hypothesised (**Hypothesis 1-3**) that growth and profitability would have a positive and significant impact, and risk a negative and significant

[25] In **table 6.4**, the output from the SARD grouping is marked in orange, and positioned by relevant GICS industry groups

impact, on the EBITDA-multiple valuation with a t-statistic ≠ 0. As the t-statistics measures how many standard errors away a coefficient is from 0, values of 0 would imply a coefficient statistically equal to 0, and an insignificant relationship between the EBITDA-multiple and underlying value driver, within the relevant segmentations.

Our SLR on the whole market models a positive intercept for our growth proxy of 2.292 and coefficient of 0.01 with $p < 0.01$[26]. This tells us that the model predicts that if a firm has seen a growth rate (in EBITDA CAGR) of 0%, their EBITDA-multiple will equal ln(EV/EBITDA) of 2.292x = EV/EBITDA of 8.895x. On a market level the beta coefficient for growth is 0.01. Hence, a one percentage unit increase in growth rates will lead to a ~1.005% increase in the EBITDA multiple.[27] On a GICS sector level, the beta coefficient for growth is significant at $p < 0.01$ for 7 out of 11 sectors, whereas on a GICS industry level, the beta coefficient for growth is significant at $p < 0.01$ for 10 out of 22 industries. On a SARD grouping level, 7 groupings are significant at $p < 0.05$, where 3 of these are significant at $p < 0.01$. For groups where both SARD groupings and industry groups have coefficients with $p < 0.01$, SARD groupings have higher positive relationships (e.g. as for Diversified Financials, Capital Goods, Semiconductors & Semiconductor Equipment and Materials). Furthermore, the intercept coefficient is significant at $p < 0.01$ for all GICS- sectors and industries. Similarly, for the SARD grouping all except one is significant at $p < 0.01$.

For profitability, our SLR on the whole market models a positive intercept coefficient of 2.38 and a beta coefficient of 0.002 with $p < 0.01$ for the intercept coefficient, but $p > 0.1$ for the beta coefficient. This tells us that the model predicts that if a firm has seen a ROIC of 0%, their EBITDA-multiple will equal ln(EV/EBITDA) of 2.38 = EV/EBITDA of 10.8 and that a one percentage unit increase in ROIC will lead to a 0.2% increase in the EBITDA multiple. The intercept coefficient is positive and statistically significant with $p < 0.01$ for all different groupings, with intercepts generally being higher for SARD groupings than GICS industry groups. Sector- and industry groups hold approximately as many beta coefficients with $p < 0.1$, with the Materials Sector and Utilities Sector having beta coefficients with $p < 0.01$, and 2 industries having beta coefficients with $p < 0.01$ (Consumer Services, Materials). Only 2 SARD groupings have statistically significant profitability

[26] We use p-values and significance level interchangeably

[27] $\Delta_{proxy} = one\ percentage\ unit \rightarrow \Delta_{EV/EBITDA} = \left(exp(\bar{y}_{proxy}) - 1\right), where\ \bar{y} = coefficient$

coefficients: Energy with a coefficient of 0.107 at $p < 0.05$, and Health Care Equipment & Services with a coefficient of 0.06 at $p < 0.05$.

For risk, our SLR on the whole market models a positive intercept coefficient of 1.65 and a beta coefficient of 0.09, with both being significant at $p < 0.01$. This tells us that the model predicts that if a firm has a WACC of 0% (which to note is an arbitrary example – given that no companies operate with zero capital cost), their EBITDA-multiple will equal ln(EV/EBITDA) of 1.65 = EV/EBITDA of 5.21 and that 1 percentage unit increase in WACC will lead to a 9.42% increase in the EBITDA multiple. We note that this is higher than the beta coefficient for profitability, at a market level. On a GICS sector level, the beta coefficient is positively significant at $p < 0.05$ for 6 out of 11 sectors, and negatively significant at $p < 0.05$ for only 1 sector[28]. On a GICS industry level, the beta coefficient is positively significant at $p < 0.05$ for 10 out of 22 industries, and negatively significant at $p < 0.05$ for 1 industry. The intercept coefficient is positive and significant at $p < 0.05$ for all sectors and for 17 out of 22 industries[29]. For the SARD groupings, only 5 groupings have a statistically significant beta coefficient at $p < 0.05$, four of which have a positive coefficient[30] and one of which has a negative coefficient[31]. Lastly, 12 out of 22 SARD groupings have a positive and significant intercept coefficient at $p < 0.05$. Please refer to **Table 6.4** for a full overview of the simple linear regression output.

[28] The Energy sector/industry is the only industry depicting a negative beta coefficient, with $p < 0.05$

[29] The intercept coeffient is not significant at $p < 0.05$ for Transportation, Automobiles & Components, Food & Staples Retailing, Technology Harware & Equipment, and Telecommunication Services

[30] The beta coefficient of risk is positive and significant for Consumer Durables & Apparel, Health Care Equipment & Services, Technology Hardware & Equipment, and Telecommunication services at $p < 0.05$

[31] The beta coefficient for risk is negative and significant for Food & Staples Retailing, at $p < 0.05$

Table 6.4 – Simple linear regression (SLR) output

Simple linear regression (SLR) output

		Subset A: Growth				**Subset B: Profitability**				**Subset C: Risk**			
	N	Intercept		Coefficient		Intercept		Coefficient		Intercept		Coefficient	
Market	929	2.292****		0.01****		2.38****		0.002		1.65****		0.09****	
		(119.587)		(10.252)		(88.769)		(0.14)		(17.86)		(8.069)	
Energy Sector	26	1.257****		0.002		1.347****		-0.014		5.525****		-0.553****	
		(8.299)		(0.417)		(8.138)		(-0.653)		(5.723)		(-4.416)	
Energy	26	1.257****	2.488****	0.002	-0.003	1.347****	2.346****	-0.014	0.107***	5.525****	0.251	-0.553****	0.26
		(8.299)	10.553	(0.417)	-0.537	(8.138)	19.847	(-0.653)	2.561	(5.723)	0.105	(-4.416)	0.888
Materials Sector	56	2.011****		0.014****		2.394****		-0.028****		2.346****		-0.032	
		(27.865)		(3.388)		(21.408)		(-3.525)		(4.718)		(-0.537)	
Materials	56	2.011****	2.224****	0.014****	0.021****	2.394****	2.296****	-0.028****	-0.016	2.346****	2.194****	-0.032	-0.001
		(27.865)	40.781	(3.388)	2.832	(21.408)	9.747	(-3.525)	-0.473	(4.718)	2.934	(-0.537)	-0.008
Industrials Sector	159	2.255****		0.01****		2.323****		0.001		1.213****		0.134****	
		(55.779)		(4.119)		(36.332)		(0.294)		(4.297)		(4.014)	
Capital Goods	109	2.266****	2.31****	0.01****	0.028****	2.393****	2.111****	-0.002	0.004	1.155****	2.307****	0.142****	-0.012
		(47.543)	39.485	(3.804)	4.415	(32.226)	16.124	(-0.396)	0.75	(3.051)	4.527	(3.223)	-0.212
Commercial & Professional Services	32	2.406****	2.949****	-0.002	-0.067**	2.323****	2.472****	0.006	-0.012	1.306**	3.066***	0.136*	-0.1
		(19.006)	11.558	(-0.164)	-2.342	(16.423)	6.836	(0.621)	-0.29	(2.368)	2.659	(1.992)	-0.605
Transportation	18	2.079****	2.254****	0.008	0.026	1.783****	2.181****	0.018	-0.003	1.367	3.13	0.086	-0.122
		(16.78)	14.257	(0.771)	1.72	(8.565)	6.354	(1.565)	-0.283	(1.69)	1.006	(0.861)	-0.334
Consumer Discretionary Sector	144	2.152****		0.011****		2.036****		0.01***		1.393****		0.096****	
		(54.07)		(3.462)		(30.911)		(2.589)		(5.956)		(3.373)	
Automobiles & Components	18	1.707****	1.991****	0.025****	0.024	1.716****	2.365****	0.033	0.009	0.161	2.244*	0.223****	0.022
		(13.286)	8.325	(3.436)	1.755	(8.318)	11.525	(1.565)	0.095	(0.352)	2.014	(4.057)	0.125
Consumer Durables & Apparel	37	2.104****	2.51****	0.012*	0.028	2.131****	2.451****	-0.002	0.012	1.81****	-0.185	0.034	0.313****
		(32.866)	10.672	(2.03)	1.111	(18.801)	11.331	(-0.284)	1.5	(3.953)	-0.239	(0.652)	3.814
Consumer Services	39	2.402****	2.154****	0.007	0.038	2.233****	2.202****	0.017****	0.015	0.882***	1.614****	0.203****	0.116
		(30.943)	16.507	(1.243)	1.325	(22.986)	10.609	(2.9)	0.563	(2.413)	3.119	(4.332)	1.361
Retailing	50	2.093****	2.303****	0.004	0.029****	1.835****	2.002****	0.017**	0.008	1.463****	2.781**	0.076	-0.067
		(28.168)	26.191	(0.551)	3.175	(14.442)	7.189	(2.249)	0.6	(3.383)	2.156	(1.441)	-0.481
Consumer Staples Sector	65	2.43****		0.004		2.335****		0.012*		1.852****		0.087*	
		(37.935)		(0.921)		(27.582)		(1.91)		(5.036)		(1.682)	
Food & Staples Retailing	13	2.231****	1.397	0	0.155	2.059****	2.795****	0.021	-0.062	1.264	9.25***	0.132	-0.938**
		(20.227)	1.832	(0.044)	1.147	(10.051)	3.616	(0.972)	-0.707	(1.412)	3.081	(1.09)	-2.331
Food, Beverage & Tobacco	39	2.46****	2.103****	0.004	0.028**	2.452****	2.478****	0.004	-0.038	1.637****	2.452****	0.126*	-0.053
		(26.731)	45.073	(0.636)	2.296	(22.569)	10.866	(0.549)	-1.572	(3.395)	5.596	(1.801)	-0.745
Household & Personal Products	13	2.589****	2.365****	0.001	-0.006	2.151****	2.125****	0.033***	0.043	2.46***	2.199	0.018	0.012
		(20.065)	7.433	(0.084)	-0.301	(11.696)	8.053	(2.715)	0.577	(3.014)	1.706	(0.166)	0.057
Health Care Sector	123	2.574****		0.004*		2.673****		-0.003		0.877****		0.217****	
		(44.692)		(1.752)		(37.173)		(-0.66)		(5.133)		(10.495)	
Health Care Equipment & Services	76	2.517****	2.385****	0.01****	0.004	2.694****	2.411****	-0.003	0.06***	0.927****	0.75	0.215****	0.205****
		(34.989)	20.643	(3.174)	1.581	(30.243)	29.543	(-0.49)	2.641	(4.09)	1.44	(7.798)	3.448
Pharmaceuticals, Biotechnology & Life Science:	47	2.633****	2.473****	-0.002	0.001	2.621****	2.209****	-0.002	0.015*	0.778****	1.336*	0.223****	0.148
		(28.839)	17.69	(-0.672)	0.14	(20.706)	13.055	(-0.24)	1.767	(3.021)	1.781	(7.201)	1.543
Financials Sector	40	2.255****		0.017****		2.056****		0.014		1.682****		0.068	
		(29.305)		(3.163)		(13.774)		(1.555)		(4.039)		(1.387)	
Diversified Financials	40	2.255****	2.278****	0.017****	0.027****	2.056****	2.2****	0.014	0.006	1.682****	1.927****	0.068	0.051
		(29.305)	36.289	(3.163)	3.201	(13.774)	13.471	(1.555)	0.187	(4.039)	3.124	(1.387)	0.49
Information Technology Sector	138	2.34****		0.012****		2.557****		-0.002		1.195****		0.142****	
		(45.431)		(6.393)		(35.59)		(-0.498)		(5.057)		(5.751)	
Semiconductors & Semiconductor Equipment	37	2.269****	2.41****	0.012****	0.032****	2.23****	2.192****	0.009	0.001	1.615****	2.218*	0.075	-0.001
		(25.929)	24.676	(3.328)	4.209	(13.583)	6.099	(1.196)	0.046	(2.746)	1.934	(1.341)	-0.008
Software & Services	56	2.497****	2.673****	0.012****	0.006	2.866****	2.833****	-0.006	-0.004	1.059****	2.545****	0.183****	0.019
		(32.638)	14.115	(5.127)	0.317	(29.396)	15.91	(-1.54)	-0.625	(4.176)	4.308	(6.841)	0.314
Technology Hardware & Equipment	45	2.281****	2.524****	0.006	0.009	2.375****	2.459****	-0.001	0.026	0.274	0.669	0.238****	0.204****
		(23.13)	14.328	(1.351)	1.608	(20.444)	7.431	(-0.204)	0.975	(0.576)	0.935	(4.431)	2.957
Communication Services Sector	40	2.239****		0.011****		2.276****		0.009		1.475****		0.117****	
		(32.232)		(2.83)		(22.34)		(1.011)		(6.017)		(3.693)	
Media & Entertainment	30	2.264****	2.534****	0.009**	0.018*	2.26****	3.116****	0.011	-0.009	1.533****	3.288****	0.108****	-0.037
		(26.284)	11.282	(2.1)	1.717	(15.109)	12.117	(0.884)	-1.006	(4.941)	3.448	(2.781)	-0.432
Telecommunication Services	10	2.074****	2.126****	0.035***	0.111	2.285****	2.404****	0.004	-0.073	0.939*	-0.229	0.204***	0.448****
		(20.809)	10.64	(3.493)	1.073	(16.044)	12.658	(0.28)	-0.859	(2.183)	-0.354	(3.234)	3.921
Utilities Sector	52	2.28****		0.016****		2.676****		-0.05****		1.688****		0.117*	
		(41.616)		(3.268)		(35.614)		(-3.662)		(4.338)		(1.901)	
Utilities	52	2.28****	2.327****	0.016****	0.003	2.676****	2.527****	-0.05****	-0.041	1.688****	1.871****	0.117*	0.079
		(41.616)	21.794	(3.268)	0.47	(35.614)	20.275	(-3.662)	-1.354	(4.338)	3.232	(1.901)	0.868
Real Estate Sector	86	2.708****		0.004		2.799****		-0.016*		2.226****		0.068**	
		(74.728)		(1.287)		(56.653)		(-1.691)		(8.808)		(2.022)	
Real Estate	86	2.708****	2.331****	0.004	0.001	2.799****	2.279****	-0.016*	0.028	2.226****	2.614****	0.068**	-0.038
		(74.728)	25.902	(1.287)	0.197	(56.653)	32.24	(-1.691)	1.411	(8.808)	5.069	(2.022)	-0.524

**** p < 0.01
*** p < 0.025
** p < 0.05 ← *Our chosen significance threshold*
* p<0.1

6.2.2 Multiple Linear Regression (MLR)

We will in this section lay out the findings of our MLR which explores the relationship between the EBITDA-multiple and its fundamental underlying value drivers, jointly, with output as per **Table 6.5**. The table gives an overview of the relevant coefficients of our models, with several key statistical indicators (R^2, adjusted R^2, standard errors, F-statistics, t-statistics, and p-values. The output generated from the MLR models is subsequently later utilised to construct the predictive model. In a financial context, it is desirable to have a regression model which accurately predicts the dependent variable, based on the independent variables. In some cases, beta coefficients close to 1 and insignificant intercept coefficients may be desirable, however, the quality of the model should not be determined solely based on this output. The performance of the model should be evaluated based on the significance of the coefficients, together with the goodness-of-fit statistics and the practical significance in the context of the problem being addressed.

Starting with our MLR output on the market level, we see similarities to our SLR output, in that the statistically significant beta coefficients are growth and risk (with p < 0.01), with positive values of 0.009 and 0.069, respectively. The MLR assigns a similar beta coefficient for growth compared to the SLR output (0.009 vs. 0.01), but a slightly lower beta coefficient for risk as compared to the SLR output (0.069 vs 0.09). Furthermore, on a market level, the beta coefficient for profitability is also similar to the SLR model, which is close to zero, but negatively insignificant with p > 0.1. We have R^2 = 13.73%, adjusted R^2 = 13.45%, and a significant F-statistic at p<0.01. For our significant beta coefficients, the MLR tells us that a 1 unit (1 percentage point) increase in EBITDA CAGR will lead to a 0.9% increase in the EBITDA-multiple, whilst an increase of 1 unit in the WACC of any random will increase the EBITDA-multiple with 7.14%.[32]

Growth

Moving on to the **sector level**, we find that growth coefficients are significant at p < 0.01 for 4 groups (Industrials, Consumer Discretionary, Financials, and IT), whereas IT holds the highest R^2 at 33.71%, and with a significant F-statistic at p < 0.01. For all said GICS sector groups, the SLR predicted very similar significant (p < 0.01) beta coefficients for growth, but with the financial sector yielding the highest beta coefficient for growth at 0.018. On an **industry level**, we find the growth coefficient to

[32] $\Delta_{proxy} = one\ percentage\ unit \rightarrow \ \Delta_{EV/EBITDA} = \left(exp(\bar{y}_{proxy}) - 1\right), where\ \bar{y} = coefficient$

be significant at $p < 0.01$ for 4 industries (Capital Goods, Diversified finance, Software & Services, Semiconductor & Semiconductor Equipment), where Semiconductor & Semiconductor Equipment has the highest R^2 at 35.05%. We find the growth coefficient to be significant at $p < 0.025$ for one industry (Automobiles & Components), and at $p < 0.1$ for one industry (Consumer Durables & Apparel), where Automobiles Component has the highest R^2 at 69.70% with a 1% significant F-statistic. For the remaining 16 industries, we cannot reject the null hypothesis that the included coefficient of the covariate is equal to zero, i.e., the theoretically derived value driver, growth, does not explain any variation in the EV/EBITDA multiple in cohesion with profitability and risk.

The results from the MLR conducted on the **SARD grouping**, depicted in blue in **Table 6.5**, exhibit a result which is diverging quite substantially from the GICS industry segmentation. On a general level, the intercept coefficient for the SARD grouping is lower than the intercept coefficient on the GICS industry level, which may imply a lower degree of omitted variable bias. Furthermore, we find the growth coefficient to be significant at $p < 0.01$ for five groupings (Material, Capital Goods, Retailing, Diversified Financials, Semiconductor & Semiconductor Equipment). As compared to the GICS industry level, Semiconductor & Semiconductor Equipment has a higher R^2 at 39.11%. Furthermore, we find the growth coefficient to be significant at $p < 0.025$ for one industry (Telecommunication Services), and at a $p < 0.05$ for five industries (Commercial & Profession Services and, Automobiles & Components, Food, Beverages & Tobacco, Media & Entertainment, and Utilities). Telecommunication has the highest R^2 at 88.10%, with an F-statistic of 1.7773 and $p < 0.01$. It should be further noted that the beta coefficient for growth for the Telecommunication grouping is negative, as opposed to the output from the GICS industry grouping. For 11 out of the 22 SARD groupings, we reject the null hypothesis that growth does not explain any variation in the EV/EBITDA multiple. A result which is slightly better than the one yielded by the GICS industry grouping (18 out of 22). Furthermore, as opposed to the GICS industry grouping, the significant beta coefficients exhibited in the SARD grouping are further away from zero, implying a higher positive relationship with EV/EBITDA.

Profitability

For profitability, on a **sector level**, only the Utilities sector has a significant beta coefficient at $p < 0.01$. Furthermore, the Real Estate sector has a significant beta coefficient at $p < 0.025$, whilst Consumer Discretionary and Health care are significant at $p < 0.05$. For Utilities, MLR models a beta coefficient of -0.061 with an R^2 = 36.10% and a 1% significant F-statistic, whilst the SLR models an insignificant beta coefficient of -0.041 at $p > 0.1$ for the same industry. On an **industry level**, with the 22 industries analysed, only two exhibit a significant beta coefficient at $p < 0.01$ (Consumer Services and Utilities). The Consumer Services industry has a R^2 of 49.45% and the Utilities industry has a R^2 of 36.10%. It should be further noted that the Utilities industry exhibits a negative sign for this beta coefficient. Furthermore, three industries exhibit beta significance at $p < 0.025$ (Household & Personal Products, Health Care Equipment & Services, and Real Estate), where the Health Care sector has the highest R^2 of 49.70%. Furthermore, both Health Care Equipment sector and the Real Estate sector have a negative sign on their respective betas for profitability. Four industries exhibit a significance at $p < 0.1$ (Automobiles & Components, Retailing, Diversified Financials, Semiconductors & Semiconductor Equipment). Conclusively, in terms of profitability, 15 out of 22 industries have insignificant beta coefficients at $p > 0.1$. For these sectors, we cannot reject the null hypothesis that the included coefficient of the covariate is equal to zero. Furthermore, for the beta-significant industries, most beta coefficients are very close to zero, i.e., has a low explanatory power This result is conflicting with the theory that there is a strong positive relationship between EV/EBITDA and profitability.

For profitability, the MLR on **SARD groupings** yields worse statistical significance as compared to the GICS industries. Only one grouping exhibits a significant beta coefficient at $p < 0.01$ (Health Care Equipment & Services), and only one grouping exhibits a significant beta coefficient at $p < 0.025$ (Energy), with an R^2 of 22.27% and 26.32% respectively. Furthermore, only two groupings exhibit a significant beta coefficient at $p < 0.05$ (Pharmaceuticals, Biotechnology & Life Sciences, and Real Estate), with an R^2 of 11.03% and 2.63%, respectively. As for the MLR on the GICS industry grouping, there are several industries with a negative beta coefficient, but in the case of SARD, none of these is significant at $p < 0.1$. The Consumer Services industry which exhibited the highest R^2 from the GICS industry grouping, exhibit an R^2 of only 10.70% in the SARD groupings. Conclusively, 18 out of 22 industries have insignificant beta coefficients for profitability at $p < 0.05$. I.e., for these sectors, we cannot reject the null hypothesis that the included coefficient of the profitability covariate is equal to zero. As opposed to the GICS industry grouping, the significant beta coefficients from the SARD grouping are further away from zero and exhibit a positive sign, showing a higher positive relationship with EV/EBITDA.

Risk

For risk, 5 GICS sectors have significant beta coefficients with $p < 0.01$ (Energy, Industrials, Health Care, IT, Utilities and Real Estate). Besides that, Communication Services have a significant beta coefficient at $p < 0.05$. Out of the groups that are significant at $p < 0.01$, Health Care has the highest R^2 at 49.68%, with a 1% significant F-statistic. Furthermore, in the Health Care Sector, the MLR models a similar positive beta coefficient for risk as to the SLR, at 0.224 (vs 0.217). On an **industry level**, 8 industries exhibit beta values with $p < 0.01$[33]. Furthermore, the Capital Goods industry exhibits statistical significance at a $p < 0.025$, and Commercial & Professional Services exhibit statistical significance at a p-value of less than 0.1. Among the 10 industries with significant beta coefficients, only one has a negative sign (Energy). Furthermore, the size of the beta coefficients for risk across the different segmentations is significantly larger than the beta coefficient for growth and profitability. This finding is interesting because it conflicts with our theoretically derived hypothesis (**Hypothesis 3**).

For the beta coefficient for risk, the MLR on **SARD grouping** exhibited weaker explanatory power, with only two SARD exhibiting significant beta coefficients on $p < 0.01$ (Consumer Durables & Apparel, Telecommunication Services). Furthermore, two SARD industries exhibit significant beta coefficients on $p < 0.05$ (Health Care Equipment & Services, Technology Hardware & Equipment), and one SARD grouping exhibits a significant beta coefficient on $p < 0.1$ (Food Staples & Retailing). Out of the industries with a beta coefficient significant on $p < 0.05$, Telecommunication Services exhibit the highest R^2 of 88.10%, and Technology Hardware & Equipment exhibits the weakest R^2 of 18.54%. However, as seen in **Table 5.5**, Telecommunication Services exhibited moderate to high VIF values (5.72 for risk), implying moderate to high multicollinearity. Among the industries with a significant beta coefficient, only one has a negative sign (Food Staples & Retailing). Furthermore, as for the GICS industry groups, the beta coefficients for risk across the different industries are significantly larger than the beta coefficients for the other two value drivers. However, there are still 17 out of 22 industries exhibiting insignificant beta values for risk. I.e., for these groupings, we cannot reject the null hypothesis that the included coefficient of the risk covariate is equal to zero.

---

[33] Energy, Capital Goods, Consumer Services, Health Care Equipment & Services, Pharmaceuticals, Biotechnology & Life Sciences, Software & Services, Technology Hardware & Equipment, Utilities, and Real Estate

Goodness-of-fit statistics

In terms of Standard Error (SE), the MLR model exhibits an SE of 0.5080 with an F-statistic significant at $p < 0.01$ on a **market level**. On a **sector level,** 10 out of 11 sectors exhibit F-statistic at a 1% significance level, whereas the Consumer Staples sector exhibits an F-statistic which is significant at a 10% significance level. The model performance is more scattered on an **industry level** with 13 out of 22 industries exhibiting a significant F-statistic on a 1% significance level, 2 out of 22 industries exhibiting a significant F-statistic on a 10% significance level, and 7 industries with an F-statistic p-value larger than 10%.

On a **SARD grouping** level, 8 industries exhibited a lower SE compared to the GICS industry grouping, whereas 10 out of 22 industries exhibit a significant F-statistic on a 1% significance level. One industry exhibited a significant F-statistic on a 2.5% level (Pharmaceuticals, Biotechnology & Lifesciences), whilst two industries exhibited a significant F-statistic on a 5% level. As for the GICS industry groups, most of the SARD industry which exhibited a significant F-statistic on a 1% level had only one independent variable showing a significant beta coefficient. However, the Telecommunication Services industry has an R^2 of 88.10%, a significant F-statistic on a 1% level, and a significant beta coefficient for both growth and risk.

Upon examining the MLR regression output presented in **Table 6.5**, two GICS industry stands out. Specifically, the Automobiles & Components industry exhibits a strong performance with an R^2 value of 69.70%, which is only slightly lower in terms of adjusted R^2. The regression model for this sector demonstrates a strong explanatory power for the goodness-of-fit statistics, as indicated by the high adjusted R^2, low regression SE, and 1% significant F-statistic. Furthermore, **table 6.5** shows an intercept coefficient that is close to 1 and the beta coefficient for growth, and profitability is significant at $p<0.01$ and $p<0.1$ respectively. The other industry which stands out is Food, Beverage & Tobacco which exhibits the weakest performance in terms of R^2 at 8.59%, and an adjusted R^2 at 0.76%. This sector also exhibits a high regression standard error, together with an insignificant F-statistic at $p > 0.1$.

As with the **SARD groupings**, presented in **Table 6.6**, Telecommunication Services exhibit the strongest R^2 of 88.10%. The regression model for this SARD group demonstrates a strong goodness-of-fit as indicated by the R^2, low regression SE at 0.1762, and significant F-statistics on $p<0.01$. Moreover, in terms of the regression coefficients, growth is negative and significant at $p<0.025$, whilst risk is highly positive and significant at $p<0.01$. The SARD Real estate grouping exhibits the worst performance with an R^2 of 2.63%, relatively high SE, and insignificant beta coefficients for all three of the independent variables.

Furthermore, the intercept coefficient plays a crucial role in regression analysis as it signifies the value of the dependent variable when all the independent variables are equal to zero. In this case, the intercept coefficient holds considerable significance, as it is substantial and statistically significant at $p < 0.01$ at the aggregated market level, for all GICS sectors and most GICS industries. This implies that the dependent variable has a significant value even when all the independent variables are zero. On a market level, the intercept coefficient is 1.75, which is significant at $p < 0.01$. In addition, it is important to note that there are high disparities between sectors and industries. For instance, the Energy grouping exhibits the highest intercept coefficient of 5.578, which is significant at $p < 0.01$, while the Consumer Services grouping exhibits the lowest intercept coefficient of 0.813, which is significant on a 2.5% level. This suggests that the value of the dependent variable when all the independent variables are zero varies significantly across sectors and industries. There are also high disparities between the SARD groupings, but the intercepts have a narrower distribution with a lower mean. The Food Staples & Retailing grouping exhibits the highest intercept coefficient of 9.239, which is significant at $p < 0.05$. The Health Care Equipment & Service grouping exhibits the lowest positive intercept coefficient of 0.993, though only significant on a 10% level. Furthermore, the energy industry which exhibited the highest intercept in the GICS industry grouping, exhibited an intercept coefficient of 0.141.

Table 6.5 – Multiple linear regression output – market, sector, and industry

Multiple linear regression (MLR) output - market, sector and industry

	Subset A: Output statistics					Subset B: Coefficients			
	N	R^2	Adj. R^2	SE	F-statistic	Intercept	Growth	Profitability	Risk
Market	929	0.1373	0.1345	0.5080	49.0601****	1.75****	0.009****	-0.0002	0.069****
						(19.29)	(8.342)	(-0.088)	(5.722)
Energy Sector	26	0.4722	0.4002	0.5286	6.5602****	5.578****	0.003	-0.006	-0.561****
						(5.644)	(0.572)	(-0.3)	(-4.353)
Energy	26	0.4722	0.4002	0.5286	6.5602****	5.578****	0.003	-0.006	-0.561****
						(5.644)	(0.572)	(-0.3)	(-4.353)
Materials Sector	56	0.2300	0.1855	0.5085	5.1765****	2.406****	0.009	-0.016	-0.022
						(5.278)	(1.671)	(-1.526)	(-0.373)
Materials	56	0.2300	0.1855	0.5085	5.1765****	2.406****	0.009	-0.016	-0.022
						(5.278)	(1.671)	(-1.526)	(-0.373)
Industrials Sector	159	0.1693	0.1532	0.4272	10.5309****	1.363****	0.01****	0.005	0.099****
						(4.89)	(3.645)	(1.033)	(2.764)
Capital Goods	109	0.1862	0.1630	0.3980	8.0083****	1.29****	0.01****	0.005	0.106***
						(3.546)	(3.391)	(0.929)	(2.388)
Commercial & Professional Services	32	0.1193	0.0249	0.4916	1.2641	1.255**	0	-0.003	0.146*
						(2.067)	(0.024)	(-0.257)	(1.84)
Transportation	18	0.2368	0.0733	0.4850	1.448	2.351***	0.015	0.029	-0.086
						(2.516)	(1.378)	(1.752)	(-0.642)
Consumer Discretionary Sector	144	0.1524	0.1342	0.4569	8.389****	1.59****	0.01****	0.009**	0.055*
						(6.84)	(3.11)	(2.25)	(1.848)
Automobiles & Components	18	0.6970	0.6321	0.3381	10.7362****	1.049*	0.026***	0.039*	0.042
						(2.095)	(2.961)	(1.899)	(0.505)
Consumer Durables & Apparel	37	0.1164	0.0361	0.3984	1.4497	1.974****	0.011*	-0.004	0.022
						(4.06)	(1.851)	(-0.639)	(0.372)
Consumer Services	39	0.4945	0.4512	0.3256	11.4127****	0.813***	0.003	0.016****	0.184****
						(2.442)	(0.814)	(3.253)	(4.249)
Retailing	50	0.1171	0.0595	0.4741	2.0328	1.422****	0.001	0.015*	0.054
						(3.103)	(0.098)	(1.982)	(0.971)
Consumer Staples Sector	65	0.0985	0.0542	0.4108	2.2219*	1.926****	0.006	0.012*	0.051
						(5.279)	(1.175)	(1.813)	(0.952)
Food Staples & Retailing	13	0.1166	-0.1779	0.3839	0.3958	1.27	-0.003	0.007	0.125
						(0.961)	(-0.241)	(0.198)	(0.599)
Food, Beverage & Tobacco	39	0.0859	0.0076	0.4348	1.0968	1.652****	0.003	0.001	0.118
						(3.292)	(0.451)	(0.152)	(1.51)
Household & Personal Products	13	0.4500	0.2667	0.3241	2.455	1.928**	0.006	0.037***	0.017
						(2.665)	(0.763)	(2.706)	(0.184)
Health Care Sector	123	0.4968	0.4842	0.3648	39.1691****	0.919****	-0.001	-0.008**	0.224****
						(5.394)	(-0.316)	(-2.191)	(10.549)
Health Care Equipment & Services	76	0.4970	0.4760	0.3721	23.7144****	0.935****	0	-0.013***	0.229****
						(4.058)	(0.15)	(-2.475)	(7.339)
Pharmaceuticals, Biotechnology & Life Sciences	47	0.5382	0.5060	0.3518	16.7074****	0.817****	-0.001	-0.001	0.221****
						(2.859)	(-0.517)	(-0.121)	(6.991)
Financials Sector	40	0.2798	0.2198	0.4766	4.663****	2.19****	0.018****	0.017*	-0.02
						(5.461)	(3.197)	(1.807)	(-0.374)
Diversified Financials	40	0.2798	0.2198	0.4766	4.663****	2.19****	0.018****	0.017*	-0.02
						(5.461)	(3.197)	(1.807)	(-0.374)
Information Technology Sector	138	0.3371	0.3223	0.4627	22.7184****	1.355****	0.01****	0.001	0.108****
						(6.206)	(5.156)	(0.204)	(4.479)
Software & Services	56	0.5760	0.5516	0.3455	23.5485****	1.297****	0.006****	-0.003	0.148****
						(5.419)	(3.039)	(-1.061)	(5.505)
Technology Hardware & Equipment	45	0.3265	0.2772	0.4632	6.6239****	0.214	-0.002	-0.005	0.255****
						(0.423)	(-0.437)	(-0.873)	(4.168)
Semiconductors & Semiconductor Equipment	37	0.3505	0.2914	0.4578	5.9351****	1.844****	0.014****	0.015*	0.013
						(3.612)	(3.822)	(1.997)	(0.247)
Communication Services Sector	40	0.2986	0.2402	0.3290	5.1096****	1.577****	0.005	-0.003	0.1**
						(5.776)	(1.178)	(-0.304)	(2.311)
Telecommunication Services	10	0.6441	0.4662	0.2564	3.6196*	1.54*	0.023	-0.005	0.097
						(1.991)	(1.023)	(-0.505)	(0.719)
Media & Entertainment	30	0.2490	0.1624	0.3570	2.8736*	1.637****	0.005	0.001	0.086
						(4.886)	(1.011)	(0.049)	(1.623)
Utilities Sector	52	0.3610	0.3210	0.2056	9.0379****	1.597****	0	-0.061****	0.18****
						(4.764)	(0.003)	(-3.413)	(2.958)
Utilities	52	0.3610	0.3210	0.2056	9.0379****	1.597****	0	-0.061****	0.18****
						(4.764)	(0.003)	(-3.413)	(2.958)
Real Estate Sector	86	0.1426	0.1113	0.2688	4.5468****	1.993****	0.003	-0.025***	0.11****
						(7.802)	(1.057)	(-2.455)	(3.128)
Real Estate	86	0.1426	0.1113	0.2688	4.5468****	1.993****	0.003	-0.025***	0.11****
						(7.802)	(1.057)	(-2.455)	(3.128)

**** p < 0.01
*** p < 0.025
** p < 0.05 ← *Our chosen significance threshold*
* p<0.1

Table 6.6 – Multiple linear regression output – SARD groupings

Multiple linear regression (MLR) output - SARD groupings

		Subset A: Output statistics				**Subset B: Coefficients**			
	N	R^2	Adj. R^2	SE	F-statistic	Intercept	Growth	Profitability	Risk
Energy	26	0.2632	0.1627	0.6052	2.6192****	0.141	0.003	0.123***	0.252
						(0.064)	(0.642)	(2.533)	(0.939)
Materials	56	0.1507	0.1017	0.3996	7.3436****	2.126****	0.024****	-0.032	0.048
						(2.782)	(2.995)	(-1.012)	(0.453)
Capital Goods	109	0.1631	0.1392	0.5540	3.0749*	2.278****	0.028****	0.005	-0.01
						(4.826)	(4.436)	(1.066)	(-0.177)
Commercial & Professional Services	32	0.1566	0.0662	0.3615	6.6465****	3.223****	-0.066**	0.001	-0.041
						(2.908)	(-2.184)	(0.015)	(-0.254)
Transportation	18	0.1724	0.0049	0.5710	6.8233**	2.212	0.027	-0.005	0.024
						(0.648)	(1.621)	(-0.486)	(0.062)
Automobiles & Components	18	0.2279	0.0624	0.3755	1.7328****	1.349	0.0330**	0.11	0.044
						(1.116)	(2.024)	(1.097)	(0.262)
Consumer Durables & Apparel	37	0.3055	0.2424	0.3279	0.9724****	-0.204	0.016	0.002	0.293****
						(-0.251)	(0.738)	(0.246)	(3.12)
Consumer Services	39	0.1070	0.0283	0.2867	12.3814	1.41***	0.045	0.035	0.074
						(2.48)	(1.322)	(1.194)	(0.81)
Retailing	50	0.2141	0.1629	0.5339	1.3771	3.333****	0.032****	0.017	-0.147
						(2.815)	(3.395)	(1.354)	(-1.11)
Food Staples & Retailing	13	0.3963	0.1951	0.4388	4.8395****	9.239**	0.051	-0.06	-0.9040**
						(2.407)	(0.377)	(-0.748)	(-2.026)
Food, Beverage & Tobacco	39	0.1635	0.0917	0.2850	1.3586	2.535****	0.025**	-0.025	-0.033
						(5.952)	(2.033)	(-0.97)	(-0.466)
Household & Personal Products	13	0.0357	0.0286	0.2730	4.1773****	2.402	-0.004	0.047	-0.038
						(1.662)	(-0.166)	(0.502)	(-0.149)
Health Care Equipment & Services	76	0.2227	0.1903	0.5535	2.9139	0.993*	0.004	0.064****	0.146**
						(1.927)	(1.525)	(2.746)	(2.287)
Pharmaceuticals, Biotechnology & Life Sciences	47	0.1103	0.0482	0.3907	1.9696***	1.107	0.002	0.0140**	0.139
						(1.452)	(0.215)	(1.698)	(1.464)
Diversified Financials	40	0.2349	0.1729	0.3927	2.2795**	1.63****	0.029****	0.004	0.107
						(2.84)	(3.321)	(0.145)	(1.125)
Software & Services	56	0.0121	0.0448	0.4900	0.111	2.593****	0.006	-0.005	0.02
						(4.293)	(0.306)	(-0.697)	(0.318)
Technology Hardware & Equipment	45	0.1854	0.1258	0.5614	32.5652*	0.716	0.005	0.007	0.178**
						(0.981)	(0.907)	(0.239)	(2.266)
Semiconductors & Semiconductor Equipment	37	0.3911	0.3357	0.5104	6.8775	1.521	0.038****	0.016	0.045
						(1.614)	(4.603)	(1.591)	(0.479)
Telecommunication Services	10	0.8810	0.8215	0.1762	1.7773****	-1.832**	-0.242***	0.008	0.793****
						(-2.657)	(-3.354)	(0.215)	(5.98)
Media & Entertainment	30	0.1398	0.0406	0.5545	5.781****	3.437****	0.0181**	-0.009	-0.062
						(3.52)	(1.699)	(-0.942)	(-0.745)
Utilities	52	0.0495	0.0099	0.3548	3.7863	2.016****	0.0011**	-0.039	0.075
						(3.266)	(0.225)	(-1.219)	(0.826)
Real Estate	86	0.0263	0.0094	0.4971	18.7795****	2.358****	0.001	0.0290**	-0.017
						(4.291)	(0.484)	(1.363)	(-0.217)

**** p < 0.01
*** p < 0.025
** p < 0.05 ← *Our chosen significance threshold*
* p<0.1

6.3 White's test – Testing for Heteroscedasticity

In **section 5.5.2** we test the MLR assumptions of our models, where we under the 5th assumption test for heteroskedasticity in our dataset, thereby testing **hypothesis 5**. Per output in **Appendix 5,** we look at White's test statistics and relevant significance levels (for the EBITDA-multiple and value driver proxies, jointly for market, sector-, industry- and SARD groupings. With significant values in cases of heterogeneity, we compare the (%)-number of significant values across the significant segmentations. For the market we have 1/1 = 100% significance, for the sector groupings we have

3/11 = ~27% significant values, for the industry groupings we have 4/22 = ~18% significant values – and lastly, for SARD groupings, we have 2/11 = ~9% significant values. Hence White's test seems to indicate more homoskedastic datasets across sectors than on an aggregate market level, more homoskedastic datasets across industries than on a sector level, and finally, SARD groupings exhibit the lowest degree of heteroscedasticity.

However, when doing a two-sample t-test with a confidence interval of 95%, the only significant difference in means of White's test output is seen between SARD groupings and industry groups. Further to this, we find differences in significant levels of the different WT outputs across all segmentations. On a market level, we see a significance of $p < 0.05$. However, for all other segmentations significance is sporadic, with 3 out of 11 sectors having a WT with $p < 0.05$, 4 out of 22 industries, and 2 out of 22 SARD groupings. However, we still note the tendency of more homoskedasticity when moving from market to sector, to industry, to SARD. All in all, error terms for all segmentations except for market are deemed relatively homoskedastic, and more homoskedastic than on an aggregate market level.

6.4 MLR Prediction Accuracy

In **sections 6.1-6.4**, we presented summary statistics and the output of the linear regression conducted. In the following sections, we will evaluate the performance of the regression models by examining how accurately they predict EV/EBITDA multiples compared to the observed values and relative to peer group averages. We will use various statistical tests to measure the accuracy of the models and quantify their performance. The output from the accuracy tests is analysed on a relative basis to the different market levels and a relative basis to the peer group average. The observed multiples for the different peer groups are simply found by taking the average of the various peer groups' median multiple on the different segmentation levels: Market, GICS sectors, GICS industries, and SARD groupings. See equation 5.28

Equation 5.28

$$PGA_i = H\ Mean_i\Big(Median\big(EVE_{2020,x}, EVE_{2021,x}, EVE_{2022,x}\big) + \cdots + Median\big(EVE_{2020,y}, EVE_{2021,y}, EVE_{2022,y}\big)\Big)$$

The methodology of utilising the harmonic mean of peer group multiples is commonly employed in academic and professional contexts. It is deemed favourable due to its simplicity and ability to offer an informative summary of the market prices of companies within a given market, sector, or industry. As outlined in **sections 3.3 and 5.3.3**, there are two schools of thought when forming different peer groups. The first approach to peer group selection, which is the most used, is based on industry affiliation (Berk & DeMarzo, 2019). The second approach argues for peer group selection based on firms that share similar dynamics in their underlying value drivers (Pearl & Rosenbaum, 2009). The objective of this section is two-folded. Firstly, we wish to test the prediction accuracy of our regression-based model compared to a simple average model. Secondly, we wish to test the relative accuracy of the model on an industry level when applying the two most common peer group segmentation approaches: GICS segmentation, and segmentations based on similar dynamics in underlying value drivers. For the second approach, the SARD segmentation model is applied. Furthermore, to ensure comparability when testing the second objective, the SARD industries will consist of the same number of constituents as the comparable GICS industry.

Furthermore, as outlined in **section 5.5.3**, the predicted EV/EBITDA multiple for the different segmentation levels is estimated with equation 5.24 below, for all segmentation levels.

Equation 5.24

$$\widehat{EVE}_i = \hat{\alpha}_i + \hat{\beta}_1(Growth_i) + \hat{\beta}_2(Proftiability_i) + \hat{\beta}_3(Risk_i)$$

The intercept coefficient, $\hat{\alpha}_i$, and the slope coefficients, $\hat{\beta}_{1-3}$, for the independent variable's growth, profitability and risk for each segmentation level are taken from multivariate regression conducted in **section 6.2.2**. This means that we use the estimated regression model derived in **section 5.5.3**, together with the observed 2022 data as inputs, and plugged into the loss functions to evaluate the accuracy across the different prediction models. Finally, the exponential function, $e^{LN(\frac{EV}{EBITDA})}$, has been used to convert the natural logarithm (LN) of EV/EBITDA back to the original scale of EV/EBITDA. This is necessary to compare the predicted values with the observed values, which are also in the original scale.

6.4.1 Error Terms

As stated in **section 5.5.3**, loss functions are commonly used to assess the accuracy of prediction models. Three types of loss functions are frequently employed, Mean Squared Error (MSE), Mean

Absolute Deviation (MAD), and Mean Absolute Percentage Error (MAPE), each with distinct strengths and limitations.

MSE is a quadratic loss function, thus holding the distinct advantage of penalising outliers. The "error"-term is the distance between the observed and the predicted multiple. MSE calculates the squared distance between the observed multiple, y_i, and the predicted multiple, $\hat{y}_i$. This error function deals with outliers by assigning larger weights to large errors and lower weights to errors close to zero. The Mean Absolute Deviation (MAD) measures the average absolute distance between each observed multiple, y_i, and its corresponding predicted estimate, $\hat{y}_i$. MAD is expressed in the same units as the underlying data, i.e., multiples. Unlike the Mean Squared Error (MSE), MAD calculates the errors in absolute values and assigns equal weight to all observations. While the output of MAD is more easily interpretable, it does not penalise outliers, which is a significant drawback. Finally, MAPE is a relative error function that expresses the absolute error in percentage terms by dividing the difference between the observed multiple, y_i, and the predicted estimate, $\hat{y}_i$, by y_i. One major advantage of MAPE is its ability to assess prediction accuracy across different samples and subsamples with varying distributions, which is highly relevant for this study. Thus, we intend to compare the prediction accuracy of the models across different segmentation levels, including the market, GICS sector, GICS industry, and SARD industry. However, similar to MAD, MAPE does not penalise outliers.

Each loss functions possess different strengths and weaknesses, hence, to provide a more complete evaluation all three loss functions are used in combination. The output from the accuracy tests is analysed on a relative basis for the different market levels and on a relative basis to the peer group average, where a prediction error closer to zero signifies a higher accuracy.

6.4.2 Regression Model Accuracy: GICS Segmentations

Table 6.7, Subset A and Subset B, display the prediction errors yielded by the multivariate regression model and the peer group average. The blue-shaded numbers have the lowest prediction error relative to the other model. On an aggregated level, the regression model yields higher prediction error compared to the peer group average model for all sample groups in terms of MSE. Analysing MAD on a market level, our regression model yields a higher prediction error compared to the peer group average. For MAPE, the regression models outperform the peer group average model.

Analysing MAD on a sector level, the regression model yields a lower prediction error for 1 out of 11 sectors (Financial sector), corresponding to 9.1% of the regressions conducted on a sector level. For MAPE, the regression model is superior in 2 out of 11 sectors (Consumer Discretionary, and Financials), corresponding to 18.2% of the subsamples. Analysing MAD on an industry level, the regression model outperforms in 3 out of 22 industries (Transportation, Consumer Durables, and Diversified Financials). This corresponds to 13.64% of the regressions conducted on an industry level, which is slightly higher than the 9.1% on a sector level. In terms of MAPE on an industry level, the regression model outperforms the peer group average model in 6 out of 22 industries. This corresponds to 27.3% of the instances, which is also higher than the 18.2% on a sector basis.

Table 6.7 – MLR model accuracy – Market and GICS segmentations versus peer group averages

MLR model accuracy - Market and GICS segmentations versus peer group averages

		Subset A: Regression model			Subset B: Peer Group Average		
	N	MSE	MAD	MAPE	MSE	MAD	MAPE
Market	929	91.1845	7.0255	0.5006	46.6737	4.9679	0.5845
Energy Sector	26	n.a	249.1844	93.6973	6.4742	1.9021	0.8548
Energy	26	n.a	249.1844	93.6973	6.4742	1.9021	0.8548
Materials Sector	56	28.4945	4.6127	0.7820	25.5488	4.0805	0.6118
Materials	56	28.4945	4.6127	0.7820	25.5488	4.0805	0.6118
Industrials Sector	159	90.0224	7.6134	0.5898	32.8133	3.8133	0.4339
Capital Goods	109	89.4807	8.0377	0.6287	25.3118	3.6551	0.4035
Commercial & Professional Services	32	144.3482	8.9619	0.6375	64.0732	4.8293	0.4526
Transportation	18	15.8103	3.1866	0.6330	12.7368	3.2350	0.5314
Consumer Discretionary Sector	144	53.2323	5.1652	0.4240	28.4736	3.9280	0.4904
Automobiles & Components	18	56.9049	5.5311	0.5535	26.8073	3.7847	0.5674
Consumer Durables & Apparel	37	20.3424	2.9904	0.2884	17.3057	3.0573	0.3636
Consumer Services	39	150.6493	10.4470	0.3782	41.5761	4.4793	0.3982
Retailing	50	41.3529	4.9135	0.4518	18.4610	3.5079	0.4975
Consumer Staples Sector	65	61.5663	6.0831	0.4057	26.9297	4.1399	0.3993
Food Staples & Retailing	13	53.2604	6.3192	0.5935	13.3389	3.0798	0.3237
Food, Beverage & Tobacco	39	91.9224	8.0875	0.5540	28.2505	4.0757	0.4024
Household & Personal Products	13	79.4974	7.3599	0.4481	25.4360	4.3622	0.3518
Health Care Sector	123	255.5120	13.3401	0.7918	77.6693	6.4619	0.5180
Health Care Equipment & Services	76	278.8745	13.7903	0.7941	88.8275	6.8401	0.5222
Pharmaceuticals, Biotechnology & Life Sciences	47	222.0906	12.7901	0.8037	58.6019	5.8109	0.5063
Financials Sector	40	47.0401	4.8977	0.4317	42.9464	5.3550	0.5817
Diversified Financials	40	47.0401	4.8977	0.4317	42.9464	5.3550	0.5817
Information Technology Sector	138	183.2920	10.7367	0.6441	68.8113	6.5712	0.6107
Software & Services	56	281.8417	14.1758	0.7315	81.1410	7.2680	0.5359
Technology Hardware & Equipment	45	167.6151	10.9386	0.8615	47.9934	5.1729	0.5641
Semiconductors & Semiconductor Equipment	37	87.6249	6.7909	0.4511	48.2941	5.5097	0.5630
Communication Services Sector	40	62.7648	6.4972	0.5081	20.6002	3.6483	0.3490
Telecommunication Services	10	46.5626	5.9055	0.5056	11.7557	2.9743	0.3131
Media & Entertainment	30	64.8632	6.4517	0.4852	23.2869	3.8843	0.3619
Utilities Sector	52	52.6058	6.6714	0.5457	8.1209	2.0717	0.1932
Utilities	52	52.6058	6.6714	0.5457	8.1209	2.0717	0.1932
Real Estate Sector	86	98.0549	8.6277	0.4997	23.6473	3.5473	0.2408
Real Estate	86	98.0549	8.6277	0.4997	23.6473	3.5473	0.2408

Figures marked in blue highlights the model with the lowest error value
MSE > 500 = n.a.; MAD > 500 = n.a.; MAPE > 100 = n.a.

The peer group average model is superior, yielding a lower relative error term on an aggregated level for MSE, MAD, and MAPE across all segmentation levels: market, sector, and industry. Furthermore, a two-sample mean t-test[34] is conducted to determine whether the mean error terms of the two models across all subgroups are significantly different from each other. I.e., if the observed difference between the error terms is likely to have occurred by chance, or if it represents a statistically significant mean difference. If the t-test is significant on $p<0.05$, one can reject the null hypothesis on a 5% significance level that there is no significant difference between the mean errors. **Appendix 7** displays the p-values of the t-test for all segmentation levels. On a 5% significance level, the performance of the peer group model is significantly different from the regression model in most of the subsamples.

6.4.3 Regression Model Accuracy: SARD Segmentations

Table 6.8 presents the prediction errors resulting from the multivariate regression performed on the SARD grouping level and the corresponding peer group average. Consistent with the previous comparison, the SARD regression model exhibits higher MSE compared to the peer group average for all industries. The same pattern is observed for MAD, where the SARD regression model performs better than the peer group average in only 3 out of 22 industries, corresponding to 13.64% of the instances. However, for MAPE, the SARD regression model outperforms the peer group average in 8 out of 22 industries, representing 36.4% of instances. This percentage is higher than the 27.3% yielded by the GICS regression-peer group average comparison.

34 The two sample mean t-test is calculated using the following formula, $t = \frac{(\bar{x}_1 - \bar{x}_2) - (\bar{\mu}_1 - \bar{\mu}_2)}{s_p * \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$

Table 6.8 – MLR model accuracy – SARD groupings versus peer group averages

MLR model accuracy - SARD groupings versus peer group averages							
		Subset A: Regression model			Subset B: Peer Group Average		
	N	MSE	MAD	MAPE	MSE	MAD	MAPE
Energy	26	184.4697	11.7513	0.8616	46.3994	5.7871	0.7967
Materials	56	16.6362	3.1500	0.3518	14.9771	3.1244	0.4033
Capital Goods	109	34.9426	4.5089	0.5954	34.2970	4.6065	0.6588
Commercial & Professional Services	32	200.0092	13.4239	1.4823	19.9180	3.6258	0.3462
Transportation	18	19.9169	3.5274	0.5793	19.9643	3.5292	0.5902
Automobiles & Components	18	78.8516	7.7141	0.6134	19.3781	3.1510	0.3294
Consumer Durables & Apparel	37	303.3948	16.1431	0.9432	42.8077	4.8480	0.3247
Consumer Services	39	52.5075	6.4450	0.5756	10.9800	2.3178	0.2341
Retailing	50	337.6390	17.5209	2.7643	31.0218	4.5753	0.6438
Food Staples & Retailing	13	n.a.	n.a.	n.a.	26.8883	3.7068	0.4355
Food, Beverage & Tobacco	39	20.4698	3.8256	0.5664	5.9432	2.1688	0.2795
Household & Personal Products	13	6.2827	2.0939	0.2431	5.1254	1.7156	0.1854
Health Care Equipment & Services	76	226.8861	12.2094	0.6954	78.2597	6.7638	0.7010
Pharmaceuticals, Biotechnology & Life Sciences	47	128.3408	9.9854	0.7253	28.6468	4.1851	0.3682
Diversified Financials	40	38.2687	5.1813	0.4646	14.0078	3.0773	0.4119
Software & Services	56	76.8273	6.4528	0.3886	63.2409	6.0783	0.4626
Technology Hardware & Equipment	45	372.5902	16.5571	0.8402	98.5053	8.0573	0.6874
Semiconductors & Semiconductor Equipment	37	79.8028	6.6354	0.4937	40.4274	5.2114	0.7166
Telecommunication Services	10	130.3080	10.5205	0.9818	19.6328	3.7284	0.3822
Media & Entertainment	30	205.4677	11.8589	1.0850	95.7076	8.4310	0.6202
Utilities	52	29.7131	4.3329	0.2903	15.0901	2.9607	0.3103
Real Estate	86	36.3472	4.1615	0.4340	35.0213	4.1757	0.4823

Figures marked in blue highlights the model with the lowest error value
MSE > 500 = n.a.; MAD > 500 = n.a.; MAPE > 100 = n.a.

As for the GICS regression-peer model, the SARD regression-peer group model yields quite similar results. Even though the peer group model is on average more accurately measured on MSE, MAD and MAPE, the regression conducted on the SARD grouping increases the relative performance to the peer group average, as compared to the GICS regression-peer group average test. Furthermore, when viewing **Appendix 7.2** which displays the p-values for a t-test for the differences in the mean errors, it is evident that the null hypothesis of the mean difference being equal to zero can be rejected in far fewer instances. I.e., the SARD regression versus SARD peer group averages yields better results than GICS regression versus GICS peer group averages, however, few of the estimates are significant at a 5% level.

6.4.4 Regression Model Accuracy: GICS Industries versus SARD Groupings

For the last relative model performance accuracy, **table 6.9** depicts the predicted error resulting from the MLR performed on the GICS industry level against the MLR performed on the SARD grouping level. When analysing MSE, the SARD model has a lower error in 13 out of 22 industries, corresponding to 59.1% of the instances. Furthermore, the SARD model had a relatively lower prediction error compared to the GICS industry model in 12 out of 22 industries on both MAD and MAPE, corresponding to 54.5% of the instance for both tests.

Table 6.9 – MLR model accuracy – GICS industry groups versus SARD groupings

MLR model accuracy - GICS industry groups versus SARD groupings

		Subset A: MLR model GICS			Subset B: MLR model SARD		
	N	MSE	MAD	MAPE	MSE	MAD	MAPE
Energy	26	n.a.	n.a.	n.a.	184.4697	11.7513	0.8616
Materials	56	28.4945	4.6127	0.7820	16.6362	3.1500	0.3518
Capital Goods	109	89.4807	8.0377	0.6287	34.9426	4.5089	0.5954
Commercial & Professional Services	32	144.3482	8.9619	0.6375	200.0092	13.4239	1.4823
Transportation	18	15.8103	3.1866	0.6330	19.9169	3.5274	0.5793
Automobiles & Components	18	56.9049	5.5311	0.5535	78.8516	7.7141	0.6134
Consumer Durables & Apparel	37	20.3424	2.9904	0.2884	303.3948	16.1431	0.9432
Consumer Services	39	150.6493	10.4470	0.7829	52.5075	6.4450	0.5756
Retailing	50	41.3529	4.9135	0.4518	337.6390	17.5209	2.7643
Food Staples & Retailing	13	53.2604	6.3192	0.5935	n.a.	n.a.	n.a.
Food, Beverage & Tobacco	39	91.9224	8.0875	0.5540	20.4698	3.8256	0.5664
Household & Personal Products	13	79.4974	7.3599	0.4481	6.2827	2.0939	0.2431
Health Care Equipment & Services	76	278.8745	13.7903	0.7941	226.8861	12.2094	0.7404
Pharmaceuticals, Biotechnology & Life Sciences	47	222.0906	12.7901	0.8037	128.3408	9.9854	0.7253
Diversified Financials	40	47.0401	4.8977	0.4317	38.2687	5.1813	0.4646
Software & Services	56	281.8417	14.1758	0.7315	76.8273	6.4528	0.3886
Technology Hardware & Equipment	45	167.6151	10.9386	0.8615	372.5902	16.5571	0.8402
Semiconductors & Semiconductor Equipment	37	87.6249	6.7909	0.4511	79.8028	6.6354	0.4937
Telecommunication Services	10	46.5626	5.9055	0.5056	130.3080	10.5205	0.9818
Media & Entertainment	30	64.8632	6.4517	0.4852	205.4677	11.8589	1.0850
Utilities	52	52.6058	6.6714	0.5457	29.7131	4.3329	0.2903
Real Estate	86	98.0549	8.6277	0.4997	36.3472	4.1615	0.4340

Figures marked in blue highlights the model with the lowest error value
MSE > 500 = n.a.; MAD > 500 = n.a.; MAPE > 100 = n.a.

Looking at **Appendix 7.3** displaying the p-values from the two-sample t-test, the null hypothesis of the mean errors being equal to zero is rejected in even fewer instances as compared to the models presented in **sections 6.4.2-6.4.3**.

Based on the conducted analysis, we see tendencies that the accuracy of valuation using a regression methodology increases with greater homogeneity in the sample. Specifically, the employment of a regression approach on SARD grouping appears to yield superior outcomes in terms of accuracy in valuation. However, as the accuracy of valuation using the regression model increases, the p-values also increase, indicating that it does not produce error estimates that are significantly lower than peer group averages at a significance level of $p<0.05$.

6.5 Regression Diagnostics

Our MLR models as per **Tables 6.5** and **6.6** show significant positive relationships between the EBITDA-multiple and risk for the market, many sectors, and industries, and several SARD groupings. Furthermore, correlation matrices in **appendices 3** and **4** generally show positive correlations between LN(EV/EBITDA) and WACC, with ~40% of the values showing positive correlations with $p < 0.05$.

As this goes against our initial hypothesis, we wish to test for potential sporadic curvature of $WACC_i$ through a **shifted power transformation** of its variables across our dataset, and potentially transform the original dataset and rerun our models. Our WACC values range from a 10th percentile of 6.3% to a 90th percentile of 10.2%, and a curvilinear relationship would entail differing concavity/convexity along with increasing values of WACC, followingly changing the relationship between our explanatory variable and response variable. This would for example be if the model predicts that an increase in WACC from 8-10% lead to a higher increase in the EV/EBITDA multiple than an increase in WACC from 10-12%, whereby the 8-10% range would be more convex.

We test this by creating a "shifted" variable ($WACC_{shifted}$) for all values of $WACC$ in our dataset, which is calculated as follows:

Equation 6.1

$$WACC_{shifted,i} = (WACC_i + 1)^2 - 1$$

To study the linearity of $WACC_i$, we need to run our MLR with both variables present. For the sake of parsimony, and given that we have few significant beta coefficients for profitability, we remove this independent variable and replace it with $WACC_{shifted,i}$, whereby we get the following formula:

Equation 6.2

$$\left(\frac{EV}{EBITDA}\right)_i = \alpha + \beta_1(EBITDA\ CAGR)_i + \beta_2(WACC)_i + \beta_3\left(WACC_{shifted}\right)_i + \varepsilon_i$$

We run our regression, and can followingly run the below formula for all values of *i*:

Equation 6.3

$$\beta_2(WACC)_i * WACC_i + \beta_3\left(WACC_{shifted}\right)_i * WACC_{shifted,i}$$

This formula will let us study potential curvilinearity by seeing what values of i have sign change compared to preceding figures of i, hence illustrating that the line is changing direction (non-linearity). The output of the MLR can be seen in **Appendix 8**, where we see only 2 sectors (Materials and IT), 3 industries (Materials, Capital Goods, and Software & Services) and 1 SARD grouping (Software & Services) significant at $p < 0.05$. When both $\beta_2(WACC)_i + \beta_3\left(WACC_{shifted}\right)_i$ for a given grouping are significant, it implies the presence of non-linearity. However, it is evident from the output that this only holds for a smaller number of the subsamples of our total dataset – and we do not conduct further testing for curvilinearity across our dataset.

6.6 Robustness Test

As outlined in **section 3.1.1**, regression models attempt to model and explain differences over time which may be skewed due to changes in market conditions (Damodaran A., 2012). Damodaran (2012) further illustrated this point, by depicting varying R-squared statistics for valuation multiples on a dataset from 1987 to 1991. Harbula (2009) proved a similar point, by showing how valuation errors peaked around time of economic uncertainty. Following this notion, we tested the robustness of our model across time by re-conducting the multivariate analysis on a market level for 2018-2021. See **Table 6.10** for a 5-year regression output.

Table 6.10 – Robustness Test – 5-year trailing analysis

Robustness Test - 5 year trailing analysis

Segmentation	Year	Subset A: Output statistics					Subset B: Coefficients			
		N	R^2	Adj. R^2	SE	F-stat	Intercept	Growth	Profitability	Risk
Market	2022	929	0.1373	0.1345	0.508	49.0601****	1.75****	0.009****	-0.0002	0.069****
							(19.29)	(8.342)	(-0.088)	(5.722)
Market	2021	943	0.0043	0.0014	0.6125	1.4651	2.25****	0.001	-0.0002	0.006
							29.211	(0.739)	(1.702)	(0.743)
Market	2020	1021	0.1309	0.1285	0.5526	55.1293****	2.062****	-0.008****	0.003*	0.049****
							33.715	(-11.898)	(1.92)	(6.065)
Market	2019	1013	0.1193	0.1168	0.4486	48.0338****	2.177****	-0.007****	-0.001	0.037****
							37.81	(-11.376)	(-0.491)	(4.542)
Market	2018	895	0.0325	0.0291	0.4034	9.562****	2.403****	-0.005****	-0.001	-0.002
							40.526	(-5.214)	(-0.857)	(-0.304)

**** p < 0.01
*** p < 0.025
** p < 0.05 ← Our chosen significance threshold
* p<0.1

As depicted in **Table 6.10** the coefficients and the goodness-of-fit measures varied quite substantially across the five years. For 2022, 2020 and 2019 the model depicts an R-squared of 13.73%, 13.09%, and 11.93% respectively, which all have a significant F-statistic on a 1% significance level. Whereas, for 2021 and 2018, the model depicts an R-squared of 0.43% and 3.25% respectively. In terms of the growth coefficient, it is somewhat surprising that 2020 through 2018 depicts negative beta coefficients for growth which are all significant on a 1% significance level. In other words, growth had a negative impact on the EV/EBITDA multiple at the market level. For profitability, the coefficient showed a negative sign for all years except for 2020, though not significant on $p < 0.1$. The coefficient for risk had a positive and significant relationship with the EV/EBITDA multiple at a 1% level for 2022, 2020 and 2019. To summarise, the study has identified intertemporal differences in the relationship between EV/EBITDA and its fundamental value drivers.

6.7 Hypotheses Testing

Research question 1

In **sections 6.1-6.2** we sought to answer **research question 1** through testing **hypotheses 1-3**, that growth and profitability would have a significant positive relationship with the EBITDA-multiple, and that risk would have a significant negative relationship, all with a t-statistic ≠ 0. **Table 6.11** summarises these findings, where the hypotheses are confirmed or rejected based on one-sided t-statistics with a significance level of 5%. I.e., if the beta coefficient is significant on $p < 0.05$, we reject the null hypothesis and confirm our alternative hypothesis. **Hypothesis 1** is confirmed in 18 out of the 33 sub-samples on an aggregated market, sector, and industry level. When isolating GICS industries, **hypothesis 1** is confirmed in 10 out of 22 instances, whereas for SARD industries **hypothesis 1** is confirmed in 6 out 22 instances. **Hypothesis 2** is confirmed in only 4 out of 33 instances on an aggregated level, with 1 instance on a sector level and 3 instances on the industry level. For the SARD industries, **hypothesis 2** is confirmed in 2 cases, both being on an industry level. **Hypothesis 3** is confirmed in 2 out of 33 sub-samples on an aggregated level, whereas for SARD industries it is confirmed in 1 sub-sample.

Furthermore, **Section 6.2** also tested **hypothesis 4**, i.e., when accounting for differences amongst independent variables, relative performance in the three value drivers jointly contributes with significant explanatory power in EV/EBITDA, with F-statistics ≠ 0. We confirm or reject the hypothesis based on the F-statistics with a one-sided significance level of 5%. Here one should note that a significant F-statistics indicates that at least one of the independent variables has a significant relationship with the dependent variable and that in cohesion, the independent variables have a significant relationship with the dependent variables. Following this notion, we will not include subsamples with significant F-statistics where it is only the intercept coefficient which is statistically significant.

Table 6.11 depicts that on an aggregated market-, sector-, and industry level, 22 out of 33 sub-samples confirm **hypothesis 4**. Isolating GICS industries, **hypothesis 4** is confirmed in 12 out of 22 industries[35]. The same holds for SARD industries, where the hypothesis is also confirmed in 12 out of 22 groupings[36]. In other words, in terms of F-statistics, the model is deemed statistically useful in explaining the variation in EV/EBITDA for 22 sub-samples on an aggregated market level and for 12 sub-samples for the SARD grouping.

However, in terms of R-squared, the proportional variance in EV/EBITDA explained by the linear regression model is deemed quite poor. Furthermore, the market and many industries have a high intercept coefficient, which indicates the model has omitted variables. When analysing the R-squared on the different segmentation levels, it is evident that the R-squared increases as we go from the market to the sector, to the industry. Whilst a potential explanation for the gradually increasing R-squared could be due to the data sample becoming more homogenous, it is evident in **Table 6.6** that on a general level, the SARD industries (orange) have a lower R-squared than the GICS industries. This will be further discussed in **section 7**. Even if the model does not possess high explanatory power in the EV/EBITDA movements, it is still deemed that regressing theoretically derived value drivers against EV/EBITDA does contribute some significant explanatory power.

Table 6.11 – Hypothesis testing (H1-H4)

[35] Excluding Materials sector and Materials industry, which have a significant F-statistic, but only with a significant intercept coeffients, and insignificant beta coefficients

[36] Excluding Household & Personal Products, which have a significant F-statistic, but only with a significant intercept coeffients, and insignificant beta coefficients

Hypothesis testing (1-4)								
	Hypothesis 1		Hypothesis 2		Hypothesis 3		Hypothesis 4	
Market	*Confirmed*		Rejected		Rejected		*Confimed*	
Energy sector	Rejected		Rejected		*Confirmed*		*Confirmed*	
Energy	Rejected	Rejected	Rejected	*Confirmed*	*Confirmed*	Rejected	*Confirmed*	*Confirmed*
Materials sector	*Confirmed*		Rejected		Rejected		Rejected	
Materials	*Confirmed*	*Confirmed*	Rejected	Rejected	Rejected	Rejected	Rejected	*Confirmed*
Industrials sector	*Confirmed*		Rejected		Rejected		*Confirmed*	
Capital Goods	*Confirmed*	*Confirmed*	Rejected	Rejected	Rejected	Rejected	*Confirmed*	Rejected
Commercial & Professional Services	Rejected	Rejected	Rejected	Rejected	Rejected	Rejected	Rejected	*Confirmed*
Transportation	Rejected	Rejected	Rejected	Rejected	Rejected	Rejected	Rejected	*Confirmed*
Consumer Discretionary sector	*Confirmed*		*Confirmed*		Rejected		*Confirmed*	
Automobiles & Components	*Confirmed*	Rejected	Rejected	Rejected	Rejected	Rejected	*Confirmed*	*Confirmed*
Consumer Durables & Apparel	Rejected	Rejected	Rejected	Rejected	Rejected	Rejected	Rejected	*Confirmed*
Consumer Services	Rejected	Rejected	*Confirmed*	Rejected	Rejected	Rejected	*Confirmed*	Rejected
Retailing	Rejected	*Confirmed*	*Confirmed*	Rejected	Rejected	Rejected	Rejected	Rejected
Consumer Staples sector	Rejected		Rejected		Rejected		Rejected	
Food Staples & Retailing	Rejected	Rejected	Rejected	Rejected	Rejected	*Confirmed*	Rejected	*Confirmed*
Food, Beverage & Tobacco	Rejected	*Confirmed*	Rejected	Rejected	Rejected	Rejected	Rejected	Rejected
Household & Personal Products	Rejected	Rejected	*Confirmed*	Rejected	Rejected	Rejected	Rejected	Rejected
Health Care sector	Rejected		Rejected		Rejected		*Confirmed*	
Health Care Equipment & Services	*Confirmed*	Rejected	Rejected	*Confirmed*	Rejected	Rejected	*Confirmed*	Rejected
Pharmaceuticals, Biotechnology & Life Sciences	Rejected	Rejected	Rejected	Rejected	Rejected	Rejected	*Confirmed*	*Confirmed*
Financials sector	*Confirmed*		Rejected		Rejected		*Confirmed*	
Diversified Financials	*Confirmed*	*Confirmed*	Rejected	*Rejected*	Rejected	Rejected	*Confirmed*	*Confirmed*
Information Technology sector	*Confirmed*		Rejected		Rejected		*Confirmed*	
Semiconductors & Semiconductor Equipment	*Confirmed*	*Confirmed*	Rejected	Rejected	Rejected	Rejected	*Confirmed*	Rejected
Software & Services	*Confirmed*	Rejected	Rejected	Rejected	Rejected	Rejected	*Confirmed*	Rejected
Technology Hardware & Equipment	Rejected	Rejected	Rejected	Rejected	Rejected	Rejected	*Confirmed*	Rejected
Communication Services sector	*Confirmed*		Rejected		Rejected		*Confirmed*	
Media & Entertainment	*Confirmed*	Rejected	Rejected	Rejected	Rejected	Rejected	Rejected	*Confirmed*
Telecommunication Services	*Confirmed*	Rejected	Rejected	Rejected	Rejected	Rejected	Rejected	*Confirmed*
Utilities	*Confirmed*		Rejected		Rejected		*Confirmed*	
Utilities	*Confirmed*	Rejected	Rejected	Rejected	Rejected	Rejected	*Confirmed*	Rejected
Real Estate	Rejected		Rejected		Rejected		*Confirmed*	
Real Estate	Rejected	Rejected	Rejected	Rejected	Rejected	Rejected	*Confirmed*	*Confirmed*

Research question 2

To test **hypothesis 5,** that the estimated homoskedasticity will successively increase when moving from the market to the GICS sector, to the GICS industry, to the SARD grouping – we conducted White's test in **Section 6.3.** Through computing the White's test statistic (H_0 = homoskedasticity) for the different segmentations, we find significant p-values (indicating heteroscedasticity) on an aggregate market level, ~27% significant values on a sector-level, ~18% significant values on an industry-level, and ~9% significant values for SARD groupings. Hence, homoskedasticity seems to increase successively, and we can confirm **hypothesis 5**.

Section 6.4 included three different loss functions to test the prediction accuracy of the model and determine whether the model contains bias for over-/under valuation. To test **hypothesis 6,** that Predicted EV/EBITDA multiples developed from a regression analysis of fundamental value drivers will have significantly lower prediction errors in determining actual market multiples than estimates based on the average peer group multiples, the output from MAD, MSE and MAPE was analysed on a relative basis against average peer group multiples. The hypothesis is accepted/rejected if the predicted model has a significantly lower/higher error value than the peer group averages at a 5%

significance level. It is evident in **Table 6.12** that the regression model possesses a higher bias for the vast majority of sub-samples.

On an aggregated market level, **hypothesis 6** is rejected in all sub-samples for MSE, where 27 out of 33 sub-samples were significant on a 5% significance level. For MAD, the hypothesis is accepted in 4 sub-samples[37], however, none of these was significant on a 5% significant level. For MAPE, the same hypothesis is accepted for 8 sub-samples[38], but only three of these were significant on a 5% level. Similar results are yielded from the SARD grouping, where we test the SARD regression model against the SARD peer group average. **Table 6.12** shows that for MSE, **hypothesis 6** is rejected in all SARD industries, where 13 out of 22 sub-samples were significant on a 5% significance level (see **Appendix 7.2**). For MAD, the hypothesis is accepted in 3 SARD industries, with all of them being significant on a 5% significance level. For MAPE, the hypothesis is accepted for 7 industries, where 4 of these are significant on a 5% significance level. Whilst there are some sector and industries which performs better than the peer group average, it is evident that from a holistic viewpoint (see **table 6.12**), the peer group average has lower prediction errors than the regression model and thus **hypothesis 6** is rejected. The potential explanation as to why there is a diverging result between the different error functions will be further discussed in **section 7**.

To continue, **hypothesis 6** coincides with **hypothesis 7**, as we seek to confirm if the accuracy of the predicted EV/EBITDA multiples derived from a regression analysis of fundamental value drivers will show a reduction in significant prediction error on a relative basis as the segmentation progresses from the market level to GICS sector, to GICS industry, to SARD grouping. From **section 6.4**, we compare the output from the three loss functions across the different market levels (see **Table 6.12**). On an inter-segmentation level, the regression model yields a high error for some sectors and industries[39], however, it is evident that the error is decreasing gradually from market to sector to industry. Furthermore, as depicted in **Table 6.13**, we compare the error from the SARD regression model relative to the GICS regression (GICS), it is evident that the SARD model yields a lower

[37] Transportation, Consumer Durables & Apparel, Financial Sector, and Diversified Financials

[38] Consumer Discretionary Sector, Automobiles & Components, Consumer Durables & Apparel, Consumer Services, Retailing, Financial Sector, Diversified financials and Semiconductors & Semiconductor Equipment

[39] Sectors: Healthcare sector, Information Technology Sector; Industries: Commercial & Professional Services, Consumer Services, Health Care Equipment & Services, Pharmaceuticals and Biotechnology & Life Science

relative prediction error across all three loss functions. We see tendencies that the prediction error decreases gradually as homoskedasticity in the data sample is increased, however when looking at the p-values from the two sample mean t-test in **Appendix 7**, this hypothesis cannot be accepted on a 5% significance level. This will be further elaborated upon in **section 7**.

Table 6.12 – Hypothesis Testing (H6-H7)

Hypothesis testing (6-7)

		Subset A: Regression model GICS vs GICS Peer Group Average			Subset B: Regression SARD vs SARD Peer Group Average		
	N	MSE	MAD	MAPE	MSE	MAD	MAPE
Market	929	Higher than PGA <0.0001	Higher than PGA <0.0001	Lower than PGA 0.0057			
Energy Sector	26	Higher than PGA <0.0001	Higher than PGA <0.0001	Higher than PGA 0.0004			
Energy	26	Higher than PGA <0.0001	Higher than PGA <0.0001	Higher than PGA 0.0004	Higher than PGA 0.5432	Higher than PGA 0.0576	Higher than PGA 0.0116
Materials Sector	56	Higher than PGA 0.2204	Higher than PGA 0.0177	Higher than PGA <0.0001			
Materials	56	Higher than PGA 0.2204	Higher than PGA 0.0177	Higher than PGA <0.0001	Higher than PGA 0.0254	Higher than PGA 0.0022	Lower than PGA 0.3679
Industrials Sector	159	Higher than PGA <0.0001	Higher than PGA <0.0001	Higher than PGA 0.0052			
Capital Goods	109	Higher than PGA <0.0001	Higher than PGA <0.0001	Higher than PGA 0.0004	Higher than PGA 0.0448	Lower than PGA 0.0431	Lower than PGA 0.0478
Commercial & Professional Services	32	Higher than PGA 0.0039	Higher than PGA 0.0001	Higher than PGA 0.1027	Higher than PGA 0.0034	Higher than PGA 0.0018	Higher than PGA 0.0160
Transportation	18	Higher than PGA 0.3295	Lower than PGA 0.9099	Higher than PGA 0.1421	Higher than PGA 0.0248	Lower than PGA 0.0121	Lower than PGA 0.0498
Consumer Discretionary Sector	144	Higher than PGA <0.0001	Higher than PGA 0.0002	Lower than PGA 0.1774			
Automobiles & Components	18	Higher than PGA 0.0426	Higher than PGA 0.0655	Lower than PGA 0.9376	Higher than PGA 0.0424	Higher than PGA 0.1064	Higher than PGA 0.4135
Consumer Durables & Apparel	37	Higher than PGA 0.2180	Lower than PGA 0.8039	Lower than PGA 0.0427	Higher than PGA 0.0019	Higher than PGA 0.0032	Higher than PGA 0.0094
Consumer Services	39	Higher than PGA <0.0001	Higher than PGA <0.0001	Lower than PGA <0.0001	Higher than PGA 0.3262	Higher than PGA 0.3211	Higher than PGA 0.3219
Retailing	50	Higher than PGA 0.0003	Higher than PGA 0.0117	Lower than PGA 0.5983	Higher than PGA 0.1595	Higher than PGA 0.1569	Higher than PGA 0.1703
Consumer Staples Sector	65	Higher than PGA <0.0001	Higher than PGA 0.0013	Higher than PGA 0.9214			
Food Staples & Retailing	13	Higher than PGA 0.0111	Higher than PGA 0.0022	Higher than PGA 0.0030	Higher than PGA	Higher than PGA	Higher than PGA
Food, Beverage & Tobacco	39	Higher than PGA <0.0001	Higher than PGA <0.0001	Higher than PGA 0.1346	Higher than PGA 0.3217	Higher than PGA 0.3138	Higher than PGA 0.3085
Household & Personal Products	13	Higher than PGA 0.0265	Higher than PGA 0.0648	Higher than PGA 0.4465	Higher than PGA 0.2978	Higher than PGA 0.3036	Higher than PGA 0.4075
Health Care Sector	123	Higher than PGA <0.0001	Higher than PGA <0.0001	Higher than PGA 0.0002			
Health Care Equipment & Services	76	Higher than PGA <0.0001	Higher than PGA <0.0001	Higher than PGA 0.0016	Higher than PGA 0.3203	Higher than PGA 0.2970	Lower than PGA 0.3094
Pharmaceuticals, Biotechnology & Life Sciences	47	Higher than PGA <0.0001	Higher than PGA <0.0001	Higher than PGA 0.0206	Higher than PGA 0.0003	Higher than PGA 0.0000	Higher than PGA 0.0000
Financials Sector	40	Higher than PGA 0.3462	Lower than PGA 0.1549	Lower than PGA 0.0002			
Diversified Financials	40	Higher than PGA 0.3462	Lower than PGA 0.1549	Lower than PGA 0.0002	Higher than PGA 0.3242	Higher than PGA 0.3164	Higher than PGA 0.3157
Information Technology Sector	138	Higher than PGA <0.0001	Higher than PGA <0.0001	Higher than PGA 0.6412			
Software & Services	56	Higher than PGA <0.0001	Higher than PGA <0.0001	Higher than PGA 0.0532	Higher than PGA 0.1587	Higher than PGA 0.1493	Lower than PGA 0.1559
Technology Hardware & Equipment	45	Higher than PGA <0.0001	Higher than PGA <0.0001	Higher than PGA 0.0081	Higher than PGA 0.0283	Higher than PGA 0.0464	Higher than PGA 0.8354
Semiconductors & Semiconductor Equipment	37	Higher than PGA 0.0102	Higher than PGA 0.1040	Lower than PGA 0.2869	Higher than PGA 0.0016	Higher than PGA 0.0026	Higher than PGA 0.0326
Communication Services Sector	40	Higher than PGA 0.0001	Higher than PGA <0.0001	Higher than PGA 0.0111			
Telecommunication Services	10	Higher than PGA 0.0294	Higher than PGA 0.0258	Higher than PGA 0.1672	Higher than PGA 0.0509	Higher than PGA 0.0559	Higher than PGA 0.1667
Media & Entertainment	30	Higher than PGA 0.0012	Higher than PGA 0.0017	Higher than PGA 0.0838	Higher than PGA 0.0002	Higher than PGA <0.0001	Higher than PGA 0.0001
Utilities Sector	52	Higher than PGA <0.0001	Higher than PGA <0.0001	Higher than PGA <0.0001			
Utilities	52	Higher than PGA <0.0001	Higher than PGA <0.0001	Higher than PGA <0.0001	Higher than PGA <0.0001	Higher than PGA <0.0001	Lower than PGA <0.0001
Real Estate Sector	86	Higher than PGA <0.0001	Higher than PGA <0.0001	Higher than PGA <0.0001			
Real Estate	86	Higher than PGA <0.0001	Higher than PGA <0.0001	Higher than PGA <0.0001	Higher than PGA 0.0001	Lower than PGA <0.0001	Lower than PGA <0.0001

H_0: Mean error difference = 0
H_A: Mean error differecnce ≠ 0
Blue shaded number are not significanct on p<0.05

Table 6.13 – Hypothesis testing (H7)

		SARD MLR vs GICS MLR		
	N	MSE	MAD	MAPE
Energy	26	Lower than GICS	Lower than GICS	Lower than GICS
		0.0794	0.0530	0.3402
Materials	56	Lower than GICS	Lower than GICS	Lower than GICS
		0.2727	0.1489	0.0301
Capital Goods	109	Lower than GICS	Lower than GICS	Lower than GICS
		0.0449	0.0472	0.0586
Commercial & Professional Services	32	Higher than GICS	Higher than GICS	Higher than GICS
		0.1181	0.0307	0.8234
Transportation	18	Higher than GICS	Higher than GICS	Lower than GICS
		0.3253	0.2803	0.5767
Automobiles & Components	18	Higher than GICS	Higher than GICS	Higher than GICS
		0.8498	0.5350	0.1626
Consumer Durables & Apparel	37	Higher than GICS	Higher than GICS	Higher than GICS
		0.0856	0.0998	0.1354
Consumer Services	39	Lower than GICS	Lower than GICS	Lower than GICS
		0.3265	0.3394	0.3562
Retailing	50	Higher than GICS	Higher than GICS	Higher than GICS
		0.1595	0.1635	0.1718
Food Staples & Retailing	13	Higher than GICS	Higher than GICS	Higher than GICS
		0.3409	0.3413	0.3402
Food, Beverage & Tobacco	39	Lower than GICS	Lower than GICS	Higher than GICS
		0.3224	0.3561	0.5184
Household & Personal Products	13	Lower than GICS	Lower than GICS	Lower than GICS
		0.3379	0.2901	0.3186
Health Care Equipment & Services	76	Lower than GICS	Lower than GICS	Lower than GICS
		0.3216	0.3568	0.9105
Pharmaceuticals, Biotechnology & Life Sciences	47	Lower than GICS	Lower than GICS	Lower than GICS
		0.9373	0.5976	0.6689
Diversified Financials	40	Lower than GICS	Higher than GICS	Higher than GICS
		0.3245	0.3360	0.4023
Software & Services	56	Lower than GICS	Lower than GICS	Lower than GICS
		0.1590	0.1636	0.2193
Technology Hardware & Equipment	45	Higher than GICS	Higher than GICS	Lower than GICS
		0.5566	0.8951	0.3104
Semiconductors & Semiconductor Equipment	37	Lower than GICS	Lower than GICS	Higher than GICS
		0.3402	0.5064	0.4441
Telecommunication Services	10	Higher than GICS	Higher than GICS	Higher than GICS
		0.2988	0.2411	0.4814
Media & Entertainment	30	Higher than GICS	Higher than GICS	Higher than GICS
		0.1611	0.1565	0.2036
Utilities	52	Lower than GICS	Lower than GICS	Lower than GICS
		0.0288	0.0070	0.0085
Real Estate	86	Lower than GICS	Lower than GICS	Lower than GICS
		0.1703	0.2892	0.1605

H_0: Mean error difference = 0

H_A: Mean error differecnce ≠ 0

Blue shaded number are not significanct on p<0.05

Chapter 7
Discussion

This section will build on the findings of our quantitative study on the underlying relationship between the EBITDA-multiple and key underlying value drivers and observed variances across different groupings. We wish to discuss the fact that utilising the SARD approach seems to increase homoskedasticity in the dataset, and to what extent homoskedasticity alters the predictive power of the regression-based model. Our hypothesis-driven approach is grounded in financial and econometric theory, whereby we have suggested a best-possible methodology for exploring the relationship between value drivers and the effect of different peer group segmentations. Now we will discuss the implications of our findings, the generalisability of our models, and the reason for rejections in several hypotheses. Ultimately, we wish to discuss to what extent our results are aligned with the theory.

7.1 SARD

Knudsen et al. (2017) introduced the SARD approach as a novel fundamental peer group selection method. The authors found in their study that applying SARD for peer grouping leads to significantly more accurate multiple prediction accuracy as compared to the industry affiliation approach when using the S&P 1500 composite index as a data sample. Whilst our study applies the same methodological approach as Plenborg et al. (2017) for selecting peer groups based on the sum of absolute rank differences, it does not apply SARD for its designed purpose. In this research, we have utilised the SARD method with the objective to enhance the homogeneity among the sub-samples. The purpose was to investigate whether peer groups with similarities in key underlying value drivers, selected through the SARD approach, can result in an increase in homoskedasticity within the data samples, and subsequently determine whether enhanced homoskedasticity through the SARD approach can lead to more statistically significant output when running linear regression. It is argued that this study's application of the SARD approach extends the original approach's usefulness.

7.2 Data Input Properties

As outlined in **section 1.3**, the study was delimited in terms of sample size, data quality and time period. We used the Bloomberg Terminal (2023) to gather financial data on the S&P 1500 composite index constituents, from which we formed a trimmed census. This study has sought to reduce selection bias to the extent possible. It is argued that no sector or industry was discriminated with an excessive amount of elimination in the construction of the final data set. The elimination of data points was done on the same criteria across all sectors and industries. Also, the same elimination method was applied in the robustness test in **section 6.5**. Furthermore, in addition to the firms that were excluded for practical purposes, two outliers were excluded from the final data sample. Whilst this might affect our empirical results, the effect is deemed to be little. Furthermore, the elimination of these outliers increased the normality in the data sample.

The study utilised a three-year period (2020-2022) to aggregate the dependent and independent variables. The normalisation was conducted to address short-term cyclicality and minimise potential biases resulting from collecting data from a single period. For EV/EBITDA we used a forward-looking time horizon, utilising Bloomberg Terminal BEst EBITDA estimates. Whilst subjectivity is present in these estimates, they are based on analyst consensus and thus deemed representable. For growth, we also used the same BEst EBITDA estimates to calculate the forward-looking EBITDA CAGR. For profitability and risk, we used trailing full-year and current full-year figures, respectively. As previously noted, this was due to data availability issues in the Bloomberg Terminal.

7.3 MLR Assumptions

While we have deemed our dataset relatively linear and non-violating of MLR.1, we have through shifted power transformation found a few examples of non-linearity within the WACC (LFY) variable, which generally coincides with the significant heteroskedastic segmentations. This certain degree of non-linearity can, in theory, reduce goodness-of-fit (affecting R^2 and adjusted R^2 values) and prediction accuracy for the given segmentations and aggregate market level, and lead to a degree of bias in coefficients.

We furthermore have inferred the potential issue of simultaneity, in which correlated independent variables can potentially inflate different test statistics such as White's test for heteroskedasticity.

Variance Inflation Factor (VIF) for almost all segmentations was deemed low/moderate, except for the Telecommunication Services SARD grouping, where the auxiliary regressions are moderate-to-high for growth and profitability, and over 5 for risk. We furthermore see the mentioned SARD grouping has the overall highest MLR R^2 of ~88%, where we question the validity of the output due to the high potential of multicollinearity. Potential omitted variable bias leads to the question of whether we have highly correlated omitted variables which can explain dynamics in variance – hence potentially leading to biases in test statistics for the independent variables. But our scope pertains to assessing the relationship between the EBITDA-multiple and the theoretically grounded key value drivers. What we find is that different segmentations lead to different R^2 values, and different goodness of fit, which further argues the research gap within regressing underlying value drivers – where the choice of value drivers and peer groups seemingly has a detrimental effect on model fit, which we point out as an interesting topic for further research.

As a last point, we find some instances of non-normality in residuals, such as for the Industrials Sector and Health Care sector, which can lead to a few biased coefficients and blur interpretability, but not enough to reject the assumption of normality. This was pre-emptively remedied through a logarithmic transformation of the dependent variable, which compresses the variability of data and made the residuals more linear. Furthermore, segmentation helps us sub-divide model output and identify patterns in segmentations with better goodness-of-fit – where we also find there to be is that fewer SARD groupings show non-normality compared to industry groups with similar sample sizes.

7.4 Underlying Value Drivers

As previously elaborated upon, and with reference to **table 6.4**, we see varying evidence for our proposed **hypotheses 1-4**. For growth, we generally see slightly positive relationships with the EBITDA-multiple, however less than half of the subsample group's beta coefficients are significant. For profitability, the relationship generally hovers close to 0, with even fewer significant beta coefficients. But the most contradictory finding was the mostly positive relationship to risk, with more significant beta coefficients than for growth and profitability, respectively.

Theory suggests (see equation 2.12) that an increase in our selected risk proxy, WACC, should have a negative impact on the EBITDA multiple, given the increased cost associated with sourcing new capital[40]. However, the risk-return trade-off suggests that potential return increases for higher levels of risk. This implies our model may have a potential issue of simultaneity[41], in which risk may not be fully independent of the other independent variables.

Whilst this may be countered by finding instrument variables which are correlated to the WACC variable, but not to the error terms of the regression model, we argue this would be impractical given that firm value drivers generally have complex interrelationships. Finding an instrument variable that is not correlated to the regression models error terms, and which can be mathematically derived from the EBITDA-multiple, is not feasible and would likely lead to scrapping the use of the EBITDA-multiple in favour of other multiples. We have laid forth a theory that suggests that when doing the regression models in a standardised format across different peer group segmentations, the EBITDA multiple is the most applicable. The potential issue of simultaneity is also intuitively evident when using accounting measures, in that different items are typically tied together through the interrelationship and methodology of financial statements.

As the beta coefficients for risk were mostly positive, we conducted a shifted power transformation to test for potential curvilinearity within the independent variable set. The few segmentations, as per **Appendix 8**, that showed curvilinearity, match several segmentations that showed significant joint heteroskedasticity as per **Appendix 5** (Information Technology Sector, Software & Services, Capital Goods). Whilst these are not necessarily only due to sporadic curvilinearity in WACC, this illustrates an advantage of the SARD groupings, in which peers are chosen by the sum of absolute rank differences in the chosen ranking criteria, which we put as the independent variables.

The varying degree of goodness-of-fit, beta coefficients and significance across all the different segmentations, indicates that our proxies have different explanatory power across the market, different sectors, industries and SARD groupings. When disregarding potential omitted variables, this makes intuitive sense in that theory and empirical studies suggest that different industries are

[40] As argued by Higgins (2005), Hussain & Chakraborty (2010), Ross (2007), Loughran, Wellman (2011), and Damodaran (2012), in **section 5.2.2**

[41] Simultaneity occurs when two or more variables are mutually dependent, where they simultaneously affect each other

associated with different best multiples[42]. Furthermore, different industries may have different sensitivities to underlying value drivers. Some industries may naturally have converged towards long-term stable growth rates that are higher than others, whereby a 1% increase in growth in one industry, may have a bigger impact than a 1% increase in another. Differences can also be inter-segmental, whereby firms differ in terms of growth, profitability, and risk. A firm that e.g., has seen lower valuations due to lack of profitability, could suddenly have gotten a boost in valuation if increasing profitability – marginally more than peers. With a cross-sectional dataset, our MLR models may therefore assign different goodness-of-fit statistics, in that it may be harder to identify patterns for some companies within certain segmentations. Nevertheless, it is evident that our models do not fully explain the relationship between EBITDA multiples and our key underlying value drivers.

7.5 Peer Group Selection & Homoskedasticity

As mentioned in **section 6.7** hypothesis 5 was confirmed, as we saw fewer detections of heteroskedasticity (i.e., less $p < 0.05$) when moving from an aggregated market level to sector, to industry, to SARD groupings. We do however note that the weighted average of WT statistics for sectors (WT = 9.18) is slightly lower than for the industries (10.29).

There are several factors which could play into this discrepancy, with one of them being sample sizes. White's test is a goodness-of-fit test based on the sum of squared residuals in the MLR model (using actual independent variables values), whereby smaller sample sizes may exude higher variances in which smaller cases of heteroskedasticity may be hard to detect, but sensitivity to outliers can be larger. Furthermore, we see in **appendices 3** and **4** that our underlying value drivers are correlated to varying extents (and with few exceptions, not highly correlated), which makes sense in that they are connected through firm's activities, e.g., that higher growth may lead to higher profitability if outgrowing marginal costs. Whilst theory suggest these are key value drivers, there is an omitted variable bias in that other drivers affect EBITDA-multiples as well. Omitted variables may lead to a wider confidence interval, due to upward-biased standard errors, which ultimately makes it more difficult to draw statistical conclusions. Should some of these be highly correlated

[42] As argued by Liu , Nissim, & Thomas (2002); Lie & Lie (2002); Plenborg & Pimentel (2016); Fernandez (2001); Gupta (2018) in **section 5.2.1**

with the independent value drivers, hence potentially explaining some of the movements in the independent variable, White's test may over- or underestimate heteroskedasticity for said variable with the bias potentially being more pronounced for smaller sample sizes. So, as our p-value hypothesis testing led to marginally less significant heteroskedastic error terms for industries than sectors, the confirmation still stands.

Furthermore, for the market segmentation, the larger sample size compared to the other segmentations may lead to a potentially higher sensitivity to heteroskedastic sub-parts of the dataset. This is, as White's test now is working with the whole dataset, it might reject the null hypothesis of homoskedasticity due to some parts being heteroskedastic, e.g., the industrials sector. So, when sub-dividing the dataset, we are able (on average, and except for industry) to achieve seemingly more homoskedastic peer groups. Output from the different multiple linear regressions and subsequent accuracy tests indicates tendencies for peer groups with more homoskedastic error terms to result in marginally better model prediction accuracy than peer groups with more heteroskedasticity.

Whilst we may infer a pattern between higher homoskedasticity and better ability for the model to significantly explain the relationship between dependent and independent variables, as well as the ability to accurately predict EBITDA multiples, we have not found evidence to conclude on this interplay. First off, whilst e.g., the models segmented by SARD groupings may have more homoskedastic residuals than the industry segmentations – industry segmentations were already deemed relatively homoskedastic. So, we see tendencies, but the effect of different degrees of homoskedasticity has not necessarily been explored, and it may e.g., have a marginally diminishing effect over a certain threshold. Additionally, we are careful not to draw any definite conclusions across subsets of data of different sizes – as changing sample sizes affect test statistics to varying degrees. So, while we see tendencies, there are also ambiguities. More significant F-statistics and somewhat better prediction accuracy indicate that our models should be better at finding patterns as well as making use of independent variables to explain patterns, for GICS segmentations. Furthermore, lower R^2 for SARD groupings indicate that less of the variability in the independent variables explains the variability in the dependent variable. However, we see tendencies that more homogenous datasets seem to have better prediction accuracy (with SARD groupings being marginally more accurate) and are likely less prone to overfit model training data. Given the fact that most subsamples have significant intercept coefficients, whilst many of the subsample's beta coefficients are insignificant at $p<0.05$, we see a tendency of omitted variable bias. However, a point to note, is that increased homoskedasticity on its own, does not constitute a better dataset for linear regression models. The other MLR assumptions still need to be met.

Above indicates that when decreasing the number of firms in a peer group, the homogeneity is increased, which subsequently tends to increase the homoskedasticity in the data samples. Following this notion, a more detailed GICS classification system could have been applied. Herein, GICS 6-digit or even GICS 8-digit would divide the S&P 1500 index into 74 and 163 sub-samples respectively. Increasing the granularity of the SARD grouping to the same extent, the homogeneity is deemed to further increase which subsequently is likely to increase the homoskedasticity in the data samples. While this is expected to enhance the extent to which the variations in EV/EBITDA can be explained by its fundamental value drivers, it would affect the statistical advantages of having a larger sample size.

GICS segmentations generally exhibit slightly better goodness-of-fit statistics, i.e., variation in dependent variables is better explained by independent variables when segmenting by GICS. However, SARD groupings generally have smaller prediction errors when benchmarked against peer group averages. We note that the SARD groupings have more homoskedastic and normal error terms, meaning GICS segmentations may be more liable to overfitting the data.

7.6 Model Accuracy

As outlined in the previous section, different peer groups were formed to test the prediction accuracy for the regression-based model relative to peer group averages across different segmentation levels. This section seeks to discuss the regression-based model's accuracy from a holistic viewpoint, whilst also comparing the accuracy across the different segmentation levels relative to the peer group average model.

From **section 6.4.2-6.4.4** the accuracy test showed that the derived multivariate regression model accuracy, based on beta coefficients, is far from perfect. Whilst the methodological approach provides some statistical explanatory power of the relationship between EV/EBITDA and its fundamental value drivers, the degree of explanatory power is deemed weak in general. From a holistic viewpoint, it is evident that the more easily applicable peer group averages, significantly outperform the regression models in terms of predicting the observed EV/EBITDA multiple. When comparing the relative output of the three accuracy tests conducted, it is evident that the regression-based model performs poorly as per MSE, better in terms of MAD, and best in terms of MAPE. As outlined in **section 6.4.1**, this can be explained by the underlying construct of the different loss functions. Since MSE is a quadratic loss function, it possesses the advantage of penalising outliers by assigning larger weights to large errors and lower weights to lower errors. As outlined in **section 5.3.2** we took a conservative approach when removing outliers due to the statistical limitation of using smaller sample sizes.

Hence when comparing the EV/EBITDA MSE of our regression model relative to the EV/EBITDA MSE of peer group averages, it makes intuitive sense that the regression model is inferior. Furthermore, the notion that the regression-derived EV/EBITDA yields the lowest relative error in terms of MAPE also makes intuitive sense as MAPE compares the prediction accuracy of the regression-based model against the observed multiple, without being influenced by random variations in specific values in each sample. I.e., since MAPE is calculated using the percentage error between the predicted and observed values, it is not influenced by the magnitude of the values in the data sample, but rather by the relative error between the predicted and actual values. Nevertheless, based on the relative error output, it is deemed to be beneficial to use the average peer group model to predict EV/EBITDA, on an aggregated level.

However, as discussed in **section 7.3**, when moving from market to sector to industry to SARD industry, we see tendencies that the homoskedasticity in the data samples gradually increased. As outlined in **section 6.4** the error term for all three tests is also gradually decreasing as we move vertically across the different peer groups. Hence, we see tendencies that data sample homogeneity influences the model accuracy.

The regression-based model is deemed to have limited explanatory power on a market level. However, when moving forward to a more granular level, it is evident that the model yields lower prediction errors for some specific sectors and industries. Specifically, the materials sector and the financial sector. For the material sector, this is counterintuitive, thus it holds insignificant beta coefficients for all three value drivers. However, the intercept of the Materials Sector is high and significant at $p < 0.05$. Therefore, a probable explanation for the low prediction error is thus omitted variable bias. On the contrary, for the financial sector, the result is in line with the multiple regression output since it holds significant beta coefficients for both growth and profitability together with significant goodness-of-fit statistics on a 1% level.

On an industry level, the model yields the lowest error across the loss functions for Consumer Durables & Apparel, and Transportation. This is counterintuitive thus both industries have insignificant beta coefficients at $p > 0.05$, together with relatively low R-squared. Hence, when comparing the output from the accuracy test with the output from the multivariate regression model on an aggregated level, it is clear that insignificant beta coefficients for the underlying value drivers are in some instances able to produce significant predictions with low prediction error, relative to the peer group average model.

Finally, when comparing the prediction accuracy yielded from the SARD regression-based approach to the accuracy yielded from the GICS regression-based approach, it is evident that the SARD-based model yields a lower error for the three loss functions on an aggregated level. However, whilst the performance of the SARD-based model is superior to GICS based model in terms of error, many of the SARD industries have insignificant beta coefficients. Furthermore, viewing the two sample mean t-test statistics between the two models in **Appendix 7**, it is evident that very few of the MSE, MAD, and MAPE have a significant mean differences. Hence, on an aggregated level we cannot reject the null hypothesis that the mean difference between the SARD-based model and the GICS model is equal to 0. However, we argue that there are tendencies as to when homogeneity in the data sample is gradually increased by using SARD, the error term is decreasing relative to the peer group average model.

One could have tested the accuracy through an additional in-sample test, by regressing the observed EV/EBITDA multiple as the dependent variable, against the predicted EV/EBITDA as the independent variable. This was deemed unnecessary, as similar accuracy conclusions could be drawn from the multivariate regression output. It is important to note that the in-sample testing conducted in this research holds potential limitations and biases, such as omitted variable bias and limited generalisability. Omitted variable bias seems to present as the model generated high and significant intercept coefficients, such that it captures chance variation and outliers rather than true underlying patterns in the data. This also points towards a high degree of omitted variable bias. Limited generalisability of the specific regression-based model is accepted thus this paper is delimited to only trying to explain the relationship between EV/EBITDA and growth, profitability, and risk, for public firms in the US market. Whilst an out-of-sample test is argued to be more statically correct, it is deemed to not be applicable in this specific study. This is due to the initial data sample being a trimmed census of the S&P 1500 composite index, and it is deemed difficult to find private firms at comparable sizes thus private firms are likely to be significantly smaller, have limited public information, and trade at illiquidity discounts (Comment, 2012; Pratt, Reilly, & Schweihs, 2008).

7.7 Generalisability of Study

Although there are limitations to this study resulting from the specified delimitations in **section 1.3**, it is believed that this research provides both theoretical and practical value from a comprehensive perspective. The study's generalisability is considered in two aspects: model generalisability and methodology generalisability.

Firstly, using the S&P 1500 composite index as the data sample, we have delimited the analysis to public US firms. This was motivated by all US firms following the same accounting standards (GAAP) and the large number of firms compromising this index. However, this delimitation removes the possibility to generalise the model and its output across other geographies, since different markets have different market-specific discrepancies. Secondly, since the study is delimited to only analyse public firms, the model might have difficulties when generalised to private firms. One of the reasons is that private firms are often trading at an illiquidity discount (Comment, 2012; Pratt, Reilly, & Schweihs, 2008). Thirdly, the scope of this study is limited to the EV/EBITDA valuation multiple and its theoretically derived value drivers. Therefore, any generalisations made from the model and findings are restricted to this specific type of valuation multiple and cannot be extended to other valuation metrics. Lastly, the empirical evidence from this study indicates a weak predictive power of the theoretically derived fundamental value drivers. These findings question the strength of the relationship between growth, profitability, and risk as explanatory variables for EV/EBITDA.

Our study suggests that there may be omitted variables that are not accounted for, indicating a more complex relationship than captured by the derived model. This is further substantiated in the robustness test conducted in **section 6.6**, where the models were tested for an expanded time horizon. In reference to **Table 6.10** and in line with prior studies (Damodaran A., 2012; Harbula, 2009), we see pronounced differences in coefficients when expanding the timeline, which questions the robustness of the models. This leads us to restrict the generalisability of the models to the given timeline of the data input to which the models have been applied.

On the other hand, the methodological approach employed in this study is considered to have wide generalisability. The study adheres to common best practices within the OLS regression framework and has complied with statistical assumptions on a general basis. Furthermore, the study has found a potential extended application of the SARD methods which goes beyond the method's intended purpose. This study has applied the SARD method to segment data which is later used as input for a regression model. The study has demonstrated that the SARD method potentially can segment data such that it meets the MLR.5-6 assumptions better. More specifically, through applying the SARD approach we see tendencies of increased homoskedasticity and normality in error terms, which subsequently show tendencies of increased model prediction accuracy with less susceptibility for model overfitting. Ultimately, the application of the SARD model for segmenting data based on financial metrics has potential applications for regression analyses in a financial context.

However, to further enhance the generalisability of the SARD method in a regression context, one approach could be to implement the principle of randomness when selecting the target firms. This could be done by assigning an individual number to each firm and using a random number generator. Doing such would decrease the potential of systematic biases in the data which could affect the accuracy of the regression analysis. Hence, by using random sampling, the results are more prone to be generalisable to a wider population, subsequently allowing for more statistical inferences to be drawn from the analysis. Regardless of whether the target firm is picked with randomness or not, tendencies infer that normality and homogeneity are likely to be higher for the SARD grouping compared to the GICS segmentations.

Chapter 8
Conclusion

We have set out to explore the capacity of linear regression models to explain and predict variability in EBITDA-multiple valuations through regressing key underlying value drivers' profitability (3-year normalised ROIC), growth (3-year normalised EBITDA CAGR), and risk (WACC LFY), on a non-discriminatory trimmed census sample of S&P 1500 companies in US public equity markets. Further to this, we have analysed the impact of peer group segmentations on MLR.1-6 assumptions, goodness-of-fit statistics, and prediction accuracy. Rooted in extensive literary coverage on relative valuation, motivated by an observed research gap within the ability of linear regression models to predict valuation multiples, and inspired by a similar test having been conducted on Indian equities (Gupta, 2018) – we have conducted thorough hypothesis testing, presented findings and successively discussed their implications.

When contextualising our research within relevant precedent theory, we found that two schools of thought in financial literature argued that peer groups either should be constructed based on industry affiliation, or through similarities in key underlying value drivers. Within the second school of thought, growth, profitability, and risk were generally drawn forth as the most important value drivers. We suggested that ROIC, EBITDA CAGR and WACC were the most suitable proxies for our model. These could be mathematically derived from the EBITDA-multiple and were supported as highly suitable proxies in financial theory. We argued for the EBITDA multiple as the most appropriate multiple for our analysis of companies with varying characteristics, as it is less sensitive to differences in accounting practices and capital structure.

The SARD method as proposed by Plenborg et al. (2017), argues for higher comparability and homogeneity in peer group selection when based on the sum of absolute rank differences in chosen criteria. With scarce literature on the SARD approach's applicability in regression models, we aimed to test the impact of peer groups constructed through the SARD approach, on MLR.1-6 assumptions, goodness-of-fit and model prediction accuracy. In line with previous methodology, and in the context of our chosen key value drivers, we constructed peer groups based on similarity in growth, profitability, and risk.

Having mapped out relevant theoretical foundations and literature, and proposed appropriate methodology and hypotheses, we conducted a deductive approach of hypothesis testing to answer our research questions. Our first research question was comprised of four hypotheses, where we expected growth and profitability to have a significant positive relationship on the EBITDA-multiple, with a t-statistic ≠ 0 (**hypotheses 1 and 2**). For risk, we expected a significant negative relationship with t-statistic ≠ 0 (**hypothesis 3**). Lastly, **hypothesis 4** expected the independent variables to have a jointly significant (95% confidence interval) impact on the EBITDA-multiple with F-statistic ≠ 0. For **hypotheses 1**, **2** and **4** the hypothesis was confirmed for a range of peer group segmentations, but not all. Specifically, **Hypothesis 4** was confirmed as the F-stat was significant at $p < 0.05$ for 64% of the subsamples on the aggregated market level and for 55% of the SARD groupings. However, significant R^2 and adjusted R^2 values were on average relatively low. **Hypothesis 3** was rejected, to our surprise, WACC seemed to have a significant positive impact on the EBITDA multiple for several groupings.

Our second research question and **hypothesis 5** expected homoskedasticity to increase when moving from an aggregated market segmentation to GICS 2-digit sector group segmentations, to GICS 4-digit industry group segmentation, to SARD groupings selected based on peers with the most similar sum of absolute rank differences in 3-year normalised ROIC, 3-year normalised EBITDA CAGR and WACC LFY. With less significant ($H_0 = homoskedastic$) test statistics when moving from an aggregate market level to sector, to industry, to SARD groupings – the hypothesis was confirmed.

Hypotheses 6 through **7** pertain to the predictive power of our estimated linear regression models, benchmarked against average peer group multiples across the different segmentation levels. **Hypothesis 6** proposed that our models would have significantly lower prediction errors than simple mean averages, calculated on a mean absolute deviation (MAD), mean squared error (MSE), and mean absolute percentage error (MAPE) basis. We rejected the hypothesis, as our model in most instances did not produce significantly lower prediction errors, compared to peer group averages. For **hypothesis 7** we tested if the accuracy of predicted EBITDA-multiples derived from a regression analysis of fundamental value drivers would show a successive reduction in significant prediction errors, on a relative basis, as the segmentations progress from an aggregate market level to the GICS sector, to the GICS industry, to the SARD groupings. Albeit tendencies were observed, as prediction errors on average decreased gradually for segmentations with more homoskedastic error terms, differences were in several instances not significant (95% confidence interval) when conducting two-sample t-tests.

Without appeal to novelty, we argue our findings indicate the SARD approach can be utilised to create more homogenous datasets which better fit MLR.5-6 assumptions, hence potentially leading to lower prediction errors, when applying linear regression. Out of the different segmentations, the SARD approach tends to have more homoskedastic error terms, normal error terms, and lower prediction errors. The GICS industry segmentation, exhibits slightly better goodness-of-fit statistics, with more heteroskedastic and non-normal error terms, suggesting it is more liable to overfitting the data.

In terms of the relationship between the EBITDA-multiple and the underlying value drivers, we find varying relationships across all the different segmentations, both in terms of significance, R^2 and adjusted R^2 values, and coefficient estimates. In many instances we find significant relationships, yet explanatory power is generally deemed relatively low. Whilst we generally deem our models to satisfy the MLR.1-6 assumptions, we do find a few instances of curvilinearity and heteroskedasticity which can lead to biased output for the given segmentations as well as on an aggregate market level, as well as indications of omitted variable bias (such as all intercepts being highly positive and significant, but fewer significant beta coefficients). We find that linear regression models regressing the chosen value proxies can explain some of the variance in EBITDA-multiple valuations but argue that similar model adaptations or future extensions should ensure to contextualise findings against other types of valuations, which is a current best practice in investment banking.

Linear regression models are actively used within finance, much due to practicality and interpretability. Our research proposes a framework, and reflects on important considerations, for approaching statistical modelling of publicly available accounting data. We show that the SARD approach can be utilised to better adhere to the MLR. 5-6 assumptions of linear regression models. We suggest that the application of the SARD approach for segmenting data based on financial metrics has potential application for regression analyses in a financial context, which future extensions can build upon. Whether for similar models, extensions, or other more complex models that handle cross-sectional data, our study offers considerations that contribute to the ever-evolving realm of financial theory.

Further Research

There is a range of natural progressions from our research, either in the form of minor adjustments, extensions, or adjacencies, which will guide the structure of our suggestions. This section aims to spawn inspiration for future research exploring similar phenomena as our paper, as well as touch upon relevant considerations.

In terms of adjustments, several elements of our variables can be potentially revamped. The first one is tied to our finding of increased homoskedasticity in error terms showing potentially lower model prediction errors. By testing for different criteria in the SARD approach (where we used growth, profitability, and risk), one may explore how these affect the MLR.5 and MLR.6 assumptions[43], and subsequent goodness-of-fit statistics and prediction errors. Here it is important that the criteria are theoretically grounded and must be guided by an approach with limited selection bias, which can be mitigated by applying randomness within sampling. Here it would also be interesting to apply similar-sized samples for potentially increased comparability.

In our initial presentation of relevant precedent literature for our research, we pointed out that a range of multiples was backed by both theory and empirical evidence, such as the P/E multiple or P/B multiple, which would be interesting to explore. However, as evident from the mathematical derivation of the underlying value drivers from the EBITDA multiples, this would potentially require changing proxies for underlying value drivers - or for some multiples, even changing underlying value drivers. As such, the scope of the research will also be skewed.

If keeping the multiple constants, a natural adjustment from our research would be to explore the relationship between the EBITDA-multiple and our proposed underlying value drivers, but with other proxies. However, comparability to our research may be biased by the fact that different proxies may affect the MLR.1-6 assumptions differently – hence, findings would be interpretable on

[43] MLR 5 covers the homoskedasticity assumption, and MLR 6 covers the normality assumption.

a stand-alone basis. Such proxies can still follow our methodology, in which certain issues such as simultaneity and omitted variable bias would likely still be present.

Should it be of interest, future research may potentially change up the underlying value drivers against the EBITDA-multiple. The same point as above stands, in which comparability may be limited due to different value drivers affecting the MLR-1.6 assumptions differently, and it must be possible to derive the underlying value drivers from the EBITDA multiple. We pointed out the fact that theory suggests that different industries may best be analysed through different multiples. As such, different combinations of multiples and value drivers can be tested across different segmentations such as e.g., variations of GICS segmentations. Here it would be important to ensure comparability of output, by limiting unnecessary variations across the input dataset (e.g., lacking values for some input values for some companies). One variation of such research could be to see how different variations of such variables would be handled by MLR.1-6 assumptions for linear regression models.

In terms of extensions, variations of the abovementioned come first to mind. One may add additional underlying value drivers, which could potentially mitigate the omitted variable bias. Still, the issue of simultaneity persists, and one must ensure that MLR.1-6 assumptions still hold, e.g., that the added underlying value drivers do not exhibit perfect multicollinearity with the existing underlying value drivers of the model. Furthermore, robustness may be researched by further extending timelines, or by applying the model to datasets from other equity markets, such as e.g., the European market.

In terms of adjacencies, other models may be tested to explore similar phenomena as of our research, which we will comment on a more holistic level – rather than delving into the peculiarities, advantages and disadvantages of different models. Pertaining to cross-sectional data, one could test out models which assume curvilinear relationships, such as polynomial regression models. Should one arrange output as binary, models such as Random Forest algorithms are good at analysing complex patterns within cross-sectional data. One could also arrange one's dataset as longitudinal data and study the output of time-series models such as e.g., ARIMA models. In the case of adjacencies, it is intuitive that comparability to our research would be diminished.

Bibliography

Abid, F., & Mseddi, S. (2010). The impact of Operating and Financial Leverages and Intrinsic Business Risk on Firm Value. *Available at SSRN: Abid, Fathi and Mseddi, Slim, The Impact of Operating and Financial Leverages and Intrinsic Business Risk on Firm Value (2004). Available at SSRN: https://ssrn.com/abstract=942029 or http://dx.doi.org/10.2139/ssrn.942029*.

Abukari, K., Jog, M., & McConomy, B. (2000). The Role and the Relative Importance of Financial Statements in Equity Valuation. *SSRN Electronic Journal*.

Achleitner, A.-K., Braun, R., & Engel, N. (2011). Value creation and pricing in buyouts: Empirical evidence from Europe and North America. *Review of Financial Economics*, 146-161.

Agresti, A., Franklin, C., Klingenberg, B., & Posner, M. (2018). *Statistics - The Art and Science of Learning from Data* . Harlow: Pearson Education limited .

Alexander, C. (2008). *Quantitative Methods in Finance.* West Sussex: John Wiley & Sons Ltd.

Alford, A. (1992). The Effect of the Set of Comparable Firms on the Accuracy of the Price-Earnings Valuation Method. *Journal of Accounting Research*, 30(1), 94-108.

Ang, A., & Chen, J. (2007). CAPM over the long run: 1926-2001. *Journal of Empirical Finance*, 14, 1-40.

Asquith, P., Mikhail, M., & Au, A. (2005). Information Contest of Equity Analyst Reports. *Journal of Financial Economics, Vol. 75, No. 2*, 245-282.

Bahari, S. (2010). Qualitative versus Quantitative Research Strategies: Contrasting Epistemological and Ontological Assumptions. *Jurnal Teknologi*, 52: 17-28.

Bajaj, M., Denis, D., Ferris, S., & Sarin, A. (2001). Firm Value and Marketability Discounts. *SSRN Electronic Journals*.

Baker, M., & Ruback, R. (1999). Estimating Industry Multiples (Working paper). *Harvard University*.

Basu, S. (1977). Investment Performance of Common Stocks in Relation to Their Price-Earnings Ratios: A Test of The Efficient Market Hypothesis. *The Journal of Finance*, 663-682.

Baumol, W., & Malkiel, B. (1967). The Firm's Optimal Debt-Equity Combination and the Cost of Capital. *Quarterly Journal of Economics, Vol. 91, No. 4*, 547-578.

Baxter, N. (1967). Leverage, Risk of Ruin, and the Cost of Capital. *Journal of Finance, Vol. 22, No. 3*, 395-404.

Beaver, W., & Morse, D. (1978). What Determines Price-Earnings Ratios? *Financial Analysts Journal, 34(4)*, 65-76.

Begley, J., & Feltham, G. (2002). The Relation between Market Values, Earnings Forecasts, and Reported Earnings. *Contemporary Accounting Research 19*, 1-48.

Berk, J., & DeMarzo, P. (2019). *Corporate Finance, Global Edition (5th Edition).* Essex: Pearson Education Limited.

Bernard, V. I. (1994). *Accounting-Based Valuation Methods, Determinants of Markets to Book Ratios, and Implications for Financial Statement Analysis (Working Paper).* Michigan: University of Michigan.

Bernström, S. (2014). *Valuation: The Market Approach (1st. edition).* John Wiley & Sons.

Betton, S., Eckbo, B., & Thorburn, K. (2009). Merger negotiations and the toehold puzzle. *Journal of Financial Economics*, 91(2), 158-178.

Bhojraj, S., & Lee, C. (2002). Who Is My Peer? A Valuation-Based Approach to the Selection of Comparable Firms. *Journal of Accounting Research*, 40, No. 2: 407-439.

Bhojraj, S., Lee, C., & Oler, D. (2003). What's My Line? A Comparison of Industry Classification Schemes for Capital Market Research. *Journal of Accounting Research*, 41(5), 745-774.

Bierman, H., & Thomas, J. (1972). Ruin Considerations and Debt Issuance. *Journal of Financial and Quantitative Analysis, Vol. 7, No. 1*, 1361-1378.

Bjerring, J., Lakonishok, J., & Vermaelen, T. (1983). Stock Prices and Financial Analysts' Recommendations. *The Journal of Finance*, 38: 187-204.

Bloomberg Terminal. (2023, March 13). *Bloomberg Terminal*. Retrieved from https://www.bloomberg.com/professional/

Bonbright, J. C. (1927). The Problem of Judicial Valuation. *Columbia Law Review Vol. 27, No. 5*, 493-522.

Bonini, S., Bianchini, R., Salvi, A., & Zanetti, L. (2010). Target Price Accuracy in Equity Research. *Journal of Business Finance & Accounting*, 37: 1177-1217.

Buffet, W. (1997). *The Essays of Warren Buffet: Lessons for Corporate America.* New York: Columbia University Press.

Chan, L., Karceski, J., & Lakonishok, J. (2007). Analysts' Conflicts of Interest and Biases in Earnings Forecasts. *The Journal of Financial and Quantitative Analysis*, 42, no. 4: 893-913.

Chandra, P. (2014). *Corporate Valuation: A Guide for Analysts, Managers and Investors.* New Delhi: McGraw Hill Education .

Cheng, A., & McNamara, R. (2000). The Valuation Accuracy of the Price-Earnings and Price-Book Benchmark Valuation Methods. *Review of Quantitative Finance and Accounting*, 15, 349-370.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences (3rd ed.).* New Jersey: Lawrence Erlbaum Associates Publishers.

Comment, R. (2012). Revisiting the Illiquidity Discount for Private Companies: A New (and 'Skeptical') Restricted-Stock Study. *Journal of Applied Corporate Finance 24(1)*.

Coppola, M., McCreary, R., Marshella, T., Rowan, M., & Stumpp, P. (2000). *Putting EBITDA in Perspective.* New York: Moody's Investor Service.

Cox, E. (1962). Henry Moore and the "Statistical Complement of Pure Economics". *The American Statistician*, 16(2), 10-13.

Credit Suisse. (2016). *Managing the multiple: Weighing growth against profitability.* New York: Credit Suisse Securities LLC.

Damodaran, A. (2002). *Investment valuation: Tools and techniques for determining the value of any asset (2nd ed.).* New York: John Wiley & Sons, Inc.

Damodaran, A. (2006). *Damodaren on Valuation: Security Analysis for Investment and Corproate Finance (2nd ed.).* Haboken, New Jersey: John & Wiley & Sons, Inc. .

Damodaran, A. (2009). Valuing Young, Start-up and Growth Companies: Estimation Issues and Valuation Challenges. *SSRN Electronic Journal*.

Damodaran, A. (2012). *Corporate Valuation: Theory and Practice, 2nd Edition.* New York: John Wiley & Sons.

Damodaran, A. (2020). *Valuation: The Art and Science of Corporate Investment Decisions.* New Jersey: John Wiley & Sons.

Daoud, J. (2017). Multicollinearity and Regression Analysis. *Journal of Physics*.

Davis, N., McCarty, B., Shaw, K., & Sidani-Tabbaa, A. (1993). Transitions from Objectivism to Constructivism in Science Education. *International Journal of Science Education*, 15(6): 627-636.

Dittman, I., & Weiner, C. (2005). Selecting Comparables for the Valuation of European Firms. *Berlin: Humboldt-Universität zu Berlin*.

Dyck, A., & Zingales, L. (2005). Control Premiums and the Effectiveness of Corporate Governance Systems. *Journal of Applied Corporate Finance*, 51-72.

Easton, P. (2004). PE Ratios, PEG Ratios, and Estimating the Implied Expected Rate of Return on Equity Capital. *American Accounting Association*, 73-95.

Eberhart, A. (2004). Equity Valuation Using Multiples. *The Journal of Investing*, 13(2), 48-54.

Elton, E., Martin, J., & Seth-Grossman, S. (1986). Discrete Expectational Data and Portfolio Performance. *The Journal of Finance*, 3: 699-713.

Emory, S. J., Dengel, I. F., & Emory, J. J. (2002). Discounts for Lack of Marketability, Emory Pre-IPO Discount Studies 1980–2000, As Adjusted October 10, 2002. *Business Valuation Review*, 21(4), 190-191.

Fairfield, P., Sweeney, R., & Yohn, T. (1996). Accounting classification and the predictive content of earnings. *The Accounting Reveiw*.

Fairfield, P., Whisenant, J., & Yohn, T. (2003a). The differential persistence of accruals and cash flows for future operating income versus future profitability. *The Reveiw of Accounting Studies 8*, 221-243.

Fama, E. (1991). Effecient Capital Markets: II. *Wiley for the American Finance Association*, 1575-1617.

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

Fama, E., & French, K. (1997). Industry Costs of Equity. *Journal of Financial Economics*, 43(2), 153-193.

Fama, E., & French, K. (1998). Value versus Growth: The International Evidence. *The Journal of Finance*, 1975-1988.

Fama, E., & French, K. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, 116(1), 1-22.

Fama, E., & Miller, M. (1972). *The Theory of Finance.* New York: Holt, Rinehart, and Winston.

Feltham, G. A., & Ohlson, J. A. (1995). Valuation and Clean Surplus Accounting for Operating and Financial Activities*. *Contemporary Accounting Research Vol.1 No. 2*, 689-731.

Fernandez , P. (2001). Valuation Using Multiples: How Do Analysts Reach Their Conclusions? *Working Paper, IESE Business School, Madrid*, 1-13.

Fernandez, P. (2013). *Company Valuation Methods.* Barcelona: IESE Business School.

Flesher, D. L., & Previts, G. J. (2013). Donaldson Brown (1885-1965): The Power of an Individual and his ideas over time. *The Accounting Historians Journal*, 79-101.

Freeman, R. N., Ohlson, J. A., & Penman, S. H. (1982). Book Rate-of-Return and Predicting of Earnings Changes: An Empirical Investigation . *Journal of Accounting Research, Vol.20*, 639-653.

Goddard, E. C. (1924). Fair Value of Public Utilities. *Michigan Law Review Vol. 22, No. 8*, 777-797.

Graham, B., & Dodd, D. (1934). *Security Analysis.* New York: McGraw-Hill.

Gupta, V. (2018). Predicting Accuracy of Valuation Multiples Using Value Drivers: Evidence from Indian Listed Firms. *Theoretical Economics Letters, 8*, 755-772.

Haavelmo, T. (1944). The Probability Approach in Econometrics. *Econometrica*, 1-115.

Hammer, B., Matter, S., Scheizer, D., & Wunsche, N. (2023). Management buyouts in times of economic policy uncertainty. *Finance Research Letters*.

Harbula, P. (2009). Valuation Multiples: Accuracy and Drivers - Evidence from the European Stock Market. *Business Valuation Review*, 28(4), 186-200.

Healey, M. J. (1990). Measuring Importance. *Statistics in Medicine 9*, 633-637.

Heckman, J. (1992). Haavelmo and the Birth of Modern Econometrics: A Review of the History of Econometric Ideas by Mary Morgan. *Journal of Economic Literature*, 30(2), 876-886.

Helleiner, E. (1995). Explaining the Globalization of Financial Markets: Bringing States Back In. *Review of International Political Economy*, 315-341.

Henschke, S., & Homburg, C. (2009). Equity Valuation Using Multiples: Controlling for Differences Between Firms. *SSRN Electronic Journal*, DOI: 10.2139/ssrn.1270812.

Herrmann, V., & Richter, F. (2003). Pricing with Performance-Controlled Multiples. *Schmalenbach Business Review*, 55(3), 194-219.

Higgins, R. C. (2005). *Analysis of Financial Management. 11th edition.* Irvine: McGraw-Hill.

Hirshleifer, J. (1966). Investment Decision under Uncertainty: Applications of the State-Preference Approach. *Quarterly Journal of Economics, Vol. 80, No. 2*, 252-277.

Holm, A. (2013). *Philosophy of Science: An Introduction for Future Knowledge Workers.* Frederiksberg: Samfundslitteratur.

Hussain, T., & Chakraborty, L. (2010). Is the cost of capital an important determinant of market performance of private commercial banks in Bangladesh. *ASBBS Annual Conference* (pp. 389-399). Las Vegas : ASBBS.

Huysmans, J., Dejaeger, K., Mues, C., Vanthienen, J., & Baesens, B. (2011). An empirical evaluation of the comprehensibility of decision table, tree and rule based predictive models. *Decision Support Systems 51(1)*, 141-154.

Jensen, M., & Smith, C. (1984). *The Theory of Corporate Finance: A Historical Overview.* New York: McGraw-Hill.

K.L., W. (1996). Do Brokerage Analysts' Recommendations Have Investment Value? *The Journal of Finance*, 51, No. 1: 137-167.

Kakati, M. (2005). Stock Valuation Process: A Practitioners' view. *Finance India 19(2)*, 513-523.

Kaplan, S. N., & Ruback, R. S. (1995). The Valuation of Cash Flow Forecasts: An Empirical Analysis. *The Journal of Finance 50*, 1059–93.

Kasperzak, R., Janke, G., & Erkilet, G. (2021). How valuation approach choice affects financial analysts' target price accuracy. *Journal of Business Economics*, 92: 741-779.

Kim, M., & Ritter, J. (1999). Valuing IPOs. *Journal of Financial Economics*, 53(3), 409-437.

Kinserdal, F., Petersen, V., & Plenborg, T. (2017). *Financial Statement Analysis: Valuation, Credit Analysis, Performance Evaluation.* Bergen: Fagbokforlaget.

Kleinbaum, D. G., Kupper, L. L., & Muller, K. E. (1988). *Variable Reduction and Factor Analysis. Applied Regression Analysis and Other Multivariable Methods.* Boston: PWS Kent Publishing Co.

Koller, T., Goedhart, M., & Wessels, D. (2010). *Valuation: Measuring and Managing the Value of Companies.* New Jersey: John Wiley & Sons.

Kraus, A., & Litzenberger, R. (1973). A State-Preference Model of Optimal Financial Leverage. *Journal of Finance, Vol. 28, No. 4* , 911-922.

Lee, C., Ma, P., & Wang, C. (2015). *Search-based peer firms: Aggregating investor perceptions through internet co-searches.* Journal of Financial Economics: 116(2), 410-431.

Lie, E., & Lie, H. (2002). Multiples Used to Estimate Corporate Value. *Financial Analysts Journal*, 44-54.

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

Liu , J., Nissim, D., & Thomas, J. (2002). Equity Valuation Using Multiples. *Journal of Accounting Research*, 135-172.

Liu, P. (1990). Stock Price Reactions to The Wall Street Journal's Securities Recommendations. *The Journal of Financial and Quantitative Analysis*, 25, 3: 399-410.

Lloyd-Davies, P. (1978). Stock Prices and the Publication of Second-Hand Information. *The Journal of Business*, 51: 43-56.

Loughran, T., & Wellman, J. (2011). New Evidence on the Relaiton between the Enterprise Multiple and Average Stock Returns. *Journal of Financial and Quantitative Analysis, 46*, 1629-1650.

Luigi, P., & Sorin, V. (2009). A review of capital structure theories. *Annals of Faculty of Economics*, 315-320.

Luigi, P., & Sorin, V. (2009). A review of capital structure theories. *Annals of the University of Oradea : Economic Science. 3.* .

M., E.-S., Thorpe, R., & Lowe, A. (2002). *Management Research: an Introduction.* Thousand Oaks, CA: Sage Publications.

Maliel, B. G. (2003). The Efficient Market Hypothesis and Its Critics. *JOURNAL OF ECONOMIC PERSPECTIVES VOL. 17, NO. 1*, 59-82.

Marczyk, G., DeMatteo, D., & Festinger, D. (2005). *Essentials of Research Design and Methodology.* Hoboken, NJ: John Wiley & Sons Inc.

Markovitz, H. (1952). Portfolio Selection. *The Journal of Finance*, 7(1), 77-91.

Mauboussin, M. J., & Callahan, D. (2023). *Cost of Capital: A Practical Guide to Measuring Opportunity Cost.* New York: Morgani Stanley Investment Management, Couterpoint Global Insights.

Miller, M., & Modigliani, F. (1958). The Cost of Capital, Corporation Finance, and the Theory of Investmen. *The American Economic Review, Vol. 48*, 261-297.

Molnar, C., Casalicchio, G., & Bischl, B. (2020). Intepretable Machine Learning - A Brief History, State-of-the-Art and Challenges. *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, 417-431.

Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). *Introduction to Linear Regression Analysis. Vol. 821.* Hoboken: John Wiley & Sons.

Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica, Vol. 35*, 768-783.

Munk, C. (2021). *Financial Markets and Investments.* Copenhagen: Copenhagen Business School.

Nel, S. W. (2010). A South African Persepctive on the Multiples of Choice in the Valuation of Ordinary Shareholders' Equity: From Thoery to Practice. *Adrican Journal of Business Management, 4*, 930-941.

Nel, W. S. (2009). The Use of Multiples in the South African Equity Market: Is the Popularity of the Price Earnings Ratio Justifiable from a Sector Perspective? *Meditary Accountancy Research, 17*, 101-115.

Nissim, D. (2013). Relative Valuation of US Insurance Companies. *Review of Accounting Studies*, 18(2).

Nissmin, D., & Penman, S. (2001). Ratio analysis and equity valuation: From research to practice. *Review of Accounting Studies* , 109-154.

Officer, M. (2007). The Price of Corporate Liquidity: Acquisition Discounts for Unlisted Targets. *Journal of Financial Economics*, 83(3), 571-598.

Ohlson, J. A. (1995). Earnings, Book Values, and Dividends in Equity Valuation. *Contemporary Accounting Research*, 661-687.

Overgaard Knudsen, K., Kold, S., & Plenborg, T. (2017). Stick to the Fundamentals and Discover Your Peers. *Financial Analyst Journal 73(3)* , 85-105.

Pearl, J., & Rosenbaum, J. (2009). *Investment Banking: Valuation, Leveraged Buyouts and Mergers & Acquisitions.* New Jersey: John Wiley & Sons.

Penman, S. H. (2007). Financial statement analysis and security valuation (3rd ed). *The International Journal of Accounting, vol. 42, issue 4*, 450-453.

Penman, S., & Zhang, X. (2003). Modeling sustainable earnings and P/E ratios using financial statement information. *Working paper, Columbia University*.

Petersen, C., Plenborg, T., & Scholer, F. (2006). Issues in Valuation of Privately Held Firms. *The Journal of Private Equity*, 10(1), 33-48.

Pinto, J., Robinson, T., & Stowe, J. (2015). Equity Valuation: A survey of Professional Practice. *Working paper, CFA Institute*.

Plenborg, T., & Pimentel, R. C. (2016). Best Practices in Applying Multiples for Valuation Purposes. *The Journal of Private Equity, vol. 19, no. 3*, 55–64.

Poterba, J., & Summers, L. (1988). Mean Reversion in Stock Prices: Evidence and Implications. *National Bureau of Economic Research*.

Pratt, S., Reilly, R., & Schweihs, R. (2008). *Valuing a Business (5th ed.).* New York: McGraw-Hill.

Richardson, S., Sloan, M., & Tuna, I. (2005). Accrual reliability, earnings persistence and stock prices. *Journal of Accounting and Economics 39* , 437-485.

Robichek, A., & Myers, S. (1966). Problems in the Theory of Optimal Capital Structur. *Journal of Financial and Quantitative Analysis, Vol. 1*, 1-35.

Ross, A. (2007). *Modern Financial Management, 7th ed.* Irvine: McGraw-Hill.

Ross, M. (2017). Chapter 12 - Linear Regression. In M. Ross, *Introductory Statistics (Fourth edition)* (pp. 519-584). Academic Press.

Rossi, E., & Forte, G. (2016). *Assessing Relative Valuation in Equity Markets.* London: Palgrave Macmillan.

Rousseuuw, P. (1987). *Robust Regression and Outlier Detection.* Brussels: Bristol-Myers-Squibb.

Rubinstein, M. (1973). Corporate Financial Policy in Segmented Markets. *Journal of Financial and Quantitative Analysis, Vol. 8, No. 5*, 749-761.

S&P. (2023, March 16). *S&P Dow Jones Indices*. Retrieved from S&P Global: https://www.spglobal.com/spdji/en/indices/equity/sp-composite-1500/#overview

Schreiner, A., & Spremann, K. (2007). Multiples and Their Valuation Accuracy in European Equity Markets. *SSRN Electronic Journal*.

Scott Jr., J. H. (1976). A Theory of Optimal Capital Structure. *The Bell Journal of Economics , Spring, 1976, Vol. 7, No. 1*, 33-54.

Sharma, M., & Prashar, E. (2013). A Conceptual Framework for Relative Valuation. *The Journal of Private Equity, Vol. 16, No. 3*, 29-32.

Sharpe, W. (1964). Capital Asset PRices: A Theory of Market Equilibrium under Conditions of Risk. *Journal of Finance, Vol. 19*, 425-442.

Silber, W. (1991). Discounts on Restricted Stock: The Impact of Illiquidity on Stock Prices. *Fiancial Analysts*, 47(4), 60-64.

Smith, A. (1776). *An Inquiry into the Nature and Causes of the Wealth of Nations.* London: Methuen & Co.

Soffer, L. C., & Soffer, R. J. (2003). *Financial Statement Analysis: A Valuation Approach.* Pennsylvania State University: Prentice Hall.

Soliman, M. T. (2008). The Use of DuPont Analysis by Market Participants. *The Accounting Review, Vol. 83, No. 3*, 823-853.

Steiger, F. (2010). *The Validity of Company Valuation Using Discounted Cash Flow Methods.* Ithaca: Cornell University.

Stickel, S., & Verrecchia, R. (1994). Evidence That Trading Volume Sustains Stock Price Changes. *Finance Analyst Journal*, 50, no. 6: 57-67.

Stiglitz, J. (1969). A Reexamination of the Modigliani-Miller Theorem. *The American Economic Reviewv, Vol. 59, No. 5*, 784-793.

Stiglitz, J. (1972). Some Aspects of the Pure Theory of Corporate Finance: Bankruptcies and Take-Over. *The Bell Journal of Economics and Management Science, Vol. 3, No. 2*, 458-482.

Stiglitz, J. (1974). On the Irrelevance of Corporate Financial Policy. *The American Economic Review, Vol. 64, No. 6* , 851-866.

UBS, W. (2001). *Valuation Multiples: A Primer.* Zürich: UBS Warburg.

Ucak, A. (2015). Adam Smith: The Inspirer of Modern Growth Theories. *Procedia - Social and Behavioral Sciences 195* , 663-672.

Woolridge, J. (2020). *Introductory Economics: A Modern Approach.* Boston: Cengage Learning, Inc. .

Yee, K. (2004). Forward Versus Trailing Earnings in Equity Valuation. *Review of Accounting Studies 9*, 301–329.

Yee, K. (2008). A Bayesian Framework for Combining Valuation Estimates. *Review of Quantitative Finance and Accounting*, 30(3), 339-354.

Yin, Y., Peasnell, K., & Hunt, H. I. (2018). How do sell-side analysts obtain price-earnings multiples to value firms? . *Accounting and Business Research, 48(1)*, 108-135.

Yoo, K. (2006). The valuation accuracy of equity valuation using a combination of multiples. *Review of Accounting and Finance*, 50(2), 108-123.

Young, S., & Zeng, Y. (2015). Accounting Comparability and the Accuracy of Peer-based Valuation Models. *SSRN Electronic Journal*, https://doi.org/10.2139/ssrn.3433586.

Yousaf, M., & Dey, S. K. (2022). Best proxy to determine firm performance using financial ratios: A CHAID approach. *Review of Economic Perspectives, Vol. 22*, 219-239.

Zarowin, P. (1990). What Determines Earnings-Price Ratios: Revisted. *Journal of Accounting, Auditing, and Finance*, 439-457.

Zdaniuk, B. (2014). Ordinary Least-Squares (OLS) Model. In A. Michalos, *Encyclopedia of Quality of Life and Well-Being Research* (pp. 4515-4517). Dordrecht: Springer.

Appendix

Appendix 1 – Literature review summary table

Literature review summary table

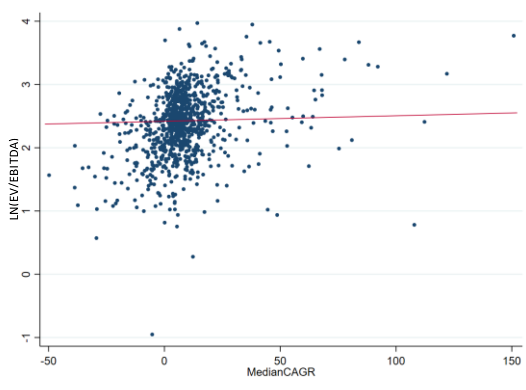
1. Valuation multiple Fundamentals	**Valuation multiples vs DCF:** - Kaplan & Ruback (1995) - Liu et al. (2002) - Damodaran, A. (2009) - Fernandez, P. (2013)	**Forward vs Trailing multiples:** - Kim & Ritter (1999) - Liu et al. (2002) - Lie & Lie (2002) - Schreiner & Spremann (2007) - Plenborg & Pimentel (2016)	**Aggregating valuation multiples:** - Baker & Ruback (1999) - Liu et al. (2002) - Herrmann & Richter (2003) - Schreiner & Spremann (2007) - Plenborg and Pimentel (2016)	**Illiquid discount:** - Reilly & Scheihs (2008) - Silber (1991) - Emory et al (2002) - Officer (2007) - Bernström (2014) - Dryck & Zingales (2005)
2. Differnt valuation multiples	**Earnings multiples:** - Abukari et al. (2000) - Nissimg & Thomas (2002) - Lie & Lie (2002)	**Cashflow-based multiple:** - Koller et al. (2010) - Baker & Ruback (1999)	**EV/EBITDA multiple advocaters:** - Rosse & Forte (2016) - Rosenbaum & Perl (2009) - Credit Suisse (2016) - UBS (2001) - Kim & Ritter (1999) - Lee et al. (2015) - Plenborg & Pimentel (2016)	**Industry specific multiples:** - Lee et al. (2015) - Baker & Ruback (1999) - Fernandez (2001) - Damodaran (2012) - Rossi & Forte (2016 - Harbula (2009)
3. Econometrics & Valuation multiples	**Economy & Statistics:** - Cox (1962) - Heckman (1992) - Haavelmo (1994)	**Equity market & statistics:** - Markowitz (1952) - Munk (2021) - Sharpe (1964) - Fama & French (1993) - Fama & French (2015)	**Non-linearity in financial markets** - Basu (1977) - Ang & Chen (2007) - Fama & French (1993) - Damodaran (2012)	**Intertemporal differences**: - Damodaran (2012) - Harbula (2009)
4. Underlying value drivers of firm value	**Growth as a value driver** - Adam Smith (1976) - UBS (2001) - Damodaran (2006) - Yin et al. (2018) - Gupta (2018) - Credit Suisse (2016)	**Profitability as a value driver:** - Bernard (1994) - Nel (2009;2010) - Koller et al. (2010) - Credit Suisse (2016) - Gupta (2018) - UBS (2001)	**Risk as a value drivers:** - Baumol & Malkiel (1967) - Rubenstein (1973) - Stiglitz (1972) - Baxter (1967) - Thomas (1972) - Kraus & Litzenberger (1973) - Robichek & Myers (1966) - Mauboussin & Callahan (2023) - Berk & Demarzo (2020) - Damodaran (2012)	
5. Proxy for growth as a fundamental value driver	**Sales growth as a proxy:** - Kakita (2005)	**EBITDA growth as a proxy:** - Zarowin (1990) - Damodaran (2006) - Achleitner (2011) - Hammer et al. (2023) - Damodaran (2012)		
6. Proxy for profitability as a fundamental value driver	**Net profit as a proxy:** - Berk & DeMarzo (2019)	**ROE as a proxy:** - Feeman et al. (1982) - Ohlson (1995) - Feltham & Ohlson (1995)	**RNOA as a proxy:** - Nissim & Penman (2001) - Fairfield et al. (1996) - Penman & Zhang (2003) - Fairfield et al (2003a) - Sloan & Tuna (2005) - Yousaf & Dey (2022)	**ROIC as a proxy:** - Koller et al. (2010) - Kinserdal et al. (2017)
7. Proxy for risk as a fundamental value drivers	**WACC as a proxy:** - Abid & Mseddi (2010) - Mauboussin & Callahan (2023) - Higgins (2005) - Hussain & Chakraborty (2010) - Ross (2007) - Lougran & Wellman (2011)			
8. Peer group selection	**Peer grouping through industry affiliation:** - Berk & DeMarzo (2019) - Alford (1992) - Fama & French (1997) - Eberhart (2004) - Bhojraj et al (2003) - Young & Zeng (2015)	**Peer grouping through dynamics in underlying value drivers:** - Berk & DeMarzo (2019) - Cheng & McNamara (2000) - Bhojraj & Lee (2002) - Dittman & Weiner (2005) - Henscke & Homburg (2009) - Cheng & McNamara (2000) - Yee (2004) - Yoo (2006)		
9. Sum of Absolute Rank Differences (SARD)* as a peer group selection approach	**SARD*:** - Knudsen, Kold, and Plenborg (2017)	**Opposers to SARD:** - Rossi & Forte (2016)		

*The mechanics of the SARD approach is elborated upon in section 2.3 and 3.4

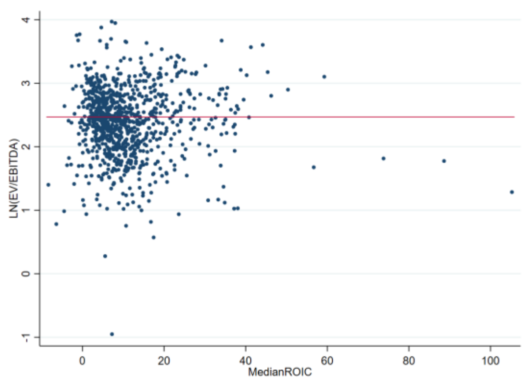
Appendix 2 – Residual scatterplots

Appendix 2.1-2.3 – Residual scatter plots

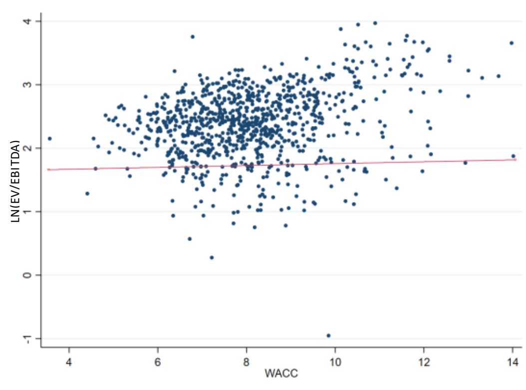
Residuals scatterplot - SLR, EBITDA CAGR (growth)



Residuals scatterplot - SLR, median ROIC (profitability)



Residuals scatterplot - SLR, WACC LFY (risk)



Appendix 3 – Summary statistics and correlation matrices

Market level

Market

Sample: Trimmed census S&P 1500

Number of firms: 929

A) Summary statistics / **B) Correlation matrix**

Variable	Mean	Median	Std. dev.	10P	90P	LN(EVE)	CAGR	ROIC	WACC
EVE	12.4583	11.2109	6.8386	5.4897	20.8065	1.0000			
CAGR	0.0902	0.0694	0.1706	-0.0573	0.2358	0.3562****	1.0000		
ROIC	0.1115	0.0875	0.1011	0.0226	0.2321	0.0300	-0.1974****	1.0000	
WACC	0.0810	0.0795	0.0155	0.0630	0.1020	0.3672***	0.2292****	0.3196****	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Sector level

Communication Services Sector

A) Summary statistics / **B) Correlation matrix**

Variable	Mean	Median	Std. dev.	10P	90P	LN(EVE)	CAGR	ROIC	WACC
EVE	11.3743	10.3753	4.5966	6.9496	17.2911	1.0000			
CAGR	0.1073	0.0765	0.1405	-0.0106	0.2107	0.4173****	1.0000		
ROIC	0.0936	0.0789	0.0685	0.0181	0.2042	0.1619	2.0000	1.0000	
WACC	0.0754	0.0703	0.0165	0.0575	0.1024	0.5139****	0.509****	0.4638****	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Consumer Discretionary Sector

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	9.8644	8.7210	5.3780	4.8445	16.1001
CAGR	0.0175	0.0231	0.1298	-0.1410	0.1598
ROIC	0.1333	0.1188	0.1029	0.0351	0.2377
WACC	0.0814	0.0806	0.0141	0.0651	0.0979

B) Correlation matrix

LN(EVE)	CAGR	ROIC	WACC
1.0000			
0.2786****	1.0000		
0.2052***	-0.0674	1.0000	
0.2712****	0.2491****	0.2877****	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Consumer Staples Sector

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	12.7849	11.9181	5.2298	7.0382	20.5770
CAGR	0.0795	0.0605	0.1142	-0.0200	0.2095
ROIC	0.1083	0.0911	0.0832	0.0281	0.2211
WACC	0.0707	0.0698	0.0101	0.0586	0.0843

B) Correlation matrix

LN(EVE)	CAGR	ROIC	WACC
1.0000			
0.1153	1.0000		
0.2339*	-0.2319*	1.0000	
0.2073*	0.1584	0.2539**	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Energy Sector

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	4.3538	3.9421	2.5948	2.1266	9.3720
CAGR	0.1194	0.0885	0.2493	-0.0702	0.2692
ROIC	0.0439	0.0528	0.0637	-0.0449	0.1447
WACC	0.0767	0.0771	0.0083	0.0669	0.0860

B) Correlation matrix

LN(EVE)	CAGR	ROIC	WACC
1.0000			
0.0848	1.0000		
-0.132	-0.6055****	1.0000	
-0.6696****	0.0923	0.0082	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Financials Sector

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	11.0577	8.4389	6.7168	5.2610	20.8411
CAGR	0.0007	0.0508	0.1479	-0.2607	0.1262
ROIC	0.1281	0.1070	0.0858	0.0410	0.2482
WACC	0.0833	0.0854	0.0177	0.0531	0.1086

B) Correlation matrix

LN(EVE)	CAGR	ROIC	WACC
1.0000			
0.4645****	1.0000		
0.2694*	0.0466	1.0000	
0.2195	0.3252**	0.4765****	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Health Care Sector

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	15.8182	13.0741	8.8486	8.0554	24.8981
CAGR	0.1557	0.1119	0.2001	-0.0108	0.4141
ROIC	0.1100	0.0855	0.0904	0.0196	0.2415
WACC	0.0807	0.0774	0.0161	0.0638	0.1020

B) Correlation matrix

LN(EVE)	CAGR	ROIC	WACC
1.0000			
0.1622*	1.0000		
-0.0912	-0.1287	1.0000	
0.6864****	0.2304***	0.0861	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Industrials Sector

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	11.5655	11.0605	5.7240	5.5581	16.9772
CAGR	0.0837	0.0655	0.1373	-0.0089	0.1760
ROIC	0.1202	0.0972	0.0856	0.0357	0.2556
WACC	0.0839	0.0843	0.0106	0.0709	0.0961

B) Correlation matrix

LN(EVE)	CAGR	ROIC	WACC
1.0000			
0.2726****	1.0000		
0.0221	-0.4494****	1.0000	
0.2946****	0.1486*	0.299****	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Information Technology Sector

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	14.6276	11.7060	8.3255	5.9001	26.6241
CAGR	0.1632	0.1102	0.2324	-0.0277	0.4460
ROIC	0.1619	0.1324	0.1458	0.0321	0.3351
WACC	0.0939	0.0927	0.0175	0.0739	0.1175

B) Correlation matrix

LN(EVE)	CAGR	ROIC	WACC
1.0000			
0.4807****	1.0000		
-0.0427	-0.2548****	1.0000	
0.4423****	0.268****	0.1279	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Materials Sector

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	9.3026	8.0531	5.1003	3.8751	15.5094
CAGR	0.0501	0.0735	0.1655	-0.1194	0.1683
ROIC	0.1127	0.1019	0.0882	0.0362	0.1691
WACC	0.0832	0.0818	0.0129	0.0678	0.1023

B) Correlation matrix

LN(EVE)	CAGR	ROIC	WACC
1.0000			
0.4187****	1.0000		
-0.4325****	-0.5914****	1.0000	
-0.0729	0.1605	0.2565*	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Real Estate Sector

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	16.0206	15.5866	4.8914	10.8398	21.4077
CAGR	0.0703	0.0564	0.1118	0.0023	0.2030
ROIC	0.0411	0.0383	0.0323	0.0086	0.0759
WACC	0.0744	0.0747	0.0090	0.0640	0.0846

B) Correlation matrix

LN(EVE)	CAGR	ROIC	WACC
1.0000			
0.139	1.0000		
-0.1815*	-0.3162****	1.0000	
0.2154**	-0.186*	0.3966****	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Utilities Sector									

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	11.6501	11.2874	2.8775	8.8054	15.0110
CAGR	0.0895	0.0766	0.0642	0.0431	0.1747
ROIC	0.0502	0.0498	0.0230	0.0303	0.0638
WACC	0.0628	0.0625	0.0055	0.0573	0.0701

B) Correlation matrix

	LN(EVE)	CAGR	ROIC	WACC
EVE	1.0000			
CAGR	0.4195****	1.0000		
ROIC	-0.4598****	-0.5946****	1.0000	
WACC	0.2597*	0.2169	0.2496*	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Industry level

Automobiles & Components

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	8.3774	6.8207	5.3277	3.2249	13.6608
CAGR	0.1080	0.1225	0.1480	-0.1274	0.3201
ROIC	0.0784	0.0652	0.0623	0.0095	0.1684
WACC	0.0812	0.0789	0.0178	0.0633	0.1021

B) Correlation matrix

	LN(EVE)	CAGR	ROIC	WACC
EVE	1.0000			
CAGR	0.6517****	1.0000		
ROIC	0.3644	-0.215	1.0000	
WACC	0.7121****	0.5115**	0.5276***	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Capital Goods

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	11.6842	11.3404	5.0543	5.6332	17.0094
CAGR	0.1038	0.0742	0.1518	-0.0062	0.2193
ROIC	0.1199	0.1008	0.0821	0.0385	0.2407
WACC	0.0857	0.0846	0.0091	0.0743	0.0969

B) Correlation matrix

	LN(EVE)	CAGR	ROIC	WACC
EVE	1.0000			
CAGR	0.3451****	1.0000		
ROIC	-0.0382	-0.4923****	1.0000	
WACC	0.2975****	0.1588*	0.1779*	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Commercial & Professional Services

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	12.5113	11.1693	8.1327	6.1113	17.1698
CAGR	0.0650	0.0606	0.0659	0.0049	0.1291
ROIC	0.1057	0.0812	0.0867	0.0313	0.2499
WACC	0.0799	0.0786	0.0125	0.0649	0.0958

B) Correlation matrix

LN(EVE)	CAGR	ROIC	WACC
1.0000			
-0.03	1.0000		
0.1126	-0.3571**	1.0000	
0.3417*	-0.1471	0.4589****	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Consumer Durables & Apparel

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	8.9581	7.6697	4.2174	5.5243	16.1001
CAGR	0.0016	0.0168	0.1122	-0.1510	0.1191
ROIC	0.1610	0.1354	0.1212	0.0491	0.2455
WACC	0.0872	0.0859	0.0131	0.0750	0.1026

B) Correlation matrix

LN(EVE)	CAGR	ROIC	WACC
1.0000			
0.3245*	1.0000		
-0.048	0.1158	1.0000	
0.1095	0.2948*	0.4705****	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Consumer Services

A) Summary statistics | **B) Correlation matrix**

Variable	Mean	Median	Std. dev.	10P	90P	LN(EVE)	CAGR	ROIC	WACC
EVE	12.7124	11.1467	6.6179	6.6172	22.2450	1.0000			
CAGR	0.0621	0.0653	0.1326	-0.0886	0.1669	0.1996	1.0000		
ROIC	0.1226	0.1028	0.1117	0.0241	0.2404	0.4302****	-0.0467	1.0000	
WACC	0.0767	0.0772	0.0127	0.0618	0.0952	0.5778****	0.2235	0.0786	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Diversified Financials

A) Summary statistics | **B) Correlation matrix**

Variable	Mean	Median	Std. dev.	10P	90P	LN(EVE)	CAGR	ROIC	WACC
EVE	11.0104	8.5239	6.6368	5.3045	20.1521	1.0000			
CAGR	-0.0041	0.0507	0.1492	-0.2435	0.1237	0.4565****	1.0000		
ROIC	0.1337	0.1101	0.0919	0.0424	0.2579	0.2446	-0.0377	1.0000	
WACC	0.0833	0.0852	0.0174	0.0591	0.1064	0.2195	0.3184**	0.4392****	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Energy

A) Summary statistics | **B) Correlation matrix**

Variable	Mean	Median	Std. dev.	10P	90P	LN(EVE)	CAGR	ROIC	WACC
EVE	4.3538	3.9421	2.5948	2.1266	9.3720	1.0000			
CAGR	0.1194	0.0885	0.2493	-0.0702	0.2692	0.0848	1.0000		
ROIC	0.0439	0.0528	0.0637	-0.0449	0.1447	-0.132	-0.6055****	1.0000	
WACC	0.0767	0.0771	0.0083	0.0669	0.0860	-0.6696****	0.0923	0.0082	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Food & Staples Retailing

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	9.9134	8.2246	3.8014	6.3434	16.3837
CAGR	0.0375	0.0460	0.0978	-0.0713	0.1640
ROIC	0.0814	0.0850	0.0464	0.0281	0.1306
WACC	0.0734	0.0738	0.0084	0.0633	0.0789

B) Correlation matrix

LN(EVE)	CAGR	ROIC	WACC
1.0000			
0.0132	1.0000		
0.2812	-0.1217	1.0000	
0.3121	0.4048	0.6054**	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Food, Beverage & Tobacco

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	13.2421	12.6716	5.3846	7.0382	23.2407
CAGR	0.0948	0.0693	0.1143	0.0085	0.2296
ROIC	0.1091	0.0893	0.0943	0.0220	0.2323
WACC	0.0681	0.0680	0.0098	0.0558	0.0814

B) Correlation matrix

LN(EVE)	CAGR	ROIC	WACC
1.0000			
0.1039	1.0000		
0.0899	-0.2855*	1.0000	
0.284*	0.1264	0.3191**	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Health Care Equipment & Services

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	16.3829	13.5839	9.4875	8.0554	27.9112
CAGR	0.1506	0.1188	0.1855	-0.0240	0.3411
ROIC	0.0974	0.0741	0.0874	0.0196	0.2082
WACC	0.0805	0.0781	0.0160	0.0641	0.1009

B) Correlation matrix

LN(EVE)	CAGR	ROIC	WACC
1.0000			
0.3461****	1.0000		
-0.0569	-0.0784	1.0000	
0.6716****	0.4409****	0.2256*	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Household & Personal Products

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	14.2847	14.0948	5.2493	7.8258	22.1894
CAGR	0.0754	0.0533	0.1264	-0.0491	0.1446
ROIC	0.1329	0.1286	0.0719	0.0527	0.2472
WACC	0.0757	0.0703	0.0107	0.0658	0.0922

B) Correlation matrix

LN(EVE)	CAGR	ROIC	WACC
1.0000			
0.0253	1.0000		
0.6334***	-0.2869	1.0000	
0.05	0.3481	-0.1021	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Materials

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	9.3026	8.0531	5.1003	3.8751	15.5094
CAGR	0.0501	0.0735	0.1655	-0.1194	0.1683
ROIC	0.1127	0.1019	0.0882	0.0362	0.1691
WACC	0.0832	0.0818	0.0129	0.0678	0.1023

B) Correlation matrix

LN(EVE)	CAGR	ROIC	WACC
1.0000			
0.4187****	1.0000		
-0.4325****	-0.5914****	1.0000	
-0.0729	0.1605	0.2565*	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Media & Entertainment

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	11.6300	10.8151	4.9081	6.9496	18.4616
CAGR	0.1211	0.0778	0.1546	-0.0028	0.2181
ROIC	0.1069	0.0939	0.0591	0.0343	0.2042
WACC	0.0783	0.0756	0.0168	0.0588	0.1038

B) Correlation matrix

LN(EVE)	CAGR	ROIC	WACC
1.0000			
0.3689**	1.0000		
0.1648	-0.0722	1.0000	
0.4653****	0.4463***	0.459***	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Pharmaceuticals, Biotechnology & Life Sciences

A) Summary statistics / **B) Correlation matrix**

Variable	Mean	Median	Std. dev.	10P	90P		LN(EVE)	CAGR	ROIC	WACC
EVE	15.0953	12.9815	7.7380	7.6880	24.5078		1.0000			
CAGR	0.1622	0.0924	0.2220	-0.0046	0.4679		-0.0996	1.0000		
ROIC	0.1364	0.1172	0.0988	0.0196	0.2744		-0.0358	-0.2209	1.0000	
WACC	0.0816	0.0771	0.0164	0.0623	0.1039		0.7317****	-0.0652	-0.0482	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Real Estate

A) Summary statistics / **B) Correlation matrix**

Variable	Mean	Median	Std. dev.	10P	90P		LN(EVE)	CAGR	ROIC	WACC
EVE	16.0206	15.5866	4.8914	10.8398	21.4077		1.0000			
CAGR	0.0703	0.0564	0.1118	0.0023	0.2030		0.139	1.0000		
ROIC	0.0411	0.0383	0.0323	0.0086	0.0759		-0.1815*	-0.3162****	1.0000	
WACC	0.0744	0.0747	0.0090	0.0640	0.0846		0.2154**	-0.186*	0.3966****	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Retailing

A) Summary statistics / **B) Correlation matrix**

Variable	Mean	Median	Std. dev.	10P	90P		LN(EVE)	CAGR	ROIC	WACC
EVE	8.9606	8.3085	4.3402	4.3695	15.5038		1.0000			
CAGR	-0.0377	-0.0231	0.1025	-0.1578	0.0901		0.0793	1.0000		
ROIC	0.1457	0.1250	0.0903	0.0449	0.2907		0.3087**	0.0623	1.0000	
WACC	0.0807	0.0821	0.0131	0.0633	0.0961		0.2036	0.3278***	0.193	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Semiconductors & Semiconductor Equipment

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	12.6405	10.5113	7.0453	5.1439	23.1829
CAGR	0.1070	0.0790	0.2279	-0.1309	0.5933
ROIC	0.1787	0.1440	0.1167	0.0632	0.3551
WACC	0.1035	0.1047	0.0160	0.0821	0.1237

B) Correlation matrix

LN(EVE)	CAGR	ROIC	WACC
1.0000			
0.4903****	1.0000		
0.1981	-0.2479	1.0000	
0.2211	0.0898	0.4067***	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Software & Services

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	17.8877	16.6727	9.0893	8.6605	31.0354
CAGR	0.2269	0.1481	0.2565	0.0459	0.4998
ROIC	0.1667	0.1313	0.1643	0.0153	0.3314
WACC	0.0926	0.0900	0.0192	0.0711	0.1175

B) Correlation matrix

LN(EVE)	CAGR	ROIC	WACC
1.0000			
0.5722****	1.0000		
-0.2051	-0.2929**	1.0000	
0.6814****	0.4113****	-0.0208	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Technology Hardware & Equipment

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	12.2046	10.1637	7.0060	5.5548	22.8291
CAGR	0.1302	0.0883	0.1859	-0.0258	0.3231
ROIC	0.1420	0.1256	0.1435	0.0324	0.2404
WACC	0.0877	0.0848	0.0128	0.0718	0.1047

B) Correlation matrix

LN(EVE)	CAGR	ROIC	WACC
1.0000			
0.2018	1.0000		
-0.0312	-0.2373	1.0000	
0.5599****	0.3966****	0.1206	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Telecommunication Services

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	10.6073	9.4717	3.6141	6.3822	15.6184
CAGR	0.0659	0.0675	0.0770	-0.0326	0.1714
ROIC	0.0537	0.0242	0.0819	0.0010	0.1779
WACC	0.0670	0.0662	0.0129	0.0487	0.0855

B) Correlation matrix

LN(EVE)	CAGR	ROIC	WACC
1.0000			
0.7772****	1.0000		
0.0984	0.248	1.0000	
0.7528***	0.8672****	0.2884	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Transportation

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	9.0240	10.4767	3.4887	4.3121	13.5114
CAGR	-0.0100	0.0191	0.0922	-0.1678	0.0838
ROIC	0.1502	0.1197	0.1023	0.0279	0.3508
WACC	0.0801	0.0798	0.0125	0.0684	0.0955

B) Correlation matrix

LN(EVE)	CAGR	ROIC	WACC
1.0000			
-0.2232	1.0000		
0.3657	-0.4253*	1.0000	
0.1614	0.0753	0.6201****	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Utilities

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	11.6501	11.2874	2.8775	8.8054	15.0110
CAGR	0.0895	0.0766	0.0642	0.0431	0.1747
ROIC	0.0502	0.0498	0.0230	0.0303	0.0638
WACC	0.0628	0.0625	0.0055	0.0573	0.0701

B) Correlation matrix

LN(EVE)	CAGR	ROIC	WACC
1.0000			
0.4195****	1.0000		
-0.4598****	-0.5946****	1.0000	
0.2597*	0.2169	0.2496*	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Appendix 4 – Summary statistics and correlation matrices (SARD)

Automobiles & Components

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	11.6069	11.5213	4.5297	6.1505	15.6233
CAGR	0.1631	0.1459	0.0649	0.0861	0.2787
ROIC	0.0201	0.0218	0.0107	0.0086	0.0318
WACC	0.0641	0.0649	0.0056	0.0546	0.0700

B) Correlation matrix

LN(EVE)	CAGR	ROIC	WACC
1.0000			
0.4019*	1.0000		
0.0237	-0.4854**	1.0000	
0.0312	0.0777	-0.2458	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Capital Goods

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	10.6147	9.7029	5.8834	3.4778	19.7186
CAGR	-0.0400	-0.0072	0.0847	-0.1849	0.0418
ROIC	0.2275	0.1923	0.1113	0.1286	0.3698
WACC	0.0868	0.0845	0.0099	0.0755	0.0988

B) Correlation matrix

LN(EVE)	CAGR	ROIC	WACC
1.0000			
0.3925****	1.0000		
0.0723	-0.0545	1.0000	
-0.0205	-0.0497	0.1586*	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Commercial & Professional Services

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	11.4623	10.5975	4.5344	7.1088	17.5483
CAGR	0.0867	0.0862	0.0220	0.0645	0.1130
ROIC	0.0869	0.0876	0.0167	0.0627	0.1070
WACC	0.0699	0.0693	0.0041	0.0651	0.0768

B) Correlation matrix

	LN(EVE)	CAGR	ROIC	WACC
EVE	1.0000			
CAGR	-0.3932**	1.0000		
ROIC	-0.0529	0.1174	1.0000	
WACC	-0.1098	0.1677	0.2235	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Consumer Durables & Apparel

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	16.9835	16.3242	6.6330	9.0704	27.9409
CAGR	0.0916	0.0883	0.0252	0.0663	0.1262
ROIC	0.2503	0.2366	0.0744	0.1660	0.3519
WACC	0.0943	0.0951	0.0065	0.0858	0.1041

B) Correlation matrix

	LN(EVE)	CAGR	ROIC	WACC
EVE	1.0000			
CAGR	0.1846	1.0000		
ROIC	0.2458	-0.0641	1.0000	
WACC	0.5419****	0.1527	0.419****	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Consumer Services

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	10.4875	10.3463	3.3613	6.6020	13.1355
CAGR	0.0416	0.0442	0.0179	0.0167	0.0637
ROIC	0.0748	0.0748	0.0176	0.0502	0.1007
WACC	0.0608	0.0620	0.0055	0.0557	0.0658

B) Correlation matrix

	LN(EVE)	CAGR	ROIC	WACC
EVE	1.0000			
CAGR	0.2737*	1.0000		
ROIC	0.0743	-0.4314****	1.0000	
WACC	0.2041	0.3128*	-0.0691	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Diversified Financials

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	10.1389	9.9993	3.8063	5.0542	15.3139
CAGR	-0.0191	0.0071	0.0741	-0.1247	0.0473
ROIC	0.0431	0.0418	0.0198	0.0191	0.0673
WACC	0.0588	0.0600	0.0067	0.0494	0.0667

B) Correlation matrix

LN(EVE)	CAGR	ROIC	WACC
1.0000			
0.4623****	1.0000		
0.0649	-0.0024	1.0000	
0.0888	-0.1834	0	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Energy

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	12.9291	11.7317	6.9466	4.1097	22.1343
CAGR	0.3777	0.2652	0.2591	0.1932	0.6810
ROIC	0.0035	0.0110	0.0287	-0.0440	0.0361
WACC	0.0819	0.0811	0.0045	0.0768	0.0884

B) Correlation matrix

LN(EVE)	CAGR	ROIC	WACC
1.0000			
-0.1089	1.0000		
0.4633***	-0.4938***	1.0000	
0.1783	0.1067	-0.0172	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Food & Staples Retailing

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	10.6629	9.3955	5.3971	5.7472	19.4630
CAGR	0.0556	0.0600	0.0103	0.0383	0.0648
ROIC	0.0873	0.0859	0.0166	0.0639	0.1057
WACC	0.0746	0.0747	0.0030	0.0695	0.0776

B) Correlation matrix

LN(EVE)	CAGR	ROIC	WACC
1.0000			
0.3269	1.0000		
-0.2084	-0.2893	1.0000	
-0.575**	-0.2893	-0.0491	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Food, Beverage & Tobacco

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	8.7518	8.5997	2.4697	5.4668	12.0283
CAGR	0.0089	0.0254	0.0384	-0.0431	0.0390
ROIC	0.0917	0.0900	0.0196	0.0643	0.1169
WACC	0.0612	0.0627	0.0069	0.0535	0.0687

B) Correlation matrix

	LN(EVE)	CAGR	ROIC	WACC
EVE	1.0000			
CAGR	0.3532**	1.0000		
ROIC	-0.2502	-0.2006	1.0000	
WACC	-0.1216	0.0165	0.3135*	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Health Care Equipment & Services

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	14.9341	13.3308	8.9052	5.5121	25.4411
CAGR	0.3479	0.2549	0.2659	0.1301	0.6803
ROIC	0.0198	0.0235	0.0299	-0.0203	0.0545
WACC	0.0867	0.0855	0.0111	0.0765	0.0987

B) Correlation matrix

	LN(EVE)	CAGR	ROIC	WACC
EVE	1.0000			
CAGR	0.1808	1.0000		
ROIC	0.2935***	-0.3032****	1.0000	
WACC	0.3721****	0.3498****	0.1415	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Household & Personal Products

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	9.9550	9.6550	2.3564	6.7908	13.0318
CAGR	0.1538	0.1588	0.0358	0.0941	0.1936
ROIC	0.0338	0.0344	0.0095	0.0220	0.0446
WACC	0.0606	0.0604	0.0034	0.0563	0.0657

B) Correlation matrix

	LN(EVE)	CAGR	ROIC	WACC
EVE	1.0000			
CAGR	-0.0904	1.0000		
ROIC	0.1715	-0.1672	1.0000	
WACC	0.0171	0.073	0.4083	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Materials

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	9.6877	9.3394	3.9051	4.5855	15.0779
CAGR	-0.0166	-0.0007	0.0709	-0.1219	0.0431
ROIC	0.0695	0.0681	0.0174	0.0447	0.0900
WACC	0.0674	0.0677	0.0052	0.0629	0.0743

B) Correlation matrix

	LN(EVE)	CAGR	ROIC	WACC
EVE	1.0000			
CAGR	0.3596****	1.0000		
ROIC	-0.0642	0.1877	1.0000	
WACC	-0.001	-0.1882	-0.1085	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Media & Entertainment

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	20.4269	20.1754	9.9503	7.0287	33.4462
CAGR	0.1903	0.1595	0.0963	0.1159	0.2836
ROIC	0.2504	0.2262	0.1118	0.1568	0.3382
WACC	0.1106	0.1058	0.0124	0.0973	0.1297

B) Correlation matrix

	LN(EVE)	CAGR	ROIC	WACC
EVE	1.0000			
CAGR	0.3087*	1.0000		
ROIC	-0.1868	-0.0825	1.0000	
WACC	-0.0813	0.1323	-0.0831	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Pharmaceuticals, Biotechnology & Life Sciences

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	13.0536	11.6287	5.4101	7.0382	21.1380
CAGR	0.1321	0.1196	0.0622	0.0788	0.1749
ROIC	0.1909	0.1871	0.0692	0.1082	0.2744
WACC	0.0779	0.0772	0.0061	0.0711	0.0854

B) Correlation matrix

	LN(EVE)	CAGR	ROIC	WACC
EVE	1.0000			
CAGR	0.0208	1.0000		
ROIC	0.2548*	-0.0451	1.0000	
WACC	0.2242	0.0041	0.0535	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Real Estate

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	11.7247	11.1778	5.9526	5.4886	19.1133
CAGR	0.2407	0.1789	0.1798	0.1028	0.5274
ROIC	0.0234	0.0254	0.0267	-0.0090	0.0547
WACC	0.0699	0.0700	0.0073	0.0606	0.0787

B) Correlation matrix

LN(EVE)	CAGR	ROIC	WACC
1.0000			
0.0214	1.0000		
0.1522	-0.1842*	1.0000	
-0.0571	0.1585	-0.2629***	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Retailing

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	10.1724	9.3943	5.6263	3.8183	18.6683
CAGR	-0.0485	-0.0255	0.0832	-0.1728	0.0394
ROIC	0.2046	0.1882	0.0646	0.1373	0.3089
WACC	0.0929	0.0930	0.0061	0.0856	0.0996

B) Correlation matrix

LN(EVE)	CAGR	ROIC	WACC
1.0000			
0.4166****	1.0000		
0.0863	-0.1251	1.0000	
-0.0693	0.0631	0.2934**	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Semiconductors & Semiconductor Equipment

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	10.8921	9.4331	6.4460	3.5822	21.7883
CAGR	-0.0622	-0.0272	0.1119	-0.2459	0.0393
ROIC	0.2888	0.2594	0.0889	0.1836	0.3816
WACC	0.1003	0.0978	0.0093	0.0898	0.1170

B) Correlation matrix

LN(EVE)	CAGR	ROIC	WACC
1.0000			
0.5797****	1.0000		
0.0077	-0.3508**	1.0000	
-0.0014	-0.1578	0.1645	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Software & Services

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	17.0754	16.0665	8.0244	7.6697	29.1066
CAGR	0.0942	0.0885	0.0345	0.0528	0.1483
ROIC	0.2515	0.2311	0.0985	0.1578	0.3519
WACC	0.0993	0.0965	0.0110	0.0875	0.1111

B) Correlation matrix

LN(EVE)	CAGR	ROIC	WACC
1.0000			
0.0432	1.0000		
-0.0848	0.1352	1.0000	
0.0427	0.2689**	0.1536	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Technology Hardware & Equipment

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	18.6460	16.4692	10.0371	6.4008	29.8434
CAGR	0.2654	0.2144	0.1546	0.1196	0.4626
ROIC	0.1175	0.1093	0.0334	0.0804	0.1553
WACC	0.1032	0.1021	0.0121	0.0909	0.1167

B) Correlation matrix

LN(EVE)	CAGR	ROIC	WACC
1.0000			
0.2381	1.0000		
0.147	-0.0674	1.0000	
0.411****	0.2912*	0.3336**	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Telecommunication Services

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	10.6873	9.0353	4.6706	6.1369	18.1787
CAGR	0.0146	0.0153	0.0134	-0.0036	0.0337
ROIC	0.0159	0.0193	0.0167	-0.0114	0.0355
WACC	0.0562	0.0599	0.0076	0.0463	0.0648

B) Correlation matrix

LN(EVE)	CAGR	ROIC	WACC
1.0000			
0.3548	1.0000		
-0.2905	-0.2328	1.0000	
0.811****	0.7922****	-0.3505	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Transportation

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	9.2858	9.4634	4.5977	2.8008	15.5056
CAGR	-0.0622	-0.0400	0.0866	-0.1935	0.0214
ROIC	0.3412	0.3362	0.1542	0.1950	0.3950
WACC	0.0852	0.0847	0.0039	0.0801	0.0898

B) Correlation matrix

LN(EVE)	CAGR	ROIC	WACC
1.0000			
0.395	1.0000		
-0.0706	0.14	1.0000	
-0.0832	-0.3287	-0.3022	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Utilities

A) Summary statistics

Variable	Mean	Median	Std. dev.	10P	90P
EVE	11.3636	11.4872	3.9225	6.6172	15.1633
CAGR	0.1540	0.1420	0.0810	0.0817	0.2404
ROIC	0.0378	0.0383	0.0161	0.0153	0.0552
WACC	0.0638	0.0636	0.0055	0.0580	0.0700

B) Correlation matrix

LN(EVE)	CAGR	ROIC	WACC
1.0000			
0.0663	1.0000		
-0.1881	-0.2238	1.0000	
0.1219	-0.049	-0.0393	1.0000

**** p<0.01

*** p<0.025

** p<0.05

* p<0.1

Appendix 5 – White's test

Appendix 5.1 – White's test: Market, GICS sector, and GICS industry

Heteroskedasticity testing with White's test, across market, GICS sector and GICS industry

		MLR		SLR - growth		SLR - profitability		SLR - risk	
	N	WT	p-value	WT	p-value	WT	p-value	WT	p-value
Market	929	9.7110**	0.0000	0.2744****	0.0000	1.5537*	0.0725	2.6503****	0.0000
Energy Sector	26	19.3487	0.7388	7.6413	0.8718	2.1007	0.4599	0.779	0.2658
Energy	26	19.3487	0.7388	7.6413	0.8718	2.1007	0.4599	0.779	0.2658
Materials Sector	56	19.059	0.3020	0.5888****	0.0010	0.048	0.3498	2.3327	0.6774
Materials	56	19.059	0.3020	0.5888****	0.0010	0.048	0.3498	2.3327	0.6774
Industrials Sector	159	7.2183****	0.0000	2.0208***	0.0219	2.8507	0.2423	0.544*	0.0660
Capital Goods	109	7.9271****	0.0000	1.5631****	0.0005	5.6408****	0.0005	4.0235	0.3115
Commercial & Professional Services	32	6.9922***	0.0247	0.2133	0.7450	2.1711	0.9763	1.4473**	0.0486
Transportation	18	2.3622	0.6144	0.1051	0.3641	2.3888	0.2404	0.9369	0.7618
Consumer Discretionary Sector	144	14.446	0.4261	2.5154	0.4577	0.1542*	0.0596	1.3341*	0.0811
Automobiles & Components	18	13.5918	0.6379	0.5899	0.8988	3.7414	0.3377	6.5979	0.4850
Consumer Durables & Apparel	37	12.0928	0.9844	2.738	0.9488	0.4113	0.3029	2.6423	0.6260
Consumer Services	39	12.835	0.1073	0.2495	0.2843	3.1404	0.9258	5.3512	0.5132
Retailing	50	14.8243	0.1376	3.3056	0.7446	0.5743	0.1540	2.3788**	0.0369
Consumer Staples Sector	65	8.8605	0.2081	6.8174	0.2544	1.0657	0.8141	2.0076	0.2668
Food Staples & Retailing	13	16.7854	0.1702	2.9992	0.8827	2.2007	0.2080	0.1075*	0.0689
Food, Beverage & Tobacco	39	6.3902*	0.0959	0.2337	0.1915	1.7466	0.7504	0.1243	0.3044
Household & Personal Products	13	20.3916	0.4502	2.8317**	0.0331	0.4917	0.5869	1.8373	0.3665
Health Care Sector	123	11.378	0.1286	8.0264	0.2232	2.4071****	0.0022	3.4188	0.9477
Health Care Equipment & Services	76	10.2754	0.1224	7.2432	0.8897	1.7482**	0.0002	1.3543*	0.9398
Pharmaceuticals, Biotechnology & Life Sciences	47	13.7097	0.1025	2.3825	0.2427	4.0995	0.7820	7.6415	0.3991
Financials Sector	40	7.1496	0.2507	0.3137***	0.0181	0.0741	0.3001	20.1335	0.1810
Diversified Financials	40	7.1496	0.2507	0.3137***	0.0181	0.0741	0.3001	2.7857	0.1810
Information Technology Sector	138	5.049	0.1330	0.2968	0.3038	4.3774	0.1288	4.8695****	0.0001
Software & Services	56	13.9256****	0.0013	4.0967	0.8548	2.1235**	0.0365	2.0009****	0.0000
Technology Hardware & Equipment	45	10	0.6216	1.1436	0.2524	0.9636	0.9636	0.8824	0.2484
Semiconductors & Semiconductor Equipment	37	17.7617	0.8300	2.6689	0.8621	1.0141	0.1121	1.3704*	0.0876
Communication Services Sector	40	19.2659**	0.0320	0.2622	0.1289	2.9593	0.3458	8.5431	0.3677
Telecommunication Services	10	14.5323	0.3505	0.1243	0.5645	1.8567	0.6177	10.231	0.6433
Media & Entertainment	30	15.0224*	0.0380	10.0907	0.2633	4.5939	0.6023	0.1343	0.5040
Utilities Sector	52	3.1245**	0.0230	0	0.8771	0	0.2277	0.1343***	0.0140
Utilities	52	3.1245**	0.0230	0	0.8771	0	0.2277	0.1343***	0.0140
Real Estate Sector	86	0*	0.0903	0****	0.0064	0	0.1006	0	0.9351
Real Estate	86	0*	0.0903	0****	0.0064	0	0.1006	0	0.9351

**** p < 0.01
*** p < 0.025
** p < 0.05 ← *Our chosen significance threshold*
*** p < 0.1**

Appendix 5.2 – White's test: SARD groupings

Heteroskedasticity testing with White's test, across SARD groupings

		MLR		SLR - growth		SLR - profitability		SLR - risk	
	N	WT	p-value	WT	p-value	WT	p-value	WT	p-value
Energy	26	16.531	0.2735	14.935	0.2928	9.6833****	0.0053	5.7568	0.1946
Materials	56	3.769	0.3020	0.792*	0.0831	1.2202**	0.0471	1.3177	0.1149
Capital Goods	109	10.5642	0.6285	2.9459	0.6489	10.3247	0.7974	9.31	0.6748
Commercial & Professional Services	32	7.1004	0.9260	1.6573	0.6730	4.0069	0.5433	0.3652	0.5175
Transportation	18	13.5435	0.7147	1.0655	0.2499	0.1373	0.7802	0.9938	0.4337
Automobiles & Components	18	5.4572	0.6267	0.3647	0.4366	1.301	0.1349	0.1853	0.8331
Consumer Durables & Apparel	37	24.972	0.1395	2.9521	0.5870	0.8103	0.9336	14.9596	0.6084
Consumer Services	39	11.4644	0.2778	1.2777	0.3697	0.0535	0.2822	0.6609**	0.0250
Retailing	50	8.2657	0.7928	0.9556	0.8333	1.829	0.5218	1.456	0.9115
Food Staples & Retailing	13	33.4624	0.2452	0.3597	0.5279	16.4375	0.9736	1.259	0.7186
Food, Beverage & Tobacco	39	19.7512	0.5076	6.0585	0.6201	12.7963	0.4007	5.377	0.4829
Household & Personal Products	13	11.8824	0.6407	1.5499	0.4374	0.1762	0.1488	1.3151	0.8303
Health Care Equipment & Services	76	4.3290	0.1224	8.5428**	0.0484	3.894	0.1217	5.421	0.1680
Pharmaceuticals, Biotechnology & Life Sciences	47	26.616	0.1025	1.6338	0.4607	2.1328	0.9157	19.8442	0.5181
Diversified Financials	40	5.5967***	0.0129	0.2602***	0.0140	0.2706	0.1427	6.9078*	0.0665
Software & Services	56	13.0358**	0.0292	7.7284*	0.0695	0.1913	0.2861	1.3965****	0.0023
Technology Hardware & Equipment	45	0	0.2851	0.8924	0.2347	0.0054	0.3202	0.0059	0.1626
Semiconductors & Semiconductor Equipment	37	0.1015	0.3505	0.2165	0.3884	0.009	0.5097	0.3113	0.1754
Telecommunication Services	10	0	0.5463	0	0.3061	0	0.1943	0	0.1185
Media & Entertainment	30	3.5654	0.2015	1.0693	0.3255	0.7745	0.0052	0.5370	0.8335
Utilities	52	0	0.3183	0	0.2101	0	0.1725	0	0.1830
Real Estate	86	0	0.2220	0	0.3606	0	0.5389	0	0.6221

**** p < 0.01
*** p < 0.025
** p < 0.05 ← *Our chosen significance threshold*
* p < 0.1

Appendix 6 – Normality testing: skewness, kurtosis, and Shapiro-Wilk

Appendix 6.1 – Normality testing: market, sector, and industry

Normality testing with skewness, kurtosis and Shapiro-Wilk test, across market, sector, and industry

		MLR			SLR - growth			SLR - profitability			SLR - WACC		
	N	Skewness	Kurtosis	Shapiro-Wilk	Skewness	Kurtosis	Shapiro-Wilk	Skewness	Kurtosis	Shapiro-Wilk	Skewness	Kurtosis	Shapiro-Wilk
Market	929	-5.4652****	3.8509****	0.9605****	-5.9901***	2.2917****	0.9735****	-5.9502****	3.3976****	0.9833****	-6.3462****	3.6748****	0.963****
Energy Sector	26	1.5860	1.3734	0.9727	-1.3271*	1.2557	0.9818	-0.8679	1.9594	0.9869	1.6159**	1.4768	0.9718
Energy	26	1.5860	1.3734	0.9727	-1.3271*	1.2557	0.9818	-0.8679	1.9594	0.9869	1.6159**	1.4768	0.9718
Materials Sector	56	-1.3444	2.2778	0.983*	-2.5170	1.6096	0.9674	-0.2297	2.0526	0.9870	-1.1436	2.0366	0.9875**
Materials	56	-1.3444	2.2778	0.983*	-2.5170	1.6096	0.9674	-0.2297	2.0526	0.9870	-1.1436	2.0366	0.9875*
Industrials Sector	159	-2.2431****	3.0136**	0.9558****	-3.519****	1.7716****	0.9735****	-2.5905**	2.5468**	0.9829**	-4.535****	2.9013****	0.9537****
Capital Goods	109	-2.2015**	2.8395**	0.9553***	-2.9477****	1.6663****	0.9739****	-2.0982****	2.3904****	0.9839****	-3.8793**	2.7398*	0.9524*
Commercial & Professional Services	32	-2.4344	2.3715	0.9514	-1.1480	1.2243	0.9846	-1.4015	2.0168	0.9814	-2.3815	2.3270	0.9534
Transportation	18	-1.7204	2.6903**	0.96*	-1.3875	1.3375	0.9766*	-1.3825	2.1291	0.9739	-1.758*	2.2009	0.964**
Consumer Discretionary Sector	144	-0.7839	1.3762	0.9626	-0.6977	1.7584	0.9730	-0.6140	2.5260	0.9798	-1.6977	2.8005	0.9618
Automobiles & Components	18	-2.4044	0.5183	0.9380	-2.6929	1.7263	0.9415	-2.6918	2.5951	0.9584	-2.2104	2.2424	0.9482
Consumer Durables & Apparel	37	0.4352	1.4262	0.9681	0.7941	0.9476	0.9714	0.7783	2.0465	0.9838	0.9295	2.1537	0.9767
Consumer Services	39	-0.6711	0.9617	0.9548	-0.6577	1.3786	0.9776	-0.5788	2.1913	0.9751	-0.8014	2.3678	0.9477
Retailing	50	-2.3924	1.7001	0.9671	-1.2280	0.8626	0.9805	-1.0806	2.2462*	0.9755*	-0.9228*	2.3484*	0.9665*
Consumer Staples Sector	65	-2.7641****	1.6652	0.9647**	-2.0249	1.4225	0.9800*	-1.9940	2.2371	0.9788	-2.7868	2.4578	0.964*
Food Staples & Retailing	13	-1.6452*	2.5477*	0.9603	-0.9756	1.2469	0.9830	-1.3329	2.2435*	0.9709	-1.7992*	2.2928	0.9541
Food, Beverage & Tobacco	39	-2.5519	1.2988	0.9546	-1.6349	1.3261	0.9803*	-1.4711	2.0521*	0.9821	-2.5163	2.3636	0.9551
Household & Personal Products	13	-2.0601	2.8326	0.9467	-0.9883	1.2500	0.9829*	-1.7321	2.6430	0.9574	-1.1080	2.0674	0.9801
Health Care Sector	123	-2.5854****	2.8415****	0.9468****	-2.6442****	2.5695****	0.9803****	-2.1765****	2.4412****	0.9842****	-4.2021****	2.7676****	0.948****
Health Care Equipment & Services	76	-2.0563**	2.6741**	0.9489**	-2.5059**	2.5698144*	0.9743**	-1.7601*	2.2538	0.9842*	-3.4307*	2.5716	0.9479**
Pharmaceuticals, Biotechnology & Life Sciences	47	-2.8343	2.3972	0.9478	-1.3367	2.2748	0.9846	-1.4737	2.1030	0.9838	-2.8336	2.3986	0.9482
Financials Sector	40	-2.3513	2.5588	0.9612*	-1.3683*	1.6004	0.9633*	-1.7065	2.1373	0.9773	-2.1571	2.2666	0.9686
Diversified Financials	40	-2.3513	2.5588	0.9612*	-1.3683*	1.6004	0.9633*	-1.7065	2.1373	0.9773	-2.1571	2.2666*	0.9686
Information Technology Sector	138	-0.5197	3.0033	0.9540	-2.4256	2.7703	0.9716	-2.3438	2.4889	0.9838	-4.3026	2.8434	0.9523
Software & Services	56	-0.7880	2.6286	0.9502	-1.3280	2.5455	0.9717	-1.4692	2.1514	0.9849	-3.0412	2.4877	0.9481
Technology Hardware & Equipment	45	-0.4641	2.3808*	0.9484	-0.8336	2.3852	0.9784	-1.4604	2.0911	0.9837	-2.7665	2.3653*	0.9489
Semiconductors & Semiconductor Equipment	37	-0.3559	2.4268	0.9639	-1.9937	2.4750	0.9715	-1.5339	2.0567	0.9802	-2.1567	2.2654	0.9667
Communication Services Sector	40	-2.5517	2.4799*	0.9569	-2.0251*	1.4765	0.9722*	-1.5742	2.0715*	0.9803	-2.4976	2.3599*	0.9569**
Telecommunication Services	10	-2.2338	2.8977*	0.9326	-2.7459	2.0196	0.9010	-0.9326	2.0947*	0.9821	-1.7850	2.3660	0.9477
Media & Entertainment	30	-2.2312***	2.3593	0.9594***	-1.7278*	1.3966568*	0.9747*	-1.4502*	2.0327	0.9794**	-2.2076***	2.2832****	0.9589****
Utilities Sector	52	2.2783**	3.2191*	0.9377****	-2.5997****	1.6448****	0.9637****	1.8034****	2.4126**	0.9628****	-2.7579****	2.4359****	0.9569****
Utilities	52	2.2783**	3.2191*	0.9377****	-2.5997****	1.6448****	0.9637****	1.8034****	2.4126**	0.9628****	-2.7579****	2.4359****	0.9569****
Real Estate Sector	86	-2.2757****	2.6809****	0.9744****	-2.2274****	1.4710****	0.9808****	-1.294**	2.2467*	0.9875**	-2.9336****	2.5184*	0.9685****
Real Estate	86	-2.2757****	2.6809****	0.9744****	-2.2274****	1.4710****	0.9808****	-1.294**	2.2467*	0.9875**	-2.9336****	2.5184*	0.9685****

**** p < 0.01
*** p < 0.025
** p < 0.05 ← *Our chosen significance threshold*
* p < 0.1

Appendix 6.2 – Normality testing: SARD groupings

Normality testing with skewness, kurtosis and Shapiro-Wilk test, across SARD groupings

		MLR			**SLR - growth**			**SLR - profitability**			**SLR - WACC**		
	N	Skewness	Kurtosis	Shapiro-Wilk	Skewness	Kurtosis	Shapiro-Wilk	Skewness	Kurtosis	Shapiro-Wilk	Skewness	Kurtosis	Shapiro-Wilk
Energy	26	-2.6949	2.4528	0.9101	-0.3698	0.8500	0.9944	-3.0260	2.9523	0.9156	-1.8823	1.3242	0.9558
Materials	56	-2.1994	2.7092	0.9640	-1.0990	1.9470	0.9523	-0.1278	0.7002	0.9982	-0.6643	0.7172	0.9942
Capital Goods	109	-1.1166**	1.5979*	0.9218**	-1.0786**	2.5813	0.9232**	-1.0681*	0.796	0.9926*	-0.6411****	0.6611**	0.9953****
Commercial & Professional Services	32	3.0044	3.1430	0.9148	2.9674	3.1576	0.9126	-0.225	0.7428	0.9978	0.6005	0.656	0.9943
Transportation	18	-2.7469	3.0553	0.9211	-2.5615	2.9723	0.9314	-0.3840	0.8669	0.9950	0.6371	0.795	0.9922
Automobiles & Components	18	-2.9321	3.1151	0.9014	-2.3756	2.8664	0.9407	-0.6238	1.0255	0.9905	-0.6599	0.9355	0.9909
Consumer Durables & Apparel	37	-2.1735	1.8543	0.9628	-3.3479	3.1739	0.9239	-0.9505	1.0848	0.9887	-2.0315	1.2875	0.9623
Consumer Services	38	-2.9485	2.3691*	0.8891	-3.1182*	2.4961	0.8788	-1.0725*	1.2199	0.9872	-1.8683	1.1622	0.9656
Retailing	50	-3.0836	3.0687	0.9428	-3.8873	3.3393	0.9153	-0.8927	0.8989	0.9909	0.3257	0.5973	0.9964
Food Staples & Retailing	13	0.5605	0.3101	0.9726	-3.7507	3.6221	0.8147	1.2781	1.6997	0.9641	1.1221	1.0427	0.9742
Food, Beverage & Tobacco	39	-1.8235	2.4886	0.9683	-3.4041	3.1872	0.9241	0.9409	1.0982	0.9850	0.1276	0.642	0.9968
Household & Personal Products	13	-1.4323	2.4104	0.9639	-0.1573	1.1764	0.9950	-1.6735	2.3872	0.9590	-0.5038	1.0146	0.9925
Health Care Equipment & Services	76	-2.7544	1.7316	0.9336	-1.0832	0.9637	0.9934	-3.7249	3.0246	0.9418	-2.8710	1.4374	0.9521
Pharmaceuticals, Biotechnology & Life Sciences	47	-2.2772	1.4318	0.9575	-0.7077	0.759	0.9944	-1.1448	1.2073	0.9874	-2.2253	1.2679	0.9578
Diversified Financials	41	-2.8152**	2.2804	0.9052***	-2.4029*	2.1746*	0.9269****	-0.7747*	0.8572*	0.9916**	-1.2846**	0.9233**	0.9842**
Software & Services	56	-1.2685**	1.3271*	0.9896**	-1.0916*	1.1400	0.9919*	-0.5515	0.6888*	0.9955**	-0.9737**	0.787	0.9915****
Technology Hardware & Equipment	45	-2.2925	1.4334	0.9581	-1.2827	1.5477	0.9873	-1.7009	1.8473	0.9789	-2.3279	1.3488	0.9521
Semiconductors & Semiconductor Equipment	37	-0.5311	1.4993*	0.8766	-1.7592*	2.3563	0.9007*	-0.5860*	0.7586*	0.9939	-0.5592	0.7503*	0.9942
Telecommunication Services	10	2.4816	3.2970	0.8886	-3.7515	3.7004	0.8128	1.4019	1.9605	0.9548	-1.1211	1.3554	0.9745
Media & Entertainment	30	-1.2691	2.1525	0.9792	-2.0768	2.5005	0.9647	-0.2898	0.7493	0.9968	-0.071	0.7036	0.9968
Utilities	52	0.5466	1.6004	0.9878	-0.8351	0.837	0.9942	1.2069	1.1931	0.9822	-1.7565	1.0387	0.9766
Real Estate	86	-2.2818****	2.1641****	0.9781****	-0.8591****	0.7183**	0.9942****	-2.3372****	2.1268*	0.9771****	-0.0834****	0.5899***	0.9968****

**** p < 0.01
*** p < 0.025
** p < 0.05 ← *Our chosen significance threshold*
* p < 0.1

Appendix 7 – Two sample mean t-test on accuracy tests

Appendix 7.1 – Two sample mean t-test – accuracy test Market + GICS vs. PGA

Two sample t-test - Accuracy test - Market + GICS vs. PGA

		T-test		
	N	MSE	MAD	MAPE
Market	929	<0.0001	<0.0001	0.0057
Energy Sector	26	<0.0001	<0.0001	0.0004
Energy	26	<0.0001	<0.0001	0.0004
Materials Sector	56	0.2204	0.0177	<0.0001
Materials	56	0.2204	0.0177	<0.0001
Industrials Sector	159	<0.0001	<0.0001	0.0052
Capital Goods	109	<0.0001	<0.0001	0.0004
Commercial & Professional Services	32	0.0039	0.0001	0.1027
Transportation	18	0.3295	0.9099	0.1421
Consumer Discretionary Sector	144	<0.0001	0.0002	0.1774
Automobiles & Components	18	0.0426	0.0655	0.9376
Consumer Durables & Apparel	37	0.2180	0.8039	0.0427
Consumer Services	39	<0.0001	<0.0001	<0.0001
Retailing	50	0.0003	0.0117	0.5983
Consumer Staples Sector	65	<0.0001	0.0013	0.9214
Food Staples & Retailing	13	0.0111	0.0022	0.0030
Food, Beverage & Tobacco	39	<0.0001	<0.0001	0.1346
Household & Personal Products	13	0.0265	0.0648	0.4465
Health Care Sector	123	<0.0001	<0.0001	0.0002
Health Care Equipment & Services	76	<0.0001	<0.0001	0.0016
Pharmaceuticals, Biotechnology & Life Sciences	47	<0.0001	<0.0001	0.0206
Financials Sector	40	0.3462	0.1549	0.0002
Diversified Financials	40	0.3462	0.1549	0.0002
Information Technology Sector	138	<0.0001	<0.0001	0.6412
Software & Services	56	<0.0001	<0.0001	0.0532
Technology Hardware & Equipment	45	<0.0001	<0.0001	0.0081
Semiconductors & Semiconductor Equipment	37	0.0102	0.1040	0.2869
Communication Services Sector	40	0.0001	<0.0001	0.0111
Telecommunication Services	10	0.0294	0.0258	0.1672
Media & Entertainment	30	0.0012	0.0017	0.0838
Utilities Sector	52	<0.0001	<0.0001	<0.0001
Utilities	52	<0.0001	<0.0001	<0.0001
Real Estate Sector	86	<0.0001	<0.0001	<0.0001
Real Estate	86	<0.0001	<0.0001	<0.0001

H_0: Mean error difference = 0

H_A: Mean error differecnce ≠ 0

Blue shaded number are not significanct on p<0.05

Appendix 7.2 – Two sample mean t-test – accuracy test SARD grouping vs. PGA

Two sample t-test - Accuracy test - SARD groupings				
		T-test		
	N	MSE	MAD	MAPE
Energy	26	0.5432	0.0576	0.0116
Materials	56	0.0254	0.0022	0.3679
Capital Goods	109	0.0448	0.0431	0.0478
Commercial & Professional Services	32	0.0034	0.0018	0.0160
Transportation	18	0.0248	0.0121	0.0498
Automobiles & Components	18	0.0424	0.1064	0.4135
Consumer Durables & Apparel	37	0.0019	0.0032	0.0094
Consumer Services	39	0.3262	0.3211	0.3219
Retailing	50	0.1595	0.1569	0.1703
Food Staples & Retailing	13	0.3409	0.3393	0.3400
Food, Beverage & Tobacco	39	0.3217	0.3138	0.3085
Household & Personal Products	13	0.2978	0.3036	0.4075
Health Care Equipment & Services	76	0.3203	0.2970	0.3094
Pharmaceuticals, Biotechnology & Life Sciences	47	0.0003	0.0000	0.0000
Diversified Financials	40	0.3242	0.3164	0.3157
Software & Services	56	0.1587	0.1493	0.1559
Technology Hardware & Equipment	45	0.0283	0.0464	0.8354
Semiconductors & Semiconductor Equipment	37	0.0016	0.0026	0.0326
Telecommunication Services	10	0.0509	0.0559	0.1667
Media & Entertainment	30	0.0002	<0.0001	0.0001
Utilities	52	<0.0001	<0.0001	<0.0001
Real Estate	86	0.0001	<0.0001	<0.0001

H_0: Mean error difference = 0

H_A: Mean error differecnce ≠ 0

Blue shaded number are not significanct on p<0.05

Appendix 7.3 – Two sample mean t-test – accuracy test GICS vs. SARD grouping

Two sample t-test - Accuracy test - GICS versus SARD

		T-test		
	N	MSE	MAD	MAPE
Energy	26	0.0794	0.0530	0.3402
Materials	56	0.2727	0.1489	0.0301
Capital Goods	109	0.0449	0.0472	0.0586
Commercial & Professional Services	32	0.1181	0.0307	0.8234
Transportation	18	0.3253	0.2803	0.5767
Automobiles & Components	18	0.8498	0.5350	0.1626
Consumer Durables & Apparel	37	0.0856	0.0998	0.1354
Consumer Services	39	0.3265	0.3394	0.3562
Retailing	50	0.1595	0.1635	0.1718
Food Staples & Retailing	13	0.3409	0.3413	0.3402
Food, Beverage & Tobacco	39	0.3224	0.3561	0.5184
Household & Personal Products	13	0.3379	0.2901	0.3186
Health Care Equipment & Services	76	0.3216	0.3568	0.9105
Pharmaceuticals, Biotechnology & Life Sciences	47	0.9373	0.5976	0.6689
Diversified Financials	40	0.3245	0.3360	0.4023
Software & Services	56	0.1590	0.1636	0.2193
Technology Hardware & Equipment	45	0.5566	0.8951	0.3104
Semiconductors & Semiconductor Equipment	37	0.3402	0.5064	0.4441
Telecommunication Services	10	0.2988	0.2411	0.4814
Media & Entertainment	30	0.1611	0.1565	0.2036
Utilities	52	0.0288	0.0070	0.0085
Real Estate	86	0.1703	0.2892	0.1605

H_0: Mean error difference = 0

H_A: Mean error differecnce ≠ 0

Blue shaded number are not significanct on p<0.05

Appendix 8 – Shifted power transformation

Appendix 8.1 – Shifted power transformation – market, sector, and industry

Shifted power transformation -marked, sector, and industry

		Subset A: Output statistics				**Subset B: Coefficients**			
	N	R^2	Adj. R^2	SE	F-statistic	Intercept	Growth	WACC	WACC (shifted)
Market	929	0.1397	0.1369	0.5073	50.0567****	2.275****	0.009****	-1.537	0.74
						(6.745)	(8.528)	(-1.54)	(1.608)
Energy Sector	26	0.5234	0.4584	0.5023	8.0531****	-5.007	0.004	36.677	-17.262
						(-0.736)	(1.093)	(1.546)	(-1.57)
Energy	26	0.5234	0.4584	0.5023	8.0531****	-5.007	0.004	36.677	-17.262
						(-0.736)	(1.093)	(1.546)	(-1.57)
Materials Sector	56	0.3287	0.2900	0.4748	8.4882****	-4.641**	0.016****	20.18****	-9.295****
						(-2.044)	(3.957)	(3.202)	(-3.212)
Materials	56	0.3287	0.2900	0.4748	8.4882****	-4.641**	0.016****	20.18****	-9.295****
						(-2.044)	(3.957)	(3.202)	(-3.212)
Industrials Sector	159	0.1789	0.1630	0.4248	11.2582****	3.43****	0.008****	-6.441*	3.023*
						(2.683)	(3.152)	(-1.671)	(1.7)
Capital Goods	109	0.2134	0.1909	0.3913	9.4965****	5.747****	0.007***	-12.22**	5.664**
						(2.688)	(2.44)	(-2.107)	(2.127)
Commercial & Professional Services	32	0.1419	0.0499	0.4852	1.5429	5.15	0.002	-13.04	6.097
						(1.185)	(0.137)	(-0.888)	(0.897)
Transportation	18	0.0695	-0.1299	0.5355	0.3484	1.58	0.006	-0.335	0.191
						(0.479)	(0.601)	(-0.029)	(0.035)
Consumer Discretionary Sector	144	0.1244	0.1057	0.4643	6.6316****	2.064***	0.008***	-1.599	0.773
						(2.472)	(2.609)	(-0.627)	(0.657)
Automobiles & Components	18	0.6361	0.5581	0.3706	8.1558****	1.721	0.014*	-3.302	1.589
						(1.102)	(2.013)	(-0.773)	(0.811)
Consumer Durables & Apparel	37	0.1786	0.1039	0.3841	2.3918*	4.57****	0.008	-7.039*	3.234*
						(2.985)	(1.26)	(-1.712)	(1.713)
Consumer Services	39	0.3965	0.3448	0.3558	7.666****	3.487**	0.003	-9.172*	4.346*
						(2.342)	(0.658)	(-1.746)	(1.784)
Retailing	50	0.0699	0.0092	0.4866	1.1523	-0.869	0.002	8.148	-3.736
						(-0.426)	(0.259)	(1.193)	(-1.182)
Consumer Staples Sector	65	0.0961	0.0516	0.4114	2.1608	5.3****	0.001	-13.338*	6.242*
						(2.677)	(0.235)	(-1.753)	(1.764)
Food Staples & Retailing	13	0.3441	0.1255	0.3307	1.5741	10.553*	-0.003	-34.724	16.195
						(1.963)	(-0.296)	(-1.774)	(1.782)
Food, Beverage & Tobacco	39	0.1181	0.0425	0.4271	1.5626	4.226*	0.001	-10.256	4.832
						(1.822)	(0.179)	(-1.127)	(1.141)
Household & Personal Products	13	0.0228	-0.3030	0.4321	0.0699	6.428	-0.001	-14.171	6.579
						(0.697)	(-0.102)	(-0.431)	(0.431)
Health Care Sector	123	0.4777	0.4645	0.3717	36.276****	1.173*	0	-0.66	0.403
						(1.938)	(0.021)	(-0.383)	(0.509)
Health Care Equipment & Services	76	0.4556	0.4329	0.3871	20.0838****	1.264*	0.002	-0.649	0.393
						(1.746)	(0.61)	(-0.324)	(0.428)
Pharmaceuticals, Biotechnology & Life Sciences	47	0.5384	0.5062	0.3517	16.7202****	1.027	-0.001	-0.459	0.314
						(0.814)	(-0.495)	(-0.121)	(0.18)
Financials Sector	40	0.2219	0.1570	0.4954	3.4217**	2.887*	0.016****	-2.819	1.314
						(1.917)	(2.829)	(-0.578)	(0.584)
Diversified Financials	40	0.2219	0.1570	0.4954	3.4217**	2.887*	0.016****	-2.819	1.314
						(1.917)	(2.829)	(-0.578)	(0.584)
Information Technology Sector	138	0.3688	0.3547	0.4515	26.1005****	-0.8	0.009****	5.472****	-2.448***
						(-0.935)	(5.453)	(2.654)	(-2.602)
Software & Services	56	0.6941	0.6764	0.2935	39.322****	-1.73***	0.007****	7.878****	-3.535****
						(-2.568)	(4.241)	(4.737)	(-4.65)
Technology Hardware & Equipment	45	0.3275	0.2783	0.4629	6.6545****	-2.279	-0.001	6.983	-3.088
						(-0.806)	(-0.236)	(0.941)	(-0.909)
Semiconductors & Semiconductor Equipment	37	0.2738	0.2078	0.4841	4.1472***	2.341	0.011****	-1.485	0.702
						(0.942)	(3.168)	(-0.276)	(0.287)
Communication Services Sector	40	0.2990	0.2406	0.3289	5.1192****	1.207	0.005	1.459	-0.632
						(1.011)	(1.188)	(0.358)	(-0.336)
Telecommunication Services	10	0.7691	0.6537	0.2065	6.6629***	-1.508	0.038	17.257	-8.083
						(-0.871)	(1.938)	(1.918)	(-1.908)
Media & Entertainment	30	0.2591	0.1737	0.3546	3.0315**	2.737	0.006	-3.546	1.677
						(1.463)	(1.194)	(-0.584)	(0.599)
Utilities Sector	52	0.2297	0.1816	0.2257	4.7713****	5.339*	0.014****	-19.199	9.07
						(1.82)	(2.844)	(-1.214)	(1.219)
Utilities	52	0.2297	0.1816	0.2257	4.7713****	5.339*	0.014****	-19.199	9.07
						(1.82)	(2.844)	(-1.214)	(1.219)
Real Estate Sector	86	0.0918	0.0586	0.2766	2.7641**	0.928	0.005*	4.811	-2.204
						(0.803)	(1.926)	(1.068)	(-1.05)
Real Estate	86	0.0918	0.0586	0.2766	2.7641**	0.928	0.005*	4.811	-2.204
						(0.803)	(1.926)	(1.068)	(-1.05)

**** p < 0.01
*** p < 0.025
** p < 0.05 ← *Our chosen significance threshold*
* p<0.1

* Marked blue when both WACC and WACC (shifted) are simultaneously significant above threshold

Appendix 8.2 – Shifted power transformation – SARD groupings

Shifted power transformation -SARD groupings

		Subset A: Output statistics				**Subset B: Coefficients**			
	N	R^2	Adj. R^2	SE	F-statistic	Intercept	Growth	WACC	WACC (shifted)
Energy	26	0.0567	-0.0719	0.6848	0.441****	-30.681	-0.002	98.829	-45.521
						(-0.438)	(-0.432)	(0.443)	(-0.442)
Materials	56	0.1426	0.0931	0.4015	9.501****	5.816	0.022****	-19.418	9.133
						(1.051)	(2.88)	(-0.721)	(0.723)
Capital Goods	109	0.1622	0.1383	0.5543	2.882	5.23*	0.027****	-7.558	3.459
						(1.789)	(4.334)	(-1.01)	(1.01)
Commercial & Professional Services	32	0.1658	0.0764	0.3595	7.5104****	11.963	-0.063**	-37.726	17.599
						(0.761)	(-2.098)	(-0.558)	(0.557)
Transportation	18	0.1603	-0.0196	0.5751	6.7776**	14.094	0.026	-36.963	17.059
						(0.199)	(1.494)	(-0.176)	(0.177)
Automobiles & Components	18	0.1637	-0.0155	0.3907	1.8553****	3.865	0.025	-10.698	5.039
						(0.397)	(1.618)	(-0.194)	(0.194)
Consumer Durables & Apparel	37	0.3071	0.2441	0.3276	0.8911****	-3.615	0.016	8.524	-3.754
						(-0.394)	(0.717)	(0.383)	(-0.37)
Consumer Services	39	0.0977	0.0181	0.2882	10.1073	3.641*	0.022	-14.333	6.841
						(1.843)	(0.7)	(-1.023)	(1.029)
Retailing	50	0.1843	0.1311	0.5439	0.9137	-0.444	0.03****	8.778	-4.052
						(-0.036)	(3.181)	(0.289)	(-0.292)
Food Staples & Retailing	13	0.4851	0.3135	0.4052	4.876****	-90.11	0.099	397.951	-185.878
						(-1.361)	(0.831)	(1.483)	(-1.486)
Food, Beverage & Tobacco	39	0.1751	0.1044	0.2831	1.2276	0.416	0.029***	13.407	-6.369
						(0.24)	(2.399)	(1.198)	(-1.203)
Household & Personal Products	13	0.2955	0.0606	0.2333	3.4641****	-32.058	-0.004	196.731*	-92.711*
						(-1.784)	(-0.223)	(1.914)	(-1.914)
Health Care Equipment & Services	76	0.1435	0.1079	0.5810	2.685	2.166	0.001	-3.326	1.611
						(0.669)	(0.329)	(-0.404)	(0.428)
Pharmaceuticals, Biotechnology & Life Sciences	47	0.1038	0.0413	0.3921	2.8264***	13.665*	0.004	-43.075	20.032
						(1.76)	(0.409)	(-1.592)	(1.597)
Diversified Financials	40	0.2890	0.2313	0.3786	2.476*	-6.462	0.025****	52.701	-24.867
						(-1.334)	(2.874)	(1.688)	(-1.685)
Software & Services	56	0.1487	0.0995	0.4549	1.2582*	-7.831**	0.009	20.268****	-9.147****
						(-2.225)	(0.509)	(2.986)	(-2.984)
Technology Hardware & Equipment	45	0.1907	0.1315	0.5595	30.243*	-2.117	0.005	5.444	-2.368
						(-0.42)	(0.853)	(0.591)	(-0.571)
Semiconductors & Semiconductor Equipment	37	0.3934	0.3383	0.5094	4.0223	18.193*	0.031****	-33.556	15.221
						(1.804)	(3.953)	(-1.631)	(1.634)
Telecommunication Services	10	0.8801	0.8202	0.1769	1.6605****	-2.094	-0.242***	2.982	-1.04
						(-0.355)	(-3.266)	(0.073)	(-0.054)
Media & Entertainment	30	0.1622	0.0656	0.5473	6.3616***	-7.523	0.015	18.296	-8.23
						(-0.889)	(1.335)	(1.263)	(-1.267)
Utilities	52	0.0393	-0.0207	0.3567	5.0129	7.089	0.001	-28.539	13.465
						(1.308)	(0.213)	(-0.978)	(0.981)
Real Estate	86	0.0043	-0.0321	0.5027	23.8494****	2.319	0.001	1.344	-0.649
						(0.649)	(0.292)	(0.082)	(-0.084)

**** p < 0.01
*** p < 0.025
** p < 0.05   ← *Our chosen significance threshold*
* p<0.1

* Marked blue when both WACC and WACC (shifted) are simultaneously significant above threshold