SYSTEMATIC PEER SELECTION

A COMPARATIVE STUDY OF METHODS USED IN RELATIVE VALUATION

HAMPUS ADEBÄCK

DANTE HAQUÉS

Master Thesis

Stockholm School of Economics

2022

Systematic Peer Selection : A Comparative Study of Methods Used in Relative Valuation

Abstract:

With the purpose to provide insights for practical use and to contribute to an arguably under-served area of valuation research, this study performs a comparative analysis of systematic peer selection methods used in relative valuation. By using U.S. data between 2010-2020 and estimating valuations with the forward P/E multiple, we compare prediction errors of three methods presented in prior research with an additional comparison of the industry classifications SIC, GICS and the previously unexplored TNIC. We find that considering the business factors of industry classification (GICS) and product similarity (TNIC), is most important to improve valuation accuracy. As a result, we encourage future research to expand the use of different business factors in systematic methods. Further, we do not find support of the suggested notion that considering several fundamental factors in peer selection is able to replace the information provided by business factors. With support from theory, our results rather suggest these factors are not comparatively priced across industries. When considering several fundamental factors within industries, we find marginally increasing benefits on accuracy when using a non-linear multivariate method. Lastly, by analysing results over time, we find that since relative valuation is based on market values, its valuation accuracy is dependent on the stability of the overall market.

Keywords:

systematic peer selection, relative valuation, industry classification, product similarity, fundamental factors

Authors:

Hampus Adebäck (24121)
Dante Haqués (41790)

Tutors:

Stina Skogsvik, Assistant Professor, Department of Accounting

Examiner:

Department of Accounting

Master Thesis
Master Program in Accounting, Valuation and Financial Management
Stockholm School of Economics

Contents

1 Introduction **1**

2 Prior research and theory **3**
2.1 Valuation in practice . . . 3
2.2 Determining valuation accuracy . . . 4
2.3 Introduction to relative valuation . . . 5
2.3.1 The P/E valuation model . . . 6
2.3.2 Advantages and disadvantages of relative valuation . . . 7
2.4 Systematic methods to construct the peer group . . . 8
2.5 Industry classification . . . 10
2.6 Value drivers . . . 11
2.7 Formulation of the average . . . 12
2.8 Other impacting factors . . . 12
2.9 Prediction performance of different valuation models . . . 13
2.10 Summary and contributions . . . 13

3 Method and data **16**
3.1 Comparative methodological approach . . . 16
3.2 Comparative statistics . . . 18
3.3 The dependent valuation multiple: forward P/E . . . 19
3.3.1 Average formulation . . . 20
3.4 Common independent fundamental variables . . . 21
3.5 Industry classification . . . 22
3.6 Non-linear univariate method . . . 23
3.7 Linear multivariate method . . . 23
3.7.1 Independent variables . . . 25
3.8 Non-linear multivariate method . . . 26
3.8.1 Independent variables . . . 27
3.9 Empirical data . . . 28
3.9.1 Exclusions and adjustments . . . 28
3.9.2 Descriptive statistics . . . 29

4 Results and analysis **30**
4.1 Peer group construction using industry classification . . . 30
4.2 Peer selection using the non-linear univariate method . . . 32
4.3 Peer selection using the linear multivariate method . . . 33
4.3.1 Estimation regressions . . . 33
4.3.2 Prediction errors for the linear multivariate method . . . 35
4.4 Peer selection using the non-linear multivariate method . . . 37
4.4.1 Prediction errors for SADR on stand-alone basis . . . 37
4.4.2 Prediction errors of SADR and industry classification . . . 39
4.4.3 Prediction errors of SADR and TNIC . . . 39
4.5 Aggregated results and analysis . . . 42
4.5.1 Analysis across time . . . 42
4.5.2 Analysis across industries . . . 43
4.5.3 Analysis of overall accuracy . . . 44

5 Summary and conclusion **45**

6 References **47**
6.1 Published articles . . . 47
6.2 Books . . . 49

7 Appendix **50**

1. Introduction

This study performs a comparative analysis of systematic peer selection methods previously presented in relative valuation research.[1] Even though relative valuation is the most commonly used valuation method in practice, it has had limited coverage within academic research (Eaton et al., 2021). In practice, peer groups are often constructed on a subjective basis and is sometimes described by practitioners as an 'art-form' rather than a systematic method (Bhjojraj & Lee, 2002). On this note, Eaton et al. (2021) studies peer selection in practice and find that parties in M&A transactions motivates higher or lower valuations in negotiations by selecting peers accordingly. We argue that further examining systematic peer selection contributes to the dynamics of valuation negotiations in practice while expanding the academic coverage of an arguably under-served valuation research area.[2] To our knowledge, an 'objective' comparison of systematic peer selection methods without presenting a novel method has not been performed.

Obvious advantages of relative valuation is its parsimoniousness with low requirements of input information, swift application and straightforward comprehension, especially compared to more complex valuation models. This short-hand approach does however make implicit assumptions that rely on market values to be correct in aggregate, or else the valuation will fail (Sharma & Prashar, 2013). While the valuation relies on market values of most similar peers, it has further been shown that finding a 'perfect twin' is theoretically impossible (Skogsvik & Skogsvik, 2009). We attribute these shortfalls of relative valuation together with the seemingly subjective use within practice, as most probable reasons as to why it has been unfavoured in academic research.

Proponents of relative valuation do however assert additional advantages that stems from the same assumption that is also the culprit of its disadvantages, namely its reliance on market values. To include market sentiment in a valuation could be useful in short-term transactions, such as private M&A or IPOs, and also to sanity-check more comprehensive valuation models (Sharma & Prashar, 2013). In IPOs, relative valuation has shown promising performance by either outperforming or adding explanatory value to cash flow models (Kim & Ritter 1999; Kaplan & Ruback 1995). It is also generally accepted that no valuation model brings a perfect answer, which thus motivates the use of several models as each provides additional information (Yee, 2004). Therefore, regardless of one's view of the fairness of prevailing market sentiment, this suggests that the use of relative valuation adds information when performing valuations, especially in certain situations.

Out of the several aspects of relative valuation examined in prior research, possibly the most critical is peer selection (Plenborg & Pimentel, 2016; Eaton et al., 2021). Peer similarity is often determined from a combination of business and financial profile, often proxied with industry classification and fundamental factors respectively. Within prior research, it has been shown that choosing peers from the same industry improves the accuracy of relative valuation. It has also been shown that additionally filtering indus-

[1] Relative valuation assumes that one firm's valuation can be deduced from a cross-section of observed valuations of the considered firm's peers.

[2] Systematic peer selection uses pre-defined methods to avoid subjectivity.

try peers with fundamental factors, especially ROE, somewhat increases predictability (Alford, 1992; Cheng & McNamara, 2000). Previous results thus indicate that industry belonging captures much, but not all information used by the market to value firms, and that including fundamental factors in peer selection adds explanatory value.

Since industry classifications in their nature inherit problems of being static and subjective (Lee et al., 2015; Hoberg & Phillips, 2016), attempts have been made to formulate methods that reduce dependency of industry classification by selecting peers on several fundamental factors (Bhojraj & Lee 2002; Knudsen et al., 2017). There have also been attempts to redefine industry classifications using modern data analysis, either based on database search patterns (Lee et al. 2015), or text descriptions of product market (Hoberg & Phillips, 2016). When including fundamental factors in peer selection, methods either assumes a non-linear or a linear relationship between the fundamental factors and the valuation multiple. In our comparative analysis of systematic peer selection methods, we compare the non-linear univariate method based on Alford (1992) and Cheng & McNamara (2000), the linear multivariate model based on Bhojraj & Lee (2002) and the non-linear multivariate model based on Knudsen et al. (2017). Additionally, we compare the traditional industry classifications SIC and GICS to the product market based TNIC, developed by Hoberg & Phillips (2016), which to our knowledge is previously unexplored when examining valuation accuracy.[3] Eaton et al. (2021) found that TNIC has high explanatory value in predicting peers selected in practice.

Using U.S. data between 2010-2020 focusing on the forward price over earnings (P/E) multiple, we present several findings. First, we confirm previous results of increased valuation accuracy by using industry classifications in peer selection. Between industry classifications, we find marginal improvements using GICS compared to SIC, and to use TNIC compared to GICS*[4]. We therefore present novel results showing that product market similarity not only explains peer selection in practice, but also improves valuation accuracy. Using TNIC does however reduce available data, which reduces its practical relevance. By analysing results across industries and classifications, we identify the need of capturing several business factors in order to improve peer selection accuracy, which we encourage future research to further examine. Secondly, we find no value in using the linear multivariate method, attributed to limited theoretical congruence and an increased underlying volatility in the P/E multiple compared to the stock-based valuation multiples used by Bhojraj & Lee (2002). Further, we find no support for the non-linear multivariate method's ability to replace the information provided by industry classifications, as suggested by Knudsen et al. (2017). Our analysis suggests that since the market prices fundamentals differently across industries, selecting peers by only considering fundamental factors is unlikely to succeed. Combining the non-linear multivariate model with industry classification, however, does improve accuracy for up to five fundamental variables. We argue that this could be relatively easily used to improve valuation accuracy in practice. Lastly, our results across time show how relative valuation's reliance on market sentiment

[3]Abbreviations of Standard Industry Classification (SIC), Global Industry Classification Standard (GICS), and Text-based Network Industry Classifications (TNIC).

[4]GICS* represents the GICS classification applied to the limited data sample available when using TNIC

affects its performance, where increased uncertainty and volatility leads to a reduction in accuracy.

The rest of this paper is structured as follows. In Section 2, a walkthrough of prior research and theory concerning the different aspects of relative valuation is presented. In Section 3, the general method of the comparative study as well as more detailed descriptions of each systematic peer selection method and data are presented. In Section 4, our results with related discussion is presented. Lastly, in Section 5, our conclusions, limitations and suggestions for further research are presented and discussed.

2. Prior research and theory

This section presents a combination of prior research and theory concerning the most important considerations when performing relative valuation. The aim of the section is to follow a typical process of performing relative valuation while presenting the most important research findings and theory to explain each step. While our comparative analysis targets systematic peer selection models and thus focuses on the area of peer group construction, description of other steps is used to motivate choices made in other areas of our method described in Section 3.

The section is structured as follows. In Section 2.1, to build relevancy, findings on the use of different valuation methods in practice are presented. In Section 2.2, error measures used to compare accuracy of different valuation models and methods are described. In Section 2.3, an introduction to relative valuation with a further description of advantages and disadvantages is presented. In Section 2.4, a walkthrough of different systematic methods to construct the peer group from prior research is described. In Section 2.5, detailed descriptions of different industry classifications and considerations are presented. In Section 2.6, findings of choices considering value drivers are described. In Section 2.7, prior results concerning the formulation of the average are presented. In Section 2.8, factors impacting relative valuation but not specifically considered in this paper are presented. In Section 2.9, prediction error results from prior studies across different valuations methods are presented. Lastly, in Section 2.10, a summary of prior research along with a discussion of how this paper contributes to relative valuation research is presented.

2.1 Valuation in practice

Even though there might exist discrepancies between academia and practice on the topic of valuation, practitioners play an important role in the functioning of capital markets. Simultaneously, valuation methods are equally important to investment professionals' success of investment decisions in a highly competitive market environment. Over the years there has been recurring studies capturing statistics on valuation methods used in practice, either by studying analyst reports or using surveys (Pinto et al., 2018).

By examining analyst reports, Demirakos et al. (2004) study 104 reports in the UK and finds that 67% includes relative valuation, 16% discounted cash flow valuation, 10% residual income valuation and 7% other methods. Further, Asquith et al. (2005) examine 1,126 reports in the U.S. where they find that 99% used earnings multiples, 25% asset multiples, 13% discounted cash flow valuation and 4% other methods. The latest and most comprehensive collection using a survey is Pinto et al. (2018) who collected 1,980 responses from equity analysts with CFA membership. Results show that a market multiples approach is used by 93% of respondents, a present value by 79%, an asset based by 61%, real options by 5% and other methods by 13%. When asked to estimate the percentage of valuation cases where the respective methods are used, the average estimation was 69% for market multiples and 60% for present value approaches, with below 50% rates for the other methods. The authors conclude and claim relative valuation and present value methods as the two primary tools for equity valuation. Concludingly, all studies indicate an extensive use of relative valuation in practice with a concurrent high relevancy for valuation research.

When further analysing the use of different valuation multiples, results from Pinto et al. (2018) show that the two most popular ratios used by respondents are P/E at 88% and enterprise value over operating profit, most often EBITDA[5], at 77%. The popularity of P/E has been longstanding (c.f. Graham & Dodd, 1934; Block, 1999) with respondents favouring forecasted earnings over historical figures (Pinto et al., 2018).

2.2 Determining valuation accuracy

Before detailing relative valuation, it is useful to illustrate how to evaluate the accuracy of different valuation methods. To examine this, a measurement basis has to be established. In the majority of the relative valuation literature, the observable market price is used to calculate the prediction error of the valuation model (e.g., Boatsman & Baskin, 1981; Alford, 1992; Kaplan & Ruback, 1995; Kim & Ritter, 1999; Cheng & McNamara, 2000; Lie & Lie, 2002; Liu et al., 2002). Other measures are also possible, where e.g. Bhojraj & Lee (2002) measure prediction errors versus future market values up to three years, with the claim that market prices converges to intrinsic values over time. Sloan (2002) however questions the validity of this by arguing that if current market prices are unreliable, so should also future market prices. The study from Bhojraj & Lee (2002) does also include prediction errors using current market prices to allow comparability of their results to previous studies.

On the same note as Sloan (2002), Anesten et al. (2020) provide an interesting discussion on the use of prediction errors as an evaluation method. Since market prices might not be efficient, one might want to test valuation methods with their return as a trading strategy. However, as shown by Anesten et al. (Appendix 1, 2020), assuming a persistent market mispricing and a limited investment period, ranking of valuation methods based on trading strategies will in expectation be the same as ranking on current pricing prediction errors. This thus supports the use of prediction errors when comparing accuracy across different valuation methods.

[5]Earnings before interest, tax, depreciation and amortisation

For cross-sectional comparison in prediction errors, the measure needs to be on the same scale. The most common approach is to scale by the actual market price, from here on defined as the percentage error (PRE) (e.g. Alford, 1992; Liu et al., 2002).[6] When it is useful to consider a measure without economic value assigned to either over- or under predictions, absolute values of the pricing error can be considered, from here on defined as the absolute percentage error (APRE). To control for skewness caused by outliers and that stock prices are bound by zero, it is also common to use power functions in the measure, often as sensitivity tests as they are less intuitive for interpretation (Kaplan & Ruback 1995; Liu et al. 2002). For given firm at time *t* the prediction errors is expressed as follows.

$$PRE_t = \frac{V_t(Mv) - P_t}{P_t} \tag{1}$$

$$APRE_t = \left| \frac{V_t(Mv) - P_t}{P_t} \right| \tag{2}$$

where V(Mv) is the estimated market value and P is the actual market value.

When analysing the distribution of prediction errors there are two aspects to consider; precision and dispersion. To analyse precision, an average measure is used, where values closer to zero indicates lower prediction errors. In many previous studies on relative valuation, mean of yearly medians is used as the precision measure (Alford, 1992; Cheng & McNamara, 2000). To analyse dispersion, a measure for spread is used, where more observations closer to the average indicates lower dispersion. Most often, interquartile ranges are used as the measure of spread (Liu et al., 2002). In the overall assessment, there is thus a trade-off between precision and dispersion (Faber, 1999).

2.3 Introduction to relative valuation

Relative valuation is a method in which the value of a firm is deduced from valuation multiples of a peer group.[7] Valuation multiples are ratios with an expression of firm value in the numerator, either equity or enterprise value, divided by a value driver in the denominator. Firm value expressions are observed market prices, most often from public markets, but also disclosed M&A transactions or IPO offerings. There is a range of value drivers, either operational or financial, and either flow- or stock-based.

Operational drivers range from website visits to barrels of oil equivalents, while the most commonly used are financial measures. These are in turn most often accounting based, such as sales, an earnings measure, a balance sheet item or a cash flow measure. Flow-based drivers are time restricted such as earnings while stock-based are time specific such as book value of equity. The time period of value drivers are discretionary, which means either historical figures or future estimates drawing on forecasts from analysts or

[6]Abbreviated in previous research as PE and APE but here called PRE and APRE not to be confused with the P/E multiple.

[7]Also known as the market multiples approach or comparable companies analysis.

management. The following formulas and steps illustrates the usual process of performing relative valuation.

$$\begin{matrix} Valuation \\ multiple_{n,t} \end{matrix} = \frac{Observed\ firm\ value\ measure_{n,t}}{Value\ driver_{n,t+i}} \tag{3}$$

Where for each Firm *n* in the considered peer group, the valuation multiple is calculated at time *t* with their respective value measure at time *t* and value driver at time *t+i*.

$$\overline{\begin{matrix} Valuation \\ estimate_t \end{matrix}} = Avg.\left(\sum_{n=1}^{N} \begin{matrix} Valuation \\ Multiple_{n,t} \end{matrix}\right) \times Value\ driver_{t+i} \tag{4}$$

Where the valuation estimate for given firm at time *t* will be a result of an average measure of the peer group valuation multiples at time *t* multiplied with the value driver of given firm at the previously chosen point of time *t+i*.

For theoretical congruence, it is favourable to match the value driver to the right measure of firm value (Sharma & Prashar, 2013). While considering value drivers attributable to both shareholders and debtholders, such as operating earnings, enterprise value is congruently used. While considering a value driver attributable only to shareholders, such as net earnings, equity value is used instead. By combining valuation multiples with theoretical valuation models, expressions of multiples with fundamental factors as independent variables can be deduced. To provide an example of this while considering our focus on the P/E multiple, the following section will deduce a forward P/E valuation model.

2.3.1 The P/E valuation model

The theoretical concept of the P/E multiple is that the expected value of future earnings should be reflected in the price, compared to either historical or a future earnings. This implies that P/E prices expected earnings growth (Penman, 2013, p.179).

The P/E multiple for given firm at time *t* using forward earnings is described as follows, limited to non-negative numbers.

$$(P/E)_t = \frac{P_t}{Earnings_{t+1}} \tag{5}$$

Where P is the current market value and Earnings is one year forecasted. By adding a fraction of the book value of equity B, the model can be described as follows.

$$(P/E)_t = \frac{P_t/B_t}{Earnings_{t+1}/B_t} = \frac{(P/B)_t}{ROE_{t+1}} \tag{6}$$

Where the numerator is the commonly used P/B valuation multiple and the denominator is the end period book return on equity.

The Dividend Discount Model (DDM) from Williams (1939) can further be used as a simple expression of the market value.

$$P_t = \sum_{s=1}^{\infty} \frac{E_0(D_{t+s})}{(1+r_e)^s} = \sum_{s=1}^{\infty} \frac{E_0[B_{t+s-1} \times (ROE_{t+s} - g_{t+s})]}{(1+r_e)^s} \tag{7}$$

Where D is the expected next year dividend, g is the growth of book value of equity, r is the cost of equity *e*, while the RHS derivation assumes that the clean surplus relation holds.

Combining the expression of market value in Equation 7 with the earlier expression of the P/E multiple in Equation 6, our deduced model of the forward P/E for given firm at time *t* is expressed as follows.

$$\frac{P_t}{E_{t+1}} = \sum_{s=1}^{\infty} \frac{E_0[\prod_{v=0}^{s-1}(1+g_v) \times (ROE_{t+s} - g_{t+s})]}{E_0(ROE_{t+1}) \times (1+r_e)^s} \tag{8}$$

Theoretically, these are the fundamental factors that should affect observed P/E ratios where P/E is positively correlated with higher profitability and growth, while negatively correlated with higher cost of capital and thus risk.

2.3.2 Advantages and disadvantages of relative valuation

As previously discussed, the wide practical use of relative valuation could be explained by advantages for practitioners non-related to subjective use. Roosenboom (2007) investigated valuation methods used by practitioners for IPOs and found that relative valuation was preferred due to its ability to reflect market sentiment. Sharma & Prashar (2013) discuss the implications of market sentiment and argues that relative valuation is more likely to reflect market perception and investor sentiment, which is of importance for valuation of transactions with a short-term perspective. Furthermore, Kaplan & Ruback (1995) compare relative and fundamental valuation and argue that despite the two displaying similar predictability, the authors suggests using both methods in a practical setting given the additional information provided by relative valuation.

There are however also disadvantages to relative valuation stemming from its reliance on market prices. In the short-hand approach, implicit assumptions are made of considerations explicitly addressed in more complex fundamental models (Sharma & Prashar, 2013). On this note, Penman (2013, p.77) highlights the issue of circularity when solely relying on relative valuation. This issue arises since if the valuation of one firm in a peer group is dependent on the valuations of the other firms, then their valuations are in turn based on that one firm. Penman (2013) furthermore debates the use of relative valuation for public firms due to its reliance on observable market prices, which then questions why the observable market price of the considered firm would not be reliable. Instead, Penman (2013) argues relative valuation being more beneficial for either private firms, firms with insufficient trading liquidity or for some reason an unreliable market price.

Eaton et al. (2021) further examine the use of relative valuation in M&A processes, where investment banks are compensated through a success fee tied to the transaction value. Hence, the banks have a monetary incentive to inflate valuations. Even though this also can be achieved using fundamental valuation, the authors argue it is less complicated to underpin adjustments of this sort when using relative valuation.

Lastly, while Sharma & Prashar (2013) raise several advantages with relative valuation, the authors also acknowledge the difficulty finding peers as no two assets can be exactly similar. A theoretically derived argument of the same nature is put forward by Skogsvik & Skogsvik (2009), where the authors highlight the theoretical impossibility in finding an "identical twin" in peer selection. The authors however suggest that the peer selection process, if anyhow used, can be improved by controlling for industry, accounting principles and expected book return as a first step.

2.4 Systematic methods to construct the peer group

Considering the difficulties of finding peers, it is not surprising that valuation literature agrees that the methodology of selecting peers is a crucial step of relative valuation (see e.g., Plenborg & Pimentel, 2016; Eaton et al., 2021). Despite extensive efforts in finding peers with the most similar cash flow characteristic could be undergone, it contradicts the parsimonious advantage of relative valuation. There is thus a clear trade-off between accuracy provided by a certain method and its required effort (Plenborg & Pimentel, 2016).

As mentioned, when examining peer selection in relative valuation research, peer similarity is deduced from business profile via industry classification and financial profile via fundamental factors. In pursuit of increasing valuation predictability, these similarities are often considered simultaneously (Plenborg & Pimentel, 2016). When using methods to consider similarity in fundamental factors, these are either non-linear or linear in their assumption of the relationship between the fundamental factors and the valuation multiple. In addition to one business profile factor, the methods are also capable to consider either one fundamental factor, which we define as univariate, or several factors simultaneously, which we define as multivariate.[8]

Boatsman & Baskin (1981) were among the first to use a non-linear univariate method. The authors examine the effect of both industry belonging and financial factors when comparing the P/E ratio of an arbitrary firm from the same industry to another firm with a similar ten-year average growth rate. By examining 80 firms and only selecting one peer based on either industry or the financial factor, the authors find that average earnings growth has lower prediction errors than an arbitrary industry peer.

Alford (1992) presents a more comprehensive non-linear univariate method by examining how industry, risk, earnings growth and leverage, both individually and in pairs, impact

[8]Technically, the univariate methods are bivariate as data intersections between two variables are used. Most often, the first variable is industry belonging while the second is a fundamental variable, which is why we consider the method as essentially univariate.

the accuracy of constructed peer groups. For industry, peers are selected on an increasingly detailed SIC code, from one to four digits, until each group contained at least six peers. For risk (total assets as proxy) and earnings growth (ROE as proxy), the two percent most similar firms for each factor were selected. When combined, the intersection between each respective factors groups were combined to include at least six firms. Results show that industry on its own, or combined with risk or earnings growth, rendered the lowest prediction errors. This indicates that industry captures most of the information provided by risk and earnings growth. For level of detail in SIC, results showed that prediction errors decrease from one to three digits, while four-digit codes do not generate noteworthy higher prediction than three-digit codes.

Cheng & McNamara (2000) use a similar selection process to that of Alford (1992), besides limiting risk (size) and earnings growth (ROE) to include six firms in purpose to match the industry factor. The results confirmed those of Alford (1992), showing that industry, on its own or combined with risk and earnings growth, had the highest accuracy. The authors do however indicate that industry combined with earnings growth slightly outperform the two other definitions, indicating that industry does not capture all information from earnings growth. The authors also show that using a combination of the P/E and P/B multiples reduces prediction errors compared to the use of only P/E.[9]

One important consideration for both Alford (1992) and Cheng & McNamara (2000) is the least number of firms for the the peer groups. Cheng & McNamara (2000) examine this by presenting their results with a range of least number of firms (LNOF) and confirm that six firms per group has the highest accuracy. This is probably the result of the tradeoff between the detail of the industry code (1-4 digits) and the number of available firms within that industry to use in valuation. On this note, Cheng & McNamara (2000) also find that accuracy increases for four-digit industries by the numbers of firms available within the industry.

Bhojraj & Lee (2002) present a linear multivariate method that constructs "warranted multiples", or estimated multiples, with cross-sectional annual linear regression estimates on EV/S and P/B multiples, regressed on fundamental factors deduced from valuation theory. One advantage with the linear multivariate method is that several variables can be considered simultaneously. Another advantage by using it on stock-based valuation multiples is the possibility to use the method on firms with negative earnings. One disadvantage is that the model assumes a linear relationship between the multiple and the independent variables, while theoretically deduced models of multiples are not linear (Sloan, 2002). Other drawbacks include inconvenience for practical use and reduced data samples to estimate regressions on in smaller markets (Knudsen et al., 2017).

The warranted multiples can both be used as a valuation multiple directly, but also indirectly by considering the market multiples of the peer group where the warranted multiples are solely used for peer ranking similarity. Results indicate that both direct and indirect application of the warranted multiples improve accuracy compared to an industry

[9]While Cheng & McNamara focus on studying combinations of valuation multiples it is not detailed here considering our focus on peer selection.

peer group based on a two-digit SIC code. Bhojraj & Lee (2002) claim their results show potential for their method to replace industry classifications since considering several fundamental factors simultaneously adds sufficient informative value.

Finally, Knudsen et al. (2017) present a non-linear multivariate method, named the sum of absolute rank differences approach (hereinafter SARD), neither limited to one fundamental factor or assumes a linear relationship between the dependent and independent variables. The SARD approach is instead similar to the clustering algorithm usually considered as the 'manhattan distance'. The authors agree with the notion that selecting peers on fundamental variables increases predictability when considering factors of growth, profitability, and risk. Like Bhojraj & Lee (2002), Knudsen et al. (2017) suggest that the ability of their method to simultaneously regard several fundamental factors allows it to be independent of industry classification.

The SARD method ranks the absolute difference between two companies on the basis of several fundamental variables and finds step-wise decreases in APRE as each fundamental variable is added. The authors test the approach across the forecasted EV/EBIT, EV/Sales, P/B and P/E and present similar decreasing patterns in APRE across all valuation multiples. Knudsen et al. (2017) find that combining SARD with industry classification has a lower APRE than when only considering fundamental factors. However, only considering fundamental factors also outperform industry classification. The authors therefore suggest the ability of the method to replace industry classification over time, especially in markets with smaller samples of firms. Lastly, Knudsen et al. (2017) test the optimal LNOF for their results, showing the same results in the comparison between methods but a range of 6-16 firms depending on the valuation multiple considered.

2.5 Industry classification

Considering the centrality of industry classification in systematic peer selection methods, it is of importance to consider the implications of using different definitions. As previously mentioned, industry belonging is a mean to distinguish companies with similar business profile in regards to economics and operations, assumed to impact valuation on the notion that similar assets should trade at similar prices (Fama & French, 1992).

Despite its wide use in both practice and academia, there are few papers investigating similarities and differences in the different classification systems (Hrazdil et al., 2013). One study is (Bhojraj et al., 2003), which compares SIC, GICS, NAICS and the Fama & French (1997) adjusted SIC (FF). The authors find that the GICS classification is superior both in explaining stock movements and cross-sectional variation in valuation multiples. Hrazdil et al. (2013) similarly show that GICS was superior in capturing homogeneity in a peer group using linear regression.

Furthermore, there are alternative systems intended to capture additional similarities not captured by the classifications above. Lee, Ma & Wang (2015) show how peer selection based on co-search patterns on SEC's database EDGAR presents promising results to be an alternative classification. This is however less relevant for practitioners as the data is not publicly available (Plenborg & Pimentel, 2016). Another alternative classification is presented by Hoberg & Phillips (2016). As mentioned, the TNIC system uses cluster text analysis from company annual filings to determine peers based on product offerings. For each firm, the TNIC dataset offers a unique peer group of comparable companies each year, along with a 'score' variable representing the level of product similarity.[10]

The TNIC classification shows promising results in several studies. Hoberg & Phillips (2008) find that product similarity increases probability of M&A transactions. Additionally, Hoberg et al. (2014) show that product market influences financial flexibility and pay-out policies. The TNIC classification also indicates relevance in peer selection, where Eaton et al. (2021) show a higher explanatory value of TNIC than SIC industry codes when peers are subjectively chosen to motivate valuations in M&A transactions.

2.6 Value drivers

After the peer-group has been constructed, the next step is to decide value driver and the valuation multiple. As mentioned, the value driver of choice is favourably matched to the congruent value measure. Important value driver considerations are as follows.

One aspect to take into consideration is whether to use a cash flow or accrual-based value driver. Proponents of cash flow-based multiples argue that accrual measures are less reliable due to arbitrary allocation procedures. Cash flow-based drivers do however not match income to expenses for a given transaction (Plenborg & Pimentel, 2016). Liu et al. (2002) find that accrual-based multiples have higher predictability on U.S. data. The authors furthermore suggest that earnings-based multiples are superior to book value-based. Schreiner & Spremann (2007) extend their study to other countries and similarly conclude that accrual value drivers, especially earnings, performs best. Kim & Ritter (1999) examine performance of P/E and P/B multiples in valuing IPO's and conclude that P/B is a better indicator of performance. Cheng & McNamara (2000) compare performance for the same multiples and show that the P/E is the single most accurate multiple (Cheng & McNamara, 2000).

Another aspect to be considered is whether a historical or forecasted measure should be used. Plenborg & Pimentel (2016) argue that since the numerator in a valuation multiple reflects market prices which by default are forward looking, the numerator should match this. The same argument is proposed by Kim & Ritter (1999), who also demonstrate forward-looking P/E multiples outperforming its historical based counterpart. The authors also show that P/E multiples based on a two-year forecast horizon outperforms that of one-year. There indeed seems to be an academic consensus on the superiority of forecasted earnings (Lie & Lie, 2002; Liu et al., 2007; Schreiner & Spremann, 2007), with further indications that longer forecasts yield better results (Liu et al.,2002)

A final consideration to take into account is that of non-transitory items.[11] Plenborg & Pimentel (2016) discuss certain issues regardingly and bring to surface the problem with

[10] See an example of this in Appendix 1.
[11] Such as impairments and legal settlements.

subjectivity when deciding for an eventual non-recurring nature of an item. Liu et al. (2002) compare IBES earnings adjusted for non-recurring items to COMPUSTAT earnings which are not, and find that IBES earnings outperforms. Nissim (2013) also compares performance of multiples adjusted for non-recurring items, and argues that adjusted earnings do not guarantee a better performance.

2.7 Formulation of the average

After the peer group has been constructed and value drivers decided, a final step to consider is how to formulate an average of the multiples in the peer group. See the three most common averages from prior research below.

$$\begin{matrix} Arithmetic \\ mean \end{matrix} = \frac{\sum_{i=1}^{n} Multiple_i}{n} \quad \begin{matrix} Harmonic \\ mean \end{matrix} = \frac{n}{\sum_{i=1}^{n} \frac{1}{Multiple_i}} \quad Median = \begin{matrix} central \\ position \end{matrix}$$

Several studies conclude evidence for harmonic mean being the most accurate (Baker & Ruback, 1999; Liu et al., 2002; Dittman & Maug, 2006), whereas one common argument for the dominance is that it avoids impact of extreme values. Even if median also reduces the impact of outliers, most studies still find harmonic mean as superior. These findings are possibly the result of a trade-off between under-estimations and sensitivity to outliers. Herrmann & Richter (2003) identify a sharp improvement for the harmonic mean compared to the median if the 1% of outlier multiples in the peer group are excluded. Additionally, Alford (1992) and Cheng & McNamara (2000) use medians to estimate values while Knudsen et al. (2017) use the harmonic mean after confirming a in sensitivity check that it outperforms any other average measurement.

2.8 Other impacting factors

Finally, there is a plethora of other aspects that in the best of worlds should be considered both for peer group construction and when interpreting results. One is the impact of accounting regime, which either stems from a difference in geography or industry. Young & Zeng (2015) find that different accounting practices in the peer group can make dissimilar companies appear similar, which will skew the multiple from the peer group.

Another aspect is the impact of size & maturity, where several studies have found that valuation accuracy will increase with these, hence making it more challenging to value smaller firms (Alford, 1992; Kim & Ritter; 1999; Cheng & McNamara, 2000; Lie & Lie, 2002). For less mature firms, negative earnings is a common trait, which poses a challenge in using earnings-based multiples.

Further, Pratt et al. (2008) argue that the benefit of high liquidity in a publicly traded stock should be awarded with a premium, as opposed to a private firm with little to none liquidity. The authors also suggest that stock prices often reflect the value of a minority stake, which necessarily not applies to a larger chunk of the company due to the associated inherent control (Pratt et al., 2008). Also, in the process of an IPO there is often a discount present, which affects the translation of relative valuations to private firms (Kim Ritter, 1999).

Finally, different firms and industries with different cost structures poses dissimilar risk of being subject to conservative bias in valuation procedures. Runsten (1998) however showed that companies in the same industry tend to have similar Q-values, which accounts for conservative bias.

2.9 Prediction performance of different valuation models

To compare the accuracy of relative valuation to other fundamental valuation methods, it is of interest to present a summary of prediction errors from previous studies and methods. Important to consider is that prediction error measures use market prices, where relative valuation is a method that, quite reasonably, captures market sentiment to a larger extent (Sharma & Prashar, 2013). As previously mentioned however, using PRE and APRE is also common when evaluating fundamental methods. Even if there might be periods where market values divert from intrinsic values, the validity of fundamental models should hinge on their ability to reflect market prices over time. Also, as mentioned, if considering situations where market values in the short-term is desirable, such as private M&A and IPOs, the PRE is a valid measure to consider. Table 2.1 is a collection of prediction errors presented by selected studies across relative-, cash flow based- and accounting based valuation. Important factors to consider for these results that prevents direct comparison is differences in time period and methods.

As can be seen by Table 2.1, while considering points discussed above, relative valuation has a high performance compared to fundamental models. These results seemingly support the use of relative valuation in certain situations and the focus of this study.

2.10 Summary and contributions

To conclude this section, we highlight the relevance to examine relative valuation with regards to its popularity in practical use and higher relative performance compared to other valuation models, especially in certain short-term focused situations such as private M&A transactions and IPOs. We further show that there are some considerations of relative valuation where prior research is mostly aligned. These are not limited to but include; the performance of value drivers where P/E consistently performs with the highest or equally high valuation predictiveness, the higher performance of using forecasted estimates and the higher performance of using harmonic mean. Considering the popularity and performance of the P/E, we also deduce a P/E valuation model to illustrate the implied independent fundamental variables affecting the dependent P/E multiple. Where prior research is less aligned, is the area of peer selection and the dynamics of systematically assessing similarity across a combination of business profile and fundamental factors.

Our contribution to relative valuation research is two-fold. First, we compare the more traditional industry classifications SIC and GICS to the more novel product similarity classification of TNIC. Product similarity shows high relevance for peer selection in prac-

Table 2.1: Overview of prediction accuracy for relative and fundamental valuation in prior research

Author (Year)	Sample	Model Specifications	Prediction Accuracy		
			Mean PRE	*Median PRE*	*Mean APRE*
Relative Valuation					
Alford (1992)	U.S. firms 1978, 1982 and 1986	Non-linear method, historical P/E	-	-	0.239*
Cheng & McNamara (2000)	U.S. firms 1973-1992	Non-linear method, historical P/E	-	-	0.245*
Bhojraj & Lee (2002)	U.S. firms 1982-1998	Linear regression, historical P/B	-	-	0.51
Knudsen et al. (2017)	U.S. firms 1995-2014	Sum of Absolute Rank Difference, Forecasted P/E	-	-	0.306
Fundamental Valuation					
Francis et al. (2000)	U.S. firms 1989-1993	DDM (n=5)	-0.68	-0.69	0.69**
		RIV (n=5)	-0.20	-0.23	0.30**
		FCF (n=5)	0.18	-0.08	0.41**
Corteau et al. (2001)	U.S. firms 1992-1996	DDM (n=5)	-0.24	-0.31	0.40
		RIV (n=5)	-0.30	-0.34	0.37
Chang et al. (2012	U.S. firms 1980-2010	RIV (n=5)	-0.37	-0.40	0.47
		AEG (n=5)	-0.16	-0.21	0.66
Ho et al. (2017)	U.S. firms 1985-2013	RIV (n=5)	-0.06	-0.09	0.82

*Mean of medians APRE is used. **Median APRE is used

tice but is previously unexplored in relative valuation research examining the valuation accuracy of peer selection. Secondly, we perform a comparative analysis of previously presented systematic peer selection methods with the same general assumptions and definitions, based on results in prior research, and on the same dataset. To our knowledge, no comparative study without suggesting a novel method exist, which could affect the 'objectivity' of previous methods, and our study thus contributes to an area with previously mixed results (Plenborg & Pimentel, 2016). Also, revisiting previous methods gathers insights on updated market data.

To summarise, the analysis of results aims to contribute with a better understanding of the valuation information provided by industry classification or product similarity, the effect on predictability by including fundamental factors in peer selection and if it is possible for a multivariate method to compensate for the information provided by industry classifications. The relevant prior research presenting systematic peer selection methods, together with the scope of this study, are summarised in Table 2.2.

Table 2.2: Overview of previous research on systematic peer group construction for relative valuation

Author (Year)	Sample	Peer Construction			Multiples	Mean formulation	Error prediction	Results
		Methods	*Matching criteria*	*Industry*				
Boatsman & Baskin (1981)	80 U.S. firms 1957-1956	Two methods of Non-linear univariate matching	(i) Industry, (ii) Industry and earnings growth	3-digit SIC	Historcial P/E	Arithmetic Mean	Absolute Prediction Error (APRE)	Peer groups matched by industry and earnings growth have higher predictability
Alford (1992)	651 U.S. firms 1978, 1982 and 1986	Seven methods of Non-linear univariate matching	(i) Market, (ii) Industry, (iii) Total Assets (TA), (iv) Industry + TA, (v) Industry + ROE, (vi) TA + ROE	4-1-digit SIC	Historical and Forecasted P/E	Median	Absolute Prediction Error (APRE)	Peer groups matched by Industry + ROE have highest predictability
Cheng & McNamara (2000)	30,310 U.S. firm observations 1973-1992	Seven methods of Non-linear univariate matching	(i) Market, (ii) Industry, (iii) Total Assets (TA), (vi) Industry + TA, (v) Industry + ROE (vi) TA + ROE	4-1-digit SIC	Historical P/E and P/B	Median	Absolute Prediction Error (APRE)	P/E multiples applied on groups matched by Industry + ROE have highest single predictability. P/E and P/B combined have highest predictability overall
Bhojraj & Lee (2002)	1,498 U.S. firms 1982-1998	Three methods of Linear multivariate regression	(i) WM directly applied, (ii) WM ranking and WM applied, (iii) WM ranking and EV/S and P/B applied	3-digit SIC	Historical EV/Sales and P/B	Harmonic Mean	Absolute Prediction Error (APRE)	WM applied as ranking to find closest peers and application of P/B multiple performed best
Liu et al. (2002)	26,613 U.S. firm observations 1982-1999	One method of Non-linear univariate matching	Industry	IBES Based	19 variations inc. both equity and EV multiples	Harmonic Mean	Absolute Prediction Error (APRE) and Prediction Error (PRE)	Two-period forecasted P/E performed best, historical EV/S performed worst
Lie & Lie (2002)	8,261 global firms in 1998	One method of Non-linear matching	Industry	3-digit SIC	Historical EV/B, EV/S, EV/EBITDA, EV/EBIT, PE and Forecasted P/E	Arithmetic Mean and Median	Absolute Prediction Error (APRE) and Prediction Error (PRE)	One-period forecasted P/E performed best among equity multiples, EV/B performed best among EV multiples
Knudsen et al. (2017)	12,350 U.S. firm observations 1995-2014	Non-linear multivariate matching	Subsequent addition of ROE, DEBT/EBIT, Size, Growth and EBIT margin	6-digit GICS	Forecasted EV/S, EV/EBIT, P/B and P/E	Harmonic Mean	Absolute Prediction Error (APRE)	Finds that SARD using all selection variables outperforms traditional Industry-based peer groups
Adebäck & Haqués (2022)	**19,291 U.S. firm observations 2010-2020**	**Non-linear and linear univariate and multivariate matching**	**Several**	**8-2-digit GICS 4-1-digit SIC TNIC**	**Forecasted P/E**	**Harmonic Mean**	**Absolute Prediction Error (APRE)**	**To be seen in Section 4.0**

3. Method and data

In this section the method and data of our study will be presented and discussed. The aim of the section is to follow a structure of considerations ranging from general choices applicable throughout to increasingly more specific ones. Hence, the section starts with an introduction to the approach of our comparative study and error measurements. After that, the section will cover the considerations of the dependent value driver, independent fundamental variables and industry definition used across, before presenting the details of each of systematic peer selection method. Lastly, the data sample used accompanied with observation exclusions, variable adjustments and descriptive statistics is presented.

The section is structured as follows. In Section 3.1, the introductory methodological approach of the comparative study along with prediction errors is discussed. In Section 3.2, the choice of value driver is described and motivated along with the average formulation. In Section 3.3, common independent fundamental variables used across methods are described. In Section 3.4, the use of industry classifications is presented. In Section 3.5, the details of the non-linear univariate model is presented. In Section 3.6, the details of the linear multivariate method is described. In Section 3.7, the details of the linear multivariate model is presented. In Section 3.8, the data sample and descriptive statistics along with exclusions and adjustments is presented.

3.1 Comparative methodological approach

As mentioned, the main purpose of this comparative study is to measure the relative performance across the non-linear univariate (Alford, 1992; Cheng & McNamara, 2000), the linear multivariate (Bhojraj & Lee, 2002) and the non-linear multivariate (Knudsen et al., 2017) systematic peer selection methods, while also including a comparison across industry classifications. The advantage of revisiting previously presented methods in the same study is that surrounding considerations, definitions and data, apart from the peer selection method themselves, are the same which allows for isolation of effects from choice of method. To make congruent decisions for surrounding considerations, but also adjustments to ensure comparability across the different methods, a number of guiding principles have been followed in our methodological approach, as described below.

First, the aim to the largest extent is to keep the integrity of previous models intact, but in order to ensure comparability across methods, congruent considerations across models has to be made. In other words, there is a trade-off between the integrity and comparability of the methods in this study. Therefore, our results are not directly comparable or forthright replications of the previous studies, but rather based on these. The guiding principle on this note is to adjust the previous methods in accordance with indicated 'best practices' from the collection of prior research.

Second, a guiding principle is to ensure that the peer selection methods used are able to estimate the valuation of a private firm. Even if relative valuation could be used to value public firms, prior research agrees that relative valuation is often more relevant in a private setting. A key example of this principle is to exclude the use of observable market prices for independent variables in any method, but also to exclude the market multiple of the firm under consideration in any valuation prediction, such as industry or market averages. Of course, all firms included in the data of this study are publicly traded, yet their observable market prices are only used in order to evaluate performance of the valuation estimates in the prediction error measures and not as input.

As presented in Section 2.8, there is a plethora of impacting factors that can be considered when performing relative valuation. Some of these are mitigated through our method, while some are considered out-of-scope and might thus have explanatory value in explaining results. The guiding principle on this topic is to reduce impact of neglected factors on comparability between methods by giving different methods the same conditions, to include sensitivity checks if needed or prioritising the other principles.

One example of the factors not considered in detail for this study is the number of firms chosen to form the peer group for each method, which also decides the LNOF in an industry to be included in the data sample. Since Alford (1992) introduced to use six firms as LNOF, it has been standard when examining peer selection in relative valuation research. As mentioned, both Cheng & McNamara (2000) and Knudsen et al. (2017) provide tests of the optimal LNOF where the former confirms six while the latter rather find a range between 6-16 firms optimal, depending on valuation multiple. Prediction results between different methods has however not been affected in these tests, which means conforming to our principle of comparability without a detailed analysis. Also, there is a trade-off between the quality and quantity of the peer group when using LNOF for practical relevance, where Eaton et al. (2021) further show that the average number of peers in practice are between 8-10. This also means conforming to our practical relevance principle while further allowing for comparability to previous studies.

Other factors not specifically considered in the study, which may affect the translation of results to a private M&A or IPO setting, are the notions of illiquidity premium, control premium and IPO discount (Sloan, 2002). Compared to a public setting which typically reflects a minority price (Plenborg & Pimentel, 2016), these notions are, as mentioned, found to respectively have a negative, positive and negative effect on firm valuation. This study will treat these notions in expectation as zero but acknowledges that one might want to adjust results from relative valuation if a premium or discount is believed to exist. We argue that relative valuation should still add information to a valuation while also referring to prior studies showing promising performance of relative valuation in e.g. IPOs.

Some additional factors to consider will indirectly be mitigated with our method, including firm size, geography and accounting regimes. First, the relation between prediction errors and firm size will not be analysed in detail as in previous research where firm size is found to be positively correlated with lower prediction errors. It will however be included in methods that consider size as a fundamental factor. Second, effects from geographies will be mitigated considering the sole focus on U.S. data. Third, accounting regimes and adjustments will be indirectly mitigated with the use of U.S. data, through constructing peer groups and reporting results based on industries classification which mitigates domestic accounting variations, and by the use of forecasted numbers which should reduce accruals of non-recurring character.

Lastly, the three considerations of the dependent value driver, common independent fundamental variables and industry classifications will be covered in further detail below as they have higher impact on the adjustments required for the different methods. In addition, how requirements of different methods impacts data used is described last.

3.2 Comparative statistics

A central part of this comparative study is to compare results from different methods and to determine the statistical significance of the differences between them.

In line with previous studies, PRE and APRE will be used to measure precision. In results, the mean of yearly medians, medians and means of PRE and APRE will be presented. The reason as to why both the mean of yearly medians and medians are presented, is to allow for comparability to both Alford (1992) and Cheng & McNamara (2000), who present the mean of yearly medians, and to Knudsen et al. (2017) who use medians and means. However, small differences across these measures are expected.

As already presented in Section 2.2 formula (1), the PRE is expressed as.

$$PRE_t = \frac{V_t(M) - P_t}{P_t}$$

Where the mean of yearly medians, median and mean for years *y* and firm observations *n* are expressed as follows.

$$\begin{matrix} Mean\ of \\ medians \end{matrix} = \frac{\sum_{t=1}^{y} Median\ PRE_t}{y} \qquad Median = \begin{matrix} central \\ PRE_i \end{matrix} \qquad Mean = \frac{\sum_{i=1}^{n} PRE_i}{n}$$

The way the expressions of PRE are interpreted is that values close to zero represents high precision of valuation estimates, negative values represent under-predictions on average and positive values represents over-predictions on average. Since PRE includes both negative and positive values that could cancel out, it is important to consider both dispersion and APRE to not overstate precision.

As also already presented in Section 2.2 formula (2), the APRE is expressed as follows.

$$APRE_t = \left| \frac{V_t(M) - P_t}{P_t} \right|$$

Where the mean of yearly medians, median and mean are expressed the same as for PRE, while using APRE instead of PRE.

The interpretation of APRE is that a value close to zero represents high precision since it only considers the magnitude of the prediction error in any direction. The economic meaning of APRE is intuitive, since if one assumes that market prices on average correspond to true values, then the APRE is how much the valuation estimate on average deviates from the true value (Alford, 1992; Knudsen et al. (2017).

For both PRE and APRE, dispersion is considered with the IQR where a lower score is interpreted as lower dispersion. Since APRE only considers positive values, IQR values are expected to be lower than for PRE.

$$IQR = Q3 - Q1 \tag{9}$$

To ensure that difference of results across methods are statistically significant, Wilcoxon signed rank tests are performed on APRE. Since the economic meaning of APRE is most intuitive and thus relevant for discussion, it is on this measure the statistical significance analysis is performed, in line with e.g., Knudsen et al. (2017). The Wilcoxon signed rank test is the non-parametric counterpart of the more traditional t-test, however without necessarily assuming normal distribution for the dataset investigated. According to the law of large numbers, it is true that the APRE dataset should follow a normal distribution which would motivate the use of regular t-tests. However, previous studies have assessed that prediction error based datasets not necessarily follow this pattern (Knudsen et al., 2017).

The test is performed on the basis if whether introducing either a variation of a method or a new method renders statistically significant differences. First, the APRE difference in each observation for the first and second method is computed. Second, the differences are ranked in absolute terms and assigned a rank from 1 to n, depending on absolute difference. Third, each absolute rank observation is assigned a plus or minus depending on if the difference between the first and second method is positive or negative. Finally, given the subsequent introduction of additional methods and pursuit in understanding if the prediction errors of the new method is significantly lower than the former, a one sided test is performed using the following hypotheses.

$$H_0 : Median\ difference\ =\ 0$$
$$H_1 : Median\ difference\ >\ 0$$

3.3 The dependent valuation multiple: forward P/E

To focus this study on the comparative performance across methods and to simplify comprehension of presented results, only one valuation multiple is considered in the study. Of course, to ensure comparability, the same valuation multiple is used across methods. With regards to its extensive practical use (Pinto et al, 2018) and its generally most accurate performance on its own in previous studies (e.g. Cheng & McNamara, 2000; Liu et al., 2002; Schreiner & Spremann, 2007) this study will limit its focus to the P/E multiple. Despite other valuation multiples could perform at a higher level compared to P/E with certain methods, the relative performance when comparing across multiples and methods should be considered as the end result is always a prediction of a price. This reasoning should thus motivate the use of the multiple with the past superior prediction performance. Examples of multiples that are not considered in this study but used in studies on which we base our methods include P/B, EV/S and EV/EBIT (Cheng & McNamara, 2000; Bhojraj & Lee, 2002; Knudsen et al., 2017).

Furthermore, since current market values often are based on future expectations (Plenborg & Pimentel, 2016) and its previous proven accuracy (e.g. Alford, 1992; Kim & Ritter, 1999; Liu et al., 2002), the earnings measure will be based on one-year analyst forecasts to increase relevance, compared to presenting results using historical numbers. Despite forecasts beyond one-year has proven to increase accuracy (Liu et al., 2002), we deem one-year forecasts should maximise the available data sample without affecting the comparative performance of the different methods, which follows our principles of practical relevance while maintaining comparability between methods. Since Alford (1992) and Cheng & McNamara (2000) both examine current P/E ratios, our choice to use forward P/E does however somewhat affect the comparability to their results.

As previously shown in Formula 8, the forward P/E model is deduced as follows.

$$\frac{P_t}{E_{t+1}} = \sum_{s=1}^{\infty} \frac{E_0[\prod_{v=0}^{s-1}(1+g_v) \times (ROE_{t+s} - g_{t+s})]}{E_0(ROE_{t+1}) \times (1+r_e)^s}$$

Estimating P/E multiples through this fundamental model will not be covered in this study as many implications would follow. Examples of these include handling forecasts of firms not yet in competitive equilibrium (Skogsvik & Skogsvik, 2009). The model is however used indirectly in the study by indicating important independent variables to be considered in the following methods. If deemed necessary, methods can be adjusted in favour of what is suggested by this model. Essentially, what is suggested, is that P/E should be driven by profitability, growth and risk, which has been long withstanding in previous research (e.g. Alford, 1992). Fundamental factors considered important in these categories include book returns or profitability margins for profitability, growth in earnings or book values for growth, and size or cost of capital for risk.

3.3.1 Average formulation

To transfer the peer group valuation multiples to a valuation estimate, the harmonic mean will be used throughout this study, considering the results of prior studies (e.g. Baker & Ruback, 1999; Liu et al., 2002; Dittman & Maug, 2006; Knudsen et al., 2017). Even if Alford (1992) and Cheng & McNamara (2000) use median, we argue that harmonic mean has the highest relevance considering the comparison between the measures made by Knudsen et al. (2017) and when also trimming our dataset for outliers (see discussion in Section 2.7). Regardless method used to construct the peer group, the harmonic mean will then be used to estimate the valuation multiple that together with the forecasted earnings of the considered Firm with peer firms *n* will result in the valuation estimate at time *t*, illustrated below.

$$\overline{P/E}_t = \frac{n}{\sum_{i=1}^{n} \frac{1}{P/E_i}} \times E_0(Earnings_{t+i}) \qquad (10)$$

3.4 Common independent fundamental variables

As earlier emphasised, to ensure comparability across methods that in different ways incorporates fundamental factors, common definitions of independent variables are used. From the discussion above, these should in some way reflect either profitability, growth or risk. Those of which are used in two or more methods are described below, while those specific to one method are described in their respective sections describing independent variables, along with discussions of how originally used definitions might differ from the common definitions described here.

Return on equity (ROE)
Considering its high presence in prior research (Plenborg & Pimentel, 2016), perhaps the most central independent variable for the P/E multiple is the return on equity. For theoretical congruence, we will use expected book ROE since we are examining forward P/E multiples, as suggested by Skogsvik & Skogsvik (2009) and illustrated in the deduced forward P/E model in formula (10). This differs from all previous peer selection methods which use current book ROE. Despite ROE typically is seen as a measure of profitability (Cheng & McNamara, 2000), given that the clean surplus relation holds, growth is also a function of this (Skogsvik & Skogsvik, 2009). Illustratively, Alford (1992) classifies ROE as a proxy for growth. We choose to view ROE as a hybrid of mainly profitability but also an indicator of growth, which could also be the reason it has had high explanatory value in previous univariate studies. Therefore, in addition to this hybrid, we add one measure of growth and profitability each, to ensure capturing the effects of these aspects.

Implied growth
In line with previous studies (Bhojraj & Lee, 2002; Knudsen et al., 2017), the proxy for growth will be a measure using earnings growth. The way we deduce growth does however somewhat differ from previous methods. In order to reduce missing values and to maximise the length of the growth forecasts, we use a step-wise model extracting analyst earnings forecasts up to three years for all observations from IBES. With these, we annualise expected earnings growth from current earnings up to the longest forecast available for each firm observation.[12]

Operating profit margin
To the extent that ROE does not capture the aspect of profitability, like previous studies (Bhojraj & Lee, 2002; Knudsen et al., 2017), we include a measure of current operating profitability, defined the same way as in previous studies, with operating earnings over revenue. We expect it to have low explanatory value given its absence in the deduced P/E model, but choose to include it for completeness and comparability to prior studies.

Equity beta
In addition to independent variables used by previous studies, we choose to include the equity beta in the multivariate models given its indirect impact on the cost of capital. Beta values are used as a proxy for risk in equity valuation (e.g. Damodaran, 1994; Damodaran,

[12] As a result, 62% of our observations uses a three year forecast, 36% uses a two year forecast and 2% uses a one year forecast

2002), where Damodaran (2002) argues it is the best estimate for risk when valuing firms using P/E multiples. We estimate CAPM equity beta values for firm observations by measuring volatility compared to the S&P 500 during a 60-month period.[13]

Introducing the equity beta does however borderline cross our practical relevance principle as estimating CAPM betas requires firms to be publicly traded (Alford, 1992). Estimating betas for private firms is thus harder and often includes the process of using comparable firms (Damodaran 2002), which might generate circularity. We do however argue that its inclusion holds theoretical relevance while it is possible to estimate the beta of private firms by regressing historical earnings against market returns. Although, for practical use and parsimoniousness, we do acknowledge its lesser relevance.

3.5 Industry classification

As a first step of analysis, this study will in addition to a comparison of systematic peer selection methods also compare results across the industry classifications SIC, GICS and TNIC. To anchor benefits of using industry classifications, the prediction error of the average total market will be presented as a baseline (hereinafter Market). Based on the results of this comparison, the industry classification with lowest prediction errors will be used throughout the rest of the comparative study. Since SIC is used by Alford (1992), Cheng & McNamara (2000) and Bhojraj & Lee (2002) while Knudsen et al. (2017) use GICS, both are included as a control before only selecting one, even though GICS has shown superior results in other studies (e.g. Bhojraj et al. 2003).

TNIC's relevancy stems from its absence in prior research even though product market has shown high explanatory value for peer selection in practice (Eaton et al., 2021). TNIC will however not be considered across all selection methods for two reasons. First, the intersection between TNIC and the otherwise required data, reduces annual firm observations with approximately 50%, which would reduce comparability in relation to previous research and practical relevance. Secondly, most likely due to time needed to perform the text based cluster analysis, there is a time lag on TNIC data with 2019 as the most recent year, which also impacts practical relevance. In the first industry classification comparison, TNIC will be included along with a sensitivity test of the best performing classification out of SIC and GICS on the TNIC intersectional data. Also, since TNIC data contains the interesting aspect of a 'score' variable for each peer, TNIC is dedicated its own Section 4.3 where it is combined with the non-linear multivariate method.[14]

Lastly, the definition of industry classification and thus measures such as industry average will be used congruently across all considered peer selection methods. Considering previous results showing that increasing detail of industry codes by Alford (1992) and Cheng & McNamara (2000), at least up to 3-digits, and to maximise the dataset for practical relevance, the analysis will use the step-wise definition of industry used by these two studies, described in further detail in the section below. Adhering to the best practice principle,

[13]The full period available is used for firms included in the sample that have been publicly listed for less than 60 months.

[14]The TNIC 'score' variable determines the level of product market proximity to the considered firm.

this will thus be used rather than the original definitions in both of the non-linear and linear multivariate methods.

3.6 Non-linear univariate method

A non-linear univariate peer selection method in two steps will be examined based on the method and results of Alford (1992) and Cheng & McNamara (2000).

Industry
First, as mentioned as a first step used as a baseline of methods compared, will be peer selection based solely on industry averages using SIC, GICS and TNIC classifications. In both prior studies, to assure a large enough set of comparable industry firms, the least number of firms (LNOF) required in the industry group is set to six. Starting at the most detailed 4-digit SIC code, if the Number of Firms (NOF) is below six, the industry group is instead based on the 3-digit SIC, and so forth until the LNOF of six is satisfied. If a firm does not have six other firms in its 1-digit SIC industry group, the firm observation is emitted. Cheng & McNamara (2000) study the optimal LNOF requirement and found that six firms minimised the prediction error.[15] For GICS, the same step-wise approach is used but instead on its 8- to 2-digits, while the TNIC data does not contain hierarchical levels.

Industry & ROE
Secondly, as examined by both Alford (1992) and Cheng & McNamara (2000), a peer selection method combining industry belonging with one-year expected ROE is used. Similarly in this study, six firms are selected within each industry group with the closest ROE. If the industry group contains six firms, the two methods will thus be the same.

It is worth mentioning that both Alford (1992) and Cheng & McNamara (2000) evaluate several variations of non-linear peer selection methods not used in this study using different combinations of industry, ROE and total assets. However, only Industry + ROE is assessed in this study given that it received the lowest APRE in both studies.

3.7 Linear multivariate method

The linear multivariate method is largely based on the method of Bhojraj & Lee (2002) who perform annual cross-sectional regressions on either EV/S or P/B as dependent variables. A regression is run on the entire dataset for each year, with independent variables derived from fundamental derived factors to reflect profitability, growth, and risk. The coefficients from each regression is used to form what is referred to as 'warranted multiples', simply described as the multiple a given firm 'deserves' based on the actual values of the independent variables and formed coefficients.

Apart from the argument that EV/S and P/B multiples performed at a low level in previous research and the possibility to use them to value firms with negative earnings, the authors

[15] The proportion of observations satisfying the requirement for 4-digit codes in the two studies were 55% and 54% respectively, where we expect a similar result.

do not present a reason to not include P/E while acknowledging that future research could extend their method to other multiples. Building on the above discussion regarding the popularity and proven accuracy of P/E, we thus choose to perform that extension now. For the P/E multiple, we argue that most of the same independent variables used by Bhojraj & Lee (2002) for the P/B multiple should be relevant.

However, it is worthwhile to mention expected implications of adjusting the dependent variable from the P/B and EV/S to a forward P/E multiple. The use of the forward P/E has broad academic support of superiority in terms of APRE (e.g. Cheng & McNamara, 2000; Liu et al., 2002; Schreiner & Spremann, 2007). However, Damodaran (2002) argues that regressions with the P/E multiples as a dependent are more volatile compared to stock-based ones such as P/B or EV/S, due to the smaller absolute values of earnings. Damodaran (1994) illustrates this effect, where the R-squared in five regressions from year 1987 to 1991 ranges from 0.93 to 0.32. Results in this study from the linear regression approach is not believed to fluctuate to the same extent but is still subject to the issue. Also, given the use of the P/E, we will be unable to consider firms with negative earnings which will thus affect the comparability of the results of Bhojraj & Lee (2002).

As previously discussed, but important to emphasise, even though stock-based multiples might be more suitable for the linear multivariate method, we argue that the validity of the method hinges on its ability to outperform other methods with the overall best performing valuation multiple. This is a result of the fact that each method across all valuation multiples results in a price prediction. As an example, assuming that the linear univariate method outperforms industry average when considering EV/S, but the industry average P/E outperforms both EV/S predictions – a simple use of the P/E with addition of the value of net debt will make the EV/S predictions irrelevant from a practical perspective.

As already mentioned, Bhojraj & Lee (2002) have another industry definition than Alford (1992) and Cheng & McNamara (2000) by defining their industry peer groups on 2-digit SIC codes. To ensure isolation of explanatory value from fundamental factors and to increase comparability across methods, we will use the same industry definition as in the non-linear univariate method. Lastly, to maintain focus on examining systematic peer selection methods, we will not focus heavily on examining the accuracy of warranted multiples valuation directly, which is more similar to hedonic valuation models. Rather, focus will be more on it as a method to construct a peer group. When doing this, Bhojraj & Lee (2002) construct the peer group with the four closest peers or the single closest peer in terms of warranted multiples. Once again, to increase comparability between methods, we will instead form peer groups of six. The used independent variables x are detailed below to be used in the following annual estimation regressions for each firm.

$$P/E_{t+1} = \alpha_t + \sum_{x=1}^{7} \beta_{x,t} X_{x,t} + \epsilon_t \tag{11}$$

3.7.1 Independent variables

To ensure integrity and comparability of the model used in the study by Bhojraj & Lee (2002), all their variables will be used, yet with the required adjustments for congruence across methods and the addition of equity beta.

Industry mean
The harmonic mean of P/E for the industry group based on step-wise industry classification instead of the original 2-digit SIC. Price is the market capitalization while earnings is the one-year forecast of earnings. The variable is expected to be positively correlated to the forward P/E.

Adjusted operating profit margin
The industry adjusted profit margin is defined by taking the difference between the operating margin of earnings for the firm subject to valuation to the median of profit margins in the industry group. Our common definition of the operating margin is the same as the one originally used. The variable is expected to be positively correlated with the forward P/E to the extent it adds explanatory value to ROE.

Adjusted growth
The industry adjusted earnings growth is defined by taking the difference of our common definition of growth to the median of the earnings forecasts in the industry group. In their study, Bhojraj & Lee used the long-term growth in operating earnings from IBES which we found to highly reduce the dataset. The variable is expected to be positively correlated with the forward P/E.

Leverage
Book leverage is long-term debt scaled by the book value of common equity. Included for completeness without expected correlation with the forward P/E considering the results of Alford (1992).

Return on equity
Defined in line with our common definition. In addition to profit margin, this variable should provide a measure of profitability in relation to its capital requirements and is expected to be positively correlated with the forward P/E.

R&D expenses
Total research and development expenses scaled by sales as defined in Bhojraj & Lee (2002). These expenses tend to lower current profitability relative to future profitability. To the extent that this variable captures profitability beyond earnings growth, this variable is expected to be positively correlated with forward P/E (Bhojraj & Lee, 2002).[16]

Equity beta
The equity beta is added in addition to the previously presented variables used by Bhojraj & Lee (2002). The variable is expected to be negatively correlated with forecasted P/E.

[16]Missing for approximately 50% of firm observations, like Bhojraj & Lee (2002) missing values are replaced with 0 to not limit the dataset.

3.8 Non-linear multivariate method

The application of the non-linear multivariate method is based on Knudsen et al. (2017). As mentioned, the authors present a method that allows several fundamental factors to be considered without assuming a linear relationship between them and the dependent multiple. Instead, the SARD method ranks peers on the sum of absolute differences in fundamental selection variables between the target firm and potential peers. A lower absolute difference indicates a closer match to the dependent company and hence a more promising peer, hinged on the assumption that assets with similar characteristics should trade at a similar price (Fama & French, 1992). While either considering the total market, or an industry if industry classification is used, each firm is assigned a rank in terms of each variable considered. After choosing one firm on which to perform relative valuation, SARD scores are then calculated for all other firms. The SARD score formula is expressed as follows.

$$SARD_{ji} = |r_{xj} - r_{xi}| + |r_{yj} - r_{yi}| + \cdots + |r_{zj} - r_{zi}| \tag{12}$$

Where the SARD score denotes the total absolute difference between firm *i* and *j* in terms of ranks *r*, across, in theory, an unlimited amount of independent variables. In the main case, the six firms with the lowest SARD scores are then considered the most similar peers and forms the peer group.

We argue that an adjustment of this method is needed for comparability to previous research but also for theoretical logic. In the SARD method when ranking all firms in the considered selection prior to choosing the considered firm and calculating the SARD scores of other firms, Knudsen et al. (2017) disable the method to consider the closest peers in absolute terms across variables. In other words, in an example of only considering one variable and a considered firm ranked in the middle of the sample, the SARD method will choose the three peers below and above the considered firm in terms of ranking. This is not comparable with the non-linear univariate method of Alford (1992) and Cheng & McNamara (2000) that choose the six closest firms in absolute terms, which in theory could be the two firms below and four firms above the considered firm's ranking.

As a result, we use and present an adjusted method which we instead call the sum of absolute difference ranks (hereinafter SADR).[17] Illustrative in the name of the adjusted method, our approach simply changes the order of ranking and calculating differences.[18] In the SADR method, we instead choose the firm to be valued and the sample of other firms first, then we calculate the absolute difference in variables between this firm and all other firms, to finally rank them according to the absolute differences. This way, comparing the non-linear univariate method to SADR when only considering one variable, the two methods will be the same (see Appendix 2 for an illustrative example of peer selec-

[17] While acknowledging the potential confusion caused by the similarity of names, we did not want to disparage the contribution of Knudsen et al. (2017) by using a completely dissimilar name.

[18] The D changes place with the R in the abbreviation.

tion with SADR vs. SARD). As a sensitivity check, we will compare the prediction error of the best performing SADR variation to the SARD method to ensure the adjustment has not negatively affected the non-linear multivariate approach.

First, after selecting the firm to be valued and the sample, either market or industry, the absolute difference between the considered Firm *j* and all other firms *i* across all variables are calculated as follows.

$$Absolute\ differences_{j\,i} = |x_j - x_i| + |y_j - y_i| + \cdots + |z_j - z_i| \tag{13}$$

After ascendingly sorting differences, each firm is assigned a rank for each variable. The SADR score used to identify the six firms with the lowest total score is then calculated for each other firm *i* as follows.

$$SADR\ score_i = r_{x,i} + r_{y,i} + \cdots + r_{z,i} \tag{14}$$

3.8.1 Independent variables

In line with prior literature, Knudsen et al. (2017) propose profitability, growth, and risk as the three most important fundamental value drivers. To assure stability with the model presented by Knudsen et al. (2017), the following study will largely use the same fundamentally derived independent variables. Some variation will however be introduced, partly to reflect sole use of the P/E multiple as opposed to the original study, and partly to increase comparability with the other methods. The order of inclusion of the independent variables is mainly based on their presence and previous results in prior research, but also on the theoretical logic from our deduced forward P/E model.

Return on equity (ROE)
Defined in line with our common definition as the expected book return of equity. The definition differs from that of Knudsen et al. (2017), who uses reported earnings in the denominator (see discussion in Section 3.3).

Implied growth
Defined in line with our common definition of growth which somewhat differs from that of Knudsen et al. (2017) who are using a two period forecasted EPS growth scaled by one period forecasted EPS growth. To the extent possible, we thus use a longer period implied growth which we argue is more relevant when also using an forward ROE.

Net Debt over EBIT
The variable is defined as long- and short-term debt, adjusted for cash and short-term investments scaled by EBIT.[19] The definition does not differ from the authors and is used as a proxy for risk, motivated by Knudsen et al. (2017) as an often-recurring component in credit analysis.

[19]Earnings before interest and tax.

Size
In Knudsen et al. (2017) the proxy for size used is market capitalisation. Adhering to the principle of possible use in a private setting, while also acknowledging the questioning of relative valuation in a public setting by Penman (2013), we instead use total sales as a proxy for size. Since the most popular proxies for firm size is either total assets, total sales or market capitalisation, we acknowledge that total assets could have been an alternative proxy with similar results. However, Dang et al., (2018) find somewhat higher correlation between total sales and market capitalisation than for total assets.

Equity beta
Defined according to our common definition and introduced as a novel variable for this method as well, motivated by Damodaran (1994, 2002). Once again, we argue that the equity beta should serve as a proxy for risk as it in extension affects the cost of capital (see discussion in Section 3.3).

Operating profit margin
Uses our common definition of operating profit which is the same used by Knudsen et al. (2017) and that of Bhojraj & Lee (2002). Noteworthy, Knudsen et al. (2017) find marginal benefit of including the operating profit when regarding the P/E multiple, which is why we include it last in our analysis.

3.9 Empirical data

Data used in the study is derived from an intersection of the databases COMPUSTAT for historical financial values, CRSP for stock prices and IBES for analyst forecasts and beta values, over the period 2010 – 2020. The sample construction starts with the COMPUSTAT data, containing a total of 128,052 observations over the investigated period, whereas data from the other databases are subsequently merged in. Market data including stock prices and beta values from CRSP are collected three months after each observation's fiscal year end to assure that the annual report has been released and its information thus should be reflected in market prices. For the analyst forecasts, IBES compiles data for several occasions during the year given analysts revisions of outlooks. IBES data is hence collected for the period closest to that of when the CRSP data is retrieved. To reflect and adhere to practical conditions, IBES data is never collected a period later to that of when CRSP data is gathered, as this is not yet published information at the given time. Finally, after the databases have been merged, but before additional adjustments and trimmings, the dataset contains 48,395 number of observations.

3.9.1 Exclusions and adjustments

COMPUSTAT derived variables including revenue, earnings and book values are excluded should they be negative. Additionally, observations with broken fiscal years are excluded. This is motivated by the risk that market movements create variation between observations with broken and non-broken fiscal years respectively, which is not captured in the methods and therefore may impact the results. Also, using trailing quarterly results in line with Knudsen et al. (2017) could generate mismatch between earnings forecasts as these usually are given for the next fiscal year. Exclusion are instead in line with other previous research (see e.g. Alford, 1992; Lie & Lie, 2002; Other) and results in a drop of approximately 6,000 observations.

Furthermore, observations with share price less than $2 are eliminated, as in line with (e.g. Liu et al., 2002) to exclude outliers. Additionally, share prices, forecasted earnings and their resulting multiple, the P/E, are trimmed in the 1st and 99th percentile. Besides this, observations with missing values and duplicates are excluded from the sample. After the datasets have been merged and above eliminations have been made, there are a total of 19,291 firm observations for the period 2010 – 2020.

Finally, the independent variables used in the linear regression method, including Profit margin, Growth, Leverage, &D expenses, ROE and Beta are winsorised at the first and final percentile in pursuit of achieving a more robust regression.

3.9.2 Descriptive statistics

An overview of used data after above eliminations and adjustments are presented in Table 3.1. The median and interquartile range for select variables, including the forecasted P/E, for the years 2010–2020 are displayed to provide an understanding of the sample.

Additionally, given the method of subsequently decreasing industry detail to reach LNOF in both the GICS and SIC classification systems, Table 3.2 displays the number of observations for each hierarchical industry level.

Table 3.1: Descriptive statistics

Presents the median and interquartile range of selected variables end of March between 2010-2020. Forecasted P/E is the share price end of march each year over the one year forecasted earnings per share. ROE is the one period forecasted earnings divided by the last year-end book value of equity. Market cap is the share price multiplied by number of diluted shares outstanding. The mean of yearly medians is calculated as the arithmetic mean.

Year	Number of observations	**Forecasted P/E**		**ROE**		**Market Cap**		**Sales**	
		Median	Interquartile	Median	Interquartile	Median	Interquartile	Median	Interquartile
2010	1,668	10.9	12.5	11.0%	10.7%	1,182.7	3,479.1	712.5	2,815.4
2011	1,868	14.7	12.1	11.4%	12.1%	1,355.8	4,375.7	736.4	2,970.4
2012	1,849	15.3	10.5	11.7%	12.1%	1,421.3	4,239.5	819.5	3,216.4
2013	1,850	16.1	11.0	11.7%	11.6%	1,553.8	4,777.0	856.9	3,376.9
2014	1,874	16.3	10.6	11.8%	10.9%	1,858.7	5,714.2	888.2	3,510.0
2015	1,899	16.3	11.2	11.9%	12.6%	1,905.7	5,811.6	943.3	3,594.4
2016	1,741	16.4	10.5	11.9%	12.4%	1,980.0	5,625.2	962.1	3,366.4
2017	1,704	16.8	9.6	12.1%	11.7%	2,072.8	6,369.2	1,037.3	3,579.0
2018	1,678	17.7	9.77	12.3%	12.4%	2,257.0	7,106.8	1,199.5	3,760.8
2019	1,643	17.9	11.3	12.9%	11.9%	2,432.2	7,060.5	1,239.0	3,971.9
2020	1,517	18.6	11.2	13.1%	10.9%	2,539.5	5,717.2	1,331.8	4,336.3
Statistics of yearly medians									
Mean		16.1		0.12		1,869.1		975.1	
Standard deviation		2.1		0.01		448		206	

Table 3.2: Hierarchical levels of GICS and SIC

Presents the number of observations in the sample present in the different hierarchical levels for the industry classifications GICS and SIC.

	GICS		SIC	
Digit level	Number of observations	% of observations	Number of observations	% of observations
Level 1	29	0.2%	717	3.7%
Level 2	521	2.7%	3,439	17.9%
Level 3	1,923	10.0%	2,217	11.5%
Level 4	16,818	87.2%	12,885	66.9%

4. Results and analysis

The following section presents the results and analysis of our methods outlined in the preceeding section. The aim of the section is to pedagogically present systematic peer selection methods with increasing complexity of considerations used. Lastly, results are summarised and analysed with regards to time and for different industries.

The section is structured as follows. In Section 4.1, results from constructing peer groups only using industry belonging are presented and discussed. In Section 4.2, results from peer groups constructed using the non-linear univariate method are presented. In Section 4.3, results from the linear multivariate method are introduced. In Section 4.4, results from the non-linear multivariate method (SADR) is introduced, both without and with industry classification, also with TNIC. Finally, in Section 4.5 the variations in each method with the best results are compared both over time and across industries.

4.1 Peer group construction using industry classification

As mentioned, the first step of our analysis compares accuracy of the different industry classifications SIC, GICS and TNIC compared to the market average as a baseline. Resting on the notion that companies in the same industry tend to share similar economics and therefore also fundamental value drivers, results will indicate the information provided by industry classification on this when defining peers. Results presented will also define the industry classification further used in the rest of the analysis.

Presented in Table 4.1 is the APRE and PRE measurements when the average of the market or the different industry classifications are considered as the peer groups. As expected and seen in the Market column in Panel B, all investigated classification systems are statistically significantly better than using Market. This confirms prior results of industry classifications providing information about firm economics (e.g. Alford, 1992; Cheng McNamara). Market mean results in panel C further highlights the trade-off between precision and dispersion, considering a mean PRE very close to zero while simultaneously having the highest IQR and the highest mean APRE. It is therefore of importance to interpret all these measures combined.[20]

[20] Mean of medians, median and mean combined will further be referred to as precision.

Table 4.1: Prediction errors using industry classification

Presents prediction errors when using the Market and industries based on either of the three industry classifications SIC, GICS and TNIC as peer groups. SIC and GICS are applied to the whole dataset, whereas TNIC only allows application to a subset of the sample. As a sensitivity, GICS* is analysed using the TNIC data sample. The industry classifications SIC and GICS are at first used in the most detailed hierarchical level, to subsequently decrease in detail until at least six firms are present in each peer group. Panel A reports Absolute Prediction Errors (APRE) for arithmetic Mean of yearly medians, Median, arithmetic Mean and Interquartile range. Panel B reports Wilcoxon signed-rank test for pairwise APRE differences where "+" indicates that the selection method in the row is more accurate than in the column, "-" denotes the opposite. ***Denotes significance at the 1% level, **denotes significance at the 5% level, *denotes significance at the 10% level. Panel C represents Prediction Errors (PRE) for the same measurements as in Panel A.

		Industry Classification		**TNIC sample**	
	Market	SIC	GICS	TNIC	GICS*
Panel A: APRE per selection method					
Mean of medians	0.308	0.280	0.267	0.240	0.248
Median	0.305	0.279	0.264	0.237	0.248
Mean	0.422	0.405	0.384	0.341	0.369
Interquartile range	0.421	0.398	0.370	0.345	0.358
Panel B: Statistical test on difference in central tendencies using Wilcoxon test					
SIC	+***		-***	-***	-***
GICS	+***	+***		-*	+***
TNIC	+***	+***	+*		+***
GICS*	+***	+***	+***	-***	
Panel C: PRE per selection method					
Mean of medians	-0.102	-0.080	-0.060	-0.063	-0.049
Median	-0.101	-0.079	-0.058	-0.063	-0.055
Mean	0.001	0.017	0.025	-0.003	0.024
Interquartile range	0.564	0.531	0.515	0.467	0.483

As also expected considering results of prior research and its more novel adjustments to modern industries, GICS has a higher precision and lower dispersion than that of SIC. Bhojraj et al. (2003) further argues that the higher precision of GICS is partly derived from a better suitability with practical use. Finally, the higher percentage of observations in the most detailed level for GICS as opposed to SIC in Table 3.2 might also have explanatory value. Still, SIC has informational value compared to using the Market average. Following the discussion in Section 3.5, these results mean that GICS will be the industry classification used further on in the analysis.

Interestingly, as there also are no results form prior research, TNIC offers the highest precision and lowest dispersion compared to the other classifications. As mentioned, using TNIC reduces the dataset which is why the GICS classification is analysed using the same dataset, denoted as GICS*. When comparing TNIC with GICS*, their performance is more similar but TNIC is still better with statistical significance. This implies that product market does not only explain peer selection in practice, but also improves the accuracy of relative valuation, which to our knowledge has not been previously confirmed.

Given the difference in definitions between GICS and TNIC, as well as their individually high performance, a natural extension of analysis would be to investigate their combined informational value. However, given that GICS is static and TNIC dynamic on a year-by-year basis, we do not currently see a way to combine them in a practically relevant way. Interestingly, when looking at the intersection of the two, the amount of peers in the TNIC data that has the same 8- or 6-digit GICS as the considered firm are 49% and 58% respectively. The partial overlap of these two similarity measures therefore suggests a benefit of a combination, which we encourage to be further examined in future research.

In addition to industry classification, previous research suggest adding fundamental value drivers to peer selection increases predictability (e.g., Boatsman Baskin, 1981; Alford, 1992; Cheng McNamara, 2000), which is examined in the following section.

4.2 Peer selection using the non-linear univariate method

The following section evaluates the impact of introducing a fundamental value driver to the industry classification. In addition to GICS, the impact of introducing the hybrid between profitability and earnings growth, forward book ROE, is investigated below.

The PRE and APRE results for the non-linear univariate method are presented in Table 4.2. Market is still used as a baseline and Industry is using GICS from Section 4.1, while ROE and Industry & ROE have been added. For ROE, the six closest firms in terms of ROE from the whole market forms the peer group, while for Industry & ROE the six closest firms in terms of ROE within each industry forms the peer group.

Only selecting peers on ROE results in somewhat lower precision yet also somewhat lower dispersion than Market. While perhaps intuitive, when selecting fewer peers than the market average, ROE does not seem to provide enough information on surrounding factors affecting the market value of firms to identify peers.

When instead using ROE to select peers within an industry, there is a benefit in terms both precision and dispersion compared to using industry averages. This confirms prior results, mainly from Cheng & McNamara (2000), and suggest that industry classifications does not capture all information provided by fundamental factors. Considering the results of only using ROE, this means that while ROE is not comparatively priced across industries, ROE is affecting market value within industries.[21] Our results thus support the notion that a combination of business profile and fundamental factors should be used to increase the accuracy of relative valuation.

Considering these results, it is therefore worthwhile examining the proposed multivariate methods that, in addition to industry belonging, are able to consider similarity in several fundamental factors simultaneously. As mentioned, while doing this, it is interesting

[21] Since we are using book return, this is effectively confirming that there are different Q-values across industries, which is not news to valuation research, see e.g. Runsten (1998).

to examine if using several fundamental factors shows potential of replacing the information provided by industry classification. The following two sections will cover this analysis.

Table 4.2: Prediction errors for univariate non-linear approach

Presents prediction errors when using the Market, Industry using the GICS classification, the six closest peers in terms of ROE and the six closest peers in terms of ROE within the Industry, denoted as Industry & ROE. Panel A reports Absolute Prediction Errors (APRE) for arithmetic Mean of yearly medians, Median, arithmetic Mean, and Interquartile range. Panel B reports Wilcoxon signed-rank test for pairwise APRE differences where "+" indicates that the selection method in the row is more accurate than in the column, "-" denotes the opposite.***Denotes significance at the 1% level, **denotes significance at the 5% level, *denotes significance at the 10% level. Panel C represents Prediction Errors (PRE) for the same measurements as in Panel A.

	Market	Industry	ROE	Industry & ROE
Panel A: APRE per selection method				
Mean of medians	0.308	0.267	0.326	0.253
Median	0.305	0.264	0.322	0.252
Mean	0.422	0.384	0.436	0.368
Interquartile range	0.421	0.370	0.403	0.352
Panel B: Statistical test on difference in central tendencies using Wilcoxon test				
Industry	+***		+***	-***
ROE	-***	-***		-***
Industry & ROE	+***	+***	+***	
Panel C: PRE per selection method				
Mean of medians	-0.102	-0.060	-0.079	-0.042
Median	-0.101	-0.058	-0.078	-0.040
Mean	0.001	0.025	0.052	0.047
Interquartile range	0.564	0.515	0.624	0.498

4.3 Peer selection using the linear multivariate method

Following the limitation of using one fundamental value driver as selection criteria in the preceding method, the multivariate linear method provides an opportunity to regard several variables. The linear multivariate method used in this section is largely based on the method of Bhojraj & Lee (2002) and results in five different variations.

4.3.1 Estimation regressions

Before the results for the variations of the method are considered, the estimation regressions are presented. As mentioned, differences compared to Bhojraj & Lee (2002) are expected. This is partly due to the difference in time period, but mainly due to estimation regressions on P/E instead of the stock-based EV/S and P/B. The results from the regressions are presented in Table 4.3.

Table 4.3: Annual estimation regressions for warranted P/E

Presents the output for the annual regressions from the linear multivariate method using the following estimation regression.

$$P/E_{t+1} = \alpha_t + \sum_{x=1}^{7} \beta_{x,t} X_{x,t} + \epsilon_t$$

Where the dependent variable P/E is defined as price for each year the end of March divided by one year forecasted EPS. For each year, presented estimations are using data from the year previous to what is defined. The seven independent variables are defined as follows. P/E Ind. mean is the harmonic mean of P/E for the industry group with the step-wise industry definition. Adj. Profit Margin is the difference between a firm's operating profit margin and the industry operating profit margin. Adj. Growth is the difference in earnings forecast of the company being valued and the median earnings forecast for the same industry group. Leverage is long term-debt scaled by book value of common equity. RD expenses is total research and development expenses scaled by sales. ROE is one-year forecasted earnings scaled by end-period common equity. Equity Beta represents volatility compared with the associated equity index SP 500. Additionally, absolute values of t-statistics are provided in parenthesis below each yearly coefficient. The final row presents the regression run on the complete dataset, with associated Adjusted R-square.

Year	**Intercept**	**P/E Ind. mean**	**Adj. Profit Margin**	**Adj. Growth**	**Leverage**	**R&D Expense**	**ROE**	**Equity Beta**	**Adj. R2**	**# of obs.**
2010	13.939	0.656	-10.378	-5.441	0.309	41.082	-13.987	-1.976	13.04%	1,498
	(8.09)	(5.14)	(2.87)	(5.82)	(1.32)	(4.19)	(7.73)	(2.77)		
2011	13.419	0.764	-15.418	-5.741	0.285	28.046	-14.913	-0.729	11.24%	1,668
	(7.33)	(6.78)	(3.75)	(5.01)	(0.99)	(3.24)	(6.50)	(1.09)		
2012	10.385	0.956	-15.081	-1.222	1.420	44.294	-19.854	-1.509	14.73%	1,865
	(5.62)	(9.04)	(3.93)	(0.93)	(3.08)	(5.36)	(8.20)	(2.34)		
2013	8.279	0.973	-10.395	-2.262	1.195	33.896	-14.201	-1.626	18.63%	1,847
	(5.70)	(11.47)	(3.16)	(2.20)	(3.52)	(4.64)	(6.17)	(3.14)		
2014	8.268	0.987	-13.851	-2.836	1.332	36.425	-15.047	-1.619	20.31%	1,849
	(5.40)	(12.19)	(4.15)	(2.78)	(4.04)	(4.41)	(7.16)	(2.88)		
2015	9.227	0.918	-11.752	-0.882	1.226	41.622	-17.636	-1.315	21.03%	1,871
	(6.06)	(11.82)	(3.82)	(0.84)	(4.01)	(4.46)	(8.32)	(2.31)		
2016	4.826	1.153	-12.842	-3.151	1.058	34.797	-14.066	-0.540	18.49%	1,896
	(2.61)	(11.43)	(3.12)	(2.73)	(3.74)	(4.15)	(7.19)	(0.79)		
2017	5.126	1.078	-10.200	-1.540	0.646	35.660	-10.999	-1.037	24.58%	1,735
	(3.04)	(12.22)	(3.55)	(1.12)	(2.70)	(4.72)	(6.20)	(1.48)		
2018	3.380	1.105	-18.612	-3.505	0.682	37.042	-9.929	-0.603	29.07%	1,702
	(2.23)	(14.19)	(6.26)	(4.72)	(3.00)	(4.77)	(5.89)	(0.93)		
2019	5.652	1.022	-11.608	-1.917	0.520	47.302	-8.269	-1.840	28.55%	1,676
	(4.04)	(13.54)	(4.42)	(2.58)	(2.21)	(5.97)	(4.80)	(2.80)		
2020	4.880	1.049	-13.945	-0.193	0.416	53.484	-9.381	-1.035	33.73%	1,641
	(3.88)	(15.81)	(4.41)	(0.14)	(1.38)	(4.81)	(3.79)	(1.54)		
Full Set	7.095	1.019	-13.049	-2.174	0.835	41.715	-12.751	-1.409	23.23%	20,589
	(16.80)	(45.14	(13.46)	(8.18)	(10.14)	(16.61)	(21.54)	(7.56)		

As expected, we find deviations from that of Bhjoraj & Lee (2002), but also from our own expectations. While Bhojraj & Lee (2002) received criticism for unstable estimations (Sloan, 2002), ours are arguably worse. First, there are differences in the expected signs for the coefficients of the independent variables Adjusted Profit Margin, Adjusted Growth and ROE. Second, the magnitude of the coefficients varies heavily over years. Third, the adjusted R-squared is lower on both annual basis and for the full dataset. A concurrent problem with such varying estimations combined with estimations with a one year time lag, is that there should be low probability for estimations to be correct.

While Damodaran (1994, 2002) shows that regressions using P/E as dependent variable is less stable compared to e.g., EV/S or P/B, the volatility in our regression poses a problem given the methods reliance on coefficients. Despite the changing signs may partly be explained by an extent of multicollinearity (see Appendix 1), this is not believed to fully explain the high volatility. While we acknowledge that a much wider range of independent variables could be tested in different combinations, that the regressions can be adjusted by eliminating dependent variables on a year-by-year basis, or that regressions could be estimated on a longer historical periods, we have tested the method following our guiding principles of maintaining the integrity of the model and practical relevance relating to parsimoniousness.

Yet, even if the analysis of the estimated coefficients and regressions results in a negative prediction of the method to produce accurate valuation estimates, it could be that the method still captures peer similarity, perhaps especially if the warranted multiples are used for ranking. In the next section, we thus present our valuation estimation results from the linear method.

4.3.2 Prediction errors for the linear multivariate method

The results of PRE and APRE for the linear multivariate method are presented in Table 4.4. In addition to the Market and the previously presented lowest prediction error using Industry & ROE, we present five variations using the linear multivariate method. These include using the warranted multiples for each observation directly (WM), using the warranted multiples of six closest peers in regards to warranted multiples (Direct WM), using the market multiples of the six closest peers in regards to warranted multiples (Indirect WM), and Direct WM or Indirect WM with selecting peers from the same industry.

Unfortunately, the method does not show results being able to capture peer similarity following the previously mentioned estimation problems. Looking at Panel B in the Market column, only the two indirect variations show results with better performance than the market average. While combining the indirect valuation with industry classification shows the best result, using the more parsimonious Industry & ROE still outperforms any variation. From a practical perspective, we find no support using the linear method.

Interestingly, most of our prediction errors for the linear method are lower than those of Bhojraj & Lee (2002) which also finds indirect estimates performs best with the lowest overall mean APRE at 0.51 using the P/B multiple. This relates to our previous discussion

Table 4.4: Prediction errors for linear multivariate method

Presents prediction errors for the Market, the best performing non-linear variation and the five variations of the linear regression method. WM represents Warranted Multiple (WM) and is derived from the "deserved" multiple following the regression. Direct WM represents WM used as ranking to construct peer groups based on closest proximity. Indirect WM represents WM used as ranking to construct peer groups, but observed P/E multiples are applied instead of WM. Direct WM + Industry and Indirect WM + Industry represents the two methods applied in the same industry. Panel A reports Absolute Prediction Errors (APRE) for arithmetic Mean of yearly medians, Median, arithmetic Mean and Interquartile range. Panel B reports Wilcoxon signed-rank test for pairwise APRE differences where "+" indicates that the selection method in the row is more accurate than in the column, "-" denotes the opposite. ***Denotes significance at the 1% level, **denotes significance at the 5% level, *denotes significance at the 10% level. Panel C represents Prediction Errors (PRE) for the same measurements as in Panel A.

			Linear Regression Approach				
	Market	Industry & ROE	WM	Direct WM	Indirect WM	Direct WM + Industry	Indirect WM + Industry
Panel A: APE per selection method							
Mean of medians	0.308	0.253	0.382	0.382	0.303	0.382	0.271
Median	0.305	0.252	0.377	0.377	0.301	0.376	0.271
Mean	0.422	0.368	0.517	0.517	0.408	0.517	0.395
Interquartile range	0.421	0.352	0.464	0.464	0.395	0.465	0.380
Panel B: Statistical test on difference in central tendencies using Wilcoxon test							
Industry & ROE	+***		+***	+***	+***	+***	+***
WM	-***	-***		-	-***	-***	-***
Direct WM	-***	-***	+		-***	-***	-***
Indirect WM	+***	-***	+***	+***		+***	-***
Direct WM + Ind.	-***	-***	+***	+***	-***		-***
Indirect WM + Ind.	+***	-***	+***	+***	+***	+***	
Panel C: PRE per selection method							
Mean of medians	-0.102	-0.042	0.228	0.228	-0.052	0.241	-0.037
Median	-0.101	-0.040	0.221	0.222	-0.051	0.233	-0.035
Mean	0.001	0.047	0.306	0.305	0.047	0.320	0.061
Interquartile range	0.564	0.498	0.698	0.698	0.592	0.683	0.532

on choice of valuation multiples, whereas despite the linear method does not seem optimal for the P/E multiple, our results support its previously shown highest accuracy.

Essentially, our results support the points of criticism towards Bhojraj & Lee (2002) presented by Sloan (2002). These include assuming a linear relationship between variables deduced from a non-linear theoretical model, that independent variables seems to have been chosen 'ad hoc', and the methodological paradox that indirect estimates performed better than the direct warranted multiples themselves.

In light of this criticism supported by our results, it seems more relevant to consider a method allowing for several fundamental factors without assuming a linear relationship between the P/E multiple and the fundamental factors. This will be examined in the section that follows.

4.4 Peer selection using the non-linear multivariate method

The SADR approach is the final method examined which to a large extent is based on Knudsen et al.'s (2017) SARD method. It enables the mentioned use of several fundamental value drivers without assuming a linear relation. The method will be tested both on stand-alone basis, in combination with industry belonging and finally in combination with the TNIC industry classification, utilising its 'score' variable for ranking.

4.4.1 Prediction errors for SADR on stand-alone basis

The results from the SADR approach on stand-alone basis without combining with industry classification are presented in Table 4.5. The observant reader will note that the results presented in the first column using only ROE, is the same as the column for the non-linear univariate approach in Table 4.2. In large, the step-wise introduction of additional fundamental selection variables continuously decreases precision and dispersion. Furthermore, the APRE increases for the sixth and final step of the model when EBIT margin is introduced, which we attribute to limited impact of operating profit on P/E, also found by Knudsen et al. (2017).

Our results thus confirm the findings of Knudsen et al. (2017) that the non-linear multivariate method is a viable method to consider several fundamental factors simultaneously when considering valuation accuracy. However, even if our best performing variation of SADR uses five fundamental factors to identify peers, it does not outperform either using Industry on its own or Industry & ROE. Therefore, our results contradict those of Knudsen et al. (2017) on the suggested possibility for a multivariate model to compensate for the information provided by industry classifications.[22]

While not entirely certain on the source of this difference, it is not limited to but could include; the definition of industry classification, using the SADR instead of the SARD method, definitions of independent variables or time period and data.[23] On this note, we thus encourage future research to trial this difference.

Considering our more promising results of combining industry classification with one fundamental factor, and the fact that Knudsen et al. (2017) also found improved results combining SARD with industry classification, the next section will examine using the combination of SADR and industry classification.

[22]Corresponding to results for the non-linear univariate method, our results again suggests that fundamental factors are not comparatively priced across industries.

[23]Comparing SADR and SARD methods as a sensitivity test will be included in the following section.

Table 4.5: Prediction errors for Sum of Absolute Differences Ranks

Presents prediction errors for the Market, the best variations from the non-linear univariate method, the linear multivariate method and the five additional variations from the Sum of Absolute Difference Ranks (SADR). Each column for the SADR approach represents the subsequent introduction of an additional fundamental value driver. Panel A reports Absolute Prediction Errors (APRE) for arithmetic Mean of yearly medians, Median, arithmetic Mean, and Interquartile range. Panel B reports Wilcoxon signed-rank test for pairwise APRE differences where "+" indicates that the selection method in the row is more accurate than in the column, "-" denotes the opposite. ***Denotes significance at the 1% level, **denotes significance at the 5% level, *denotes significance at the 10% level. Panel C represents Prediction Errors (PRE) for the same measurements as in Panel A.

					SADR approach					
	Market	Industry	Industry & ROE	Indirect WM + Industry	ROE	ROE + Growth	ROE + Growth + Size	ROE + Growth + Size + ND / EBIT	ROE + Growth + Size + ND/EBIT + Beta	ROE + Growth + Size + ND/EBIT + Beta + EBIT%
Panel A: APRE per selection method										
Mean of medians	0.308	0.267	0.253	0.271	0.326	0.312	0.302	0.291	0.279	0.282
Median	0.305	0.264	0.252	0.271	0.322	0.310	0.299	0.289	0.278	0.281
Mean	0.422	0.384	0.368	0.395	0.436	0.422	0.416	0.400	0.392	0.397
Interquartile range	0.421	0.370	0.352	0.380	0.403	0.393	0.391	0.383	0.381	0.383
Panel B: Statistical test on difference in central tendencies using Wilcoxon test										
Industry	+***		-***	+**	+***	+***	+***	+***	+**	+***
Industry & ROE	+***	+***		+***	+***	+***	+***	+***	+***	+***
IWM industry	+***	-**	-***		+***	+***	+***	+***	+	+***
SADR 1	-***	-***	-***	-***		-***	-***	-***	-***	-***
SADR 2	-*	-***	-***	-***	+***		-***	-***	-***	-***
SADR 3	+***	-***	-***	-***	+***	+***		-***	-***	-***
SADR 4	+***	-***	-***	-***	+***	+***	+***		-***	-***
SADR 5	+***	-**	-***	-	+***	+***	+***	+***		+***
SADR 6	+***	-***	-***	-***	+***	+***	+***	+***	-***	
Panel C: PRE per selection method										
Mean of medians	-0.102	-0.060	-0.042	-0.037	-0.079	-0.078	-0.067	-0.059	-0.048	-0.050
Median	-0.101	-0.058	-0.040	-0.035	-0.078	-0.076	-0.065	-0.057	-0.046	-0.048
Mean	0.001	0.025	0.047	0.061	0.052	0.049	0.052	0.050	0.053	0.052
Interquartile range	0.564	0.515	0.498	0.532	0.624	0.593	0.583	0.562	0.543	0.553

4.4.2 Prediction errors of SADR and industry classification

The results from SADR combined with industry classification are presented in Table 4.7. Like in the previous section, the first variation of SADR only using ROE is the same as Industry & ROE as results from the non-linear univariate method.[24]

As expected, using more than one fundamental factor combined with industry classification continuously increases precision at a statistically significant level. Once again, using five fundamental factors in the SADR method yields the highest prediction and lowest dispersion, if also considering PRE. However, in comparison to the industry agnostic SADR and to the results of Knudsen et al. (2017), the marginal benefit of each additional fundamental factor is lower. Even though there is a benefit, this suggests that industry already captures parts of the information provided by the additional fundamental factors. From a parsimonious perspective, one could argue that Industry & ROE performs at a relatively high level.

As an end note, to assure the validity of our adjustments to the non-linear multivariate method, we perform a sensitivity check comparing SADR to SARD for the best performing variation of SADR, combining industry with five fundamental factors. Results show that while comparable, SADR performs marginally better both in regards to precision and dispersion, motivating our adjustment.[25]

4.4.3 Prediction errors of SADR and TNIC

Considering the promising and unexplored results of TNIC as an industry classification, the performance of the non-linear multivariate method and the existence of the 'score' variable in the TNIC data, it is of interest to examine prediction errors of a combination of SADR and TNIC. Since the TNIC score variable is a numerical measure of product similarity, it can be added as one of the independent variables used for ranking with SADR. In Table 4.7, the resulting PRE and APRE is presented when considering the additional six variations of SADR using the TNIC score as the first independent variable while excluding the non-significant operating margin.

Surprisingly, in the process of choosing the six most similar firms with regards to the TNIC score, the estimation accuracy decreases compared to using the entire TNIC peer group. In an untabulated sensitivity check increasing the number of peer firms to 12, the performance is better than using six peers, but still worse than the entire peer average. Also, the same results are found when excluding the TNIC score and only considering ROE. With a degree of speculation, we propose that these dynamics suggests the functioning of the entire TNIC peer group as a synthetic peer. Instead of a ranking of the most similar peers, if different peers captures different aspects of similarity, it would motivate these results. Examining the prospects of synthetic peers in relative valuation could thus be something for future research to further study.

[24] Which would not have been the case if using the SARD approach instead of SADR.

[25] The SARD method's results of APRE are a median of 0.248, mean of 0.365 and IQR of 0.354.

Table 4.6: Prediction errors for industry Sum of Absolute Differences Ranks

Presents prediction errors for the Market, the best variations from the non-linear univariate method, the linear multivariate method and the five additional variations from the Sum of Absolute Difference Ranks (SADR) applied after filtering for GICS industry classification. Each column for the SADR approach represents the subsequent introduction of an additional fundamental value driver. Panel A reports Absolute Prediction Errors (APRE) for arithmetic Mean of yearly medians, Median, arithmetic Mean and Interquartile range. Panel B reports Wilcoxon signed-rank test for pairwise APRE differences where "+" indicates that the selection method in the row is more accurate than in the column, "-" denotes the opposite. ***Denotes significance at the 1% level, **denotes significance at the 5% level, *denotes significance at the 10% level. Panel C represents Prediction Errors (PRE) for the same measurements as in Panel A.

					SADR Matching Same Industry					
	Market	Industry	Industry & ROE	Indirect WM + Industry	ROE	ROE + Growth	ROE + Growth + Size	ROE + Growth + Size + ND/EBIT	ROE + Growth + Size + ND/EBIT + Beta	ROE + Growth + Size + ND/EBIT + Beta + EBIT%
Panel A: APRE per selection method										
Mean of medians	0.308	0.267	0.253	0.271	0.253	0.250	0.246	0.244	0.241	0.245
Median	0.305	0.264	0.252	0.271	0.252	0.248	0.244	0.243	0.239	0.243
Mean	0.422	0.384	0.368	0.395	0.368	0.367	0.369	0.364	0.363	0.369
Interquartile range	0.421	0.370	0.352	0.380	0.352	0.353	0.358	0.353	0.353	0.355
Panel B: Statistical test on difference in central tendencies using Wilcoxon test										
Industry	+***		-***	+**	-***	-***	-***	-***	-***	-***
Industry & ROE	+***	+***		+***		-**	-**	-***	-***	-***
IWM industry	+***	-**	-***		-***	-***	-***	-***	-***	-***
SADR 1	+***	+***	-***	+***		-**	+**	-***	-***	-***
SADR 2	+***	+***	+**	+***	+**		-***	-***	-***	-***
SADR 3	+***	+***	+**	+***	+**	+***		-*	-***	-
SADR 4	+***	+***	+***	+***	+***	+***	+*		-**	-
SADR 5	+***	+***	+***	+***	+***	+***	+***	+***		+***
SADR 6	+***	+***	+***	+***	+***	+***	+	+	-***	
Panel C: PRE per selection method										
Mean of medians	-0.102	-0.060	-0.042	-0.037	-0.042	-0.044	-0.037	-0.035	-0.032	-0.032
Median	-0.101	-0.058	-0.040	-0.035	-0.040	-0.042	-0.034	-0.034	-0.030	-0.030
Mean	0.001	0.025	0.047	0.061	0.047	0.043	0.051	0.047	0.051	0.054
Interquartile range	0.564	0.515	0.498	0.532	0.498	0.486	0.481	0.475	0.472	0.483

Table 4.7: Prediction errors for TNIC Sum of Absolute Differences Ranks

Presents prediction for the Market* average applied on the TNIC dataset, GICS* applied on the TNIC dataset and TNIC from the Sum of Absolute Difference Ranks (SADR) applied after filtering for TNIC industry classification. Each column for the SADR approach represents the subsequent introduction of and additional fundamental value driver. Panel A reports Absolute Prediction Errors (APRE) for arithmetic Mean of yearly medians, arithmetic Mean, Median, and Interquartile range. Panel B reports Wilcoxon signed-rank test for pairwise APRE differences where "+" indicates that the selection method in the row is more accurate than in the column, "-" denotes the opposite. ***Denotes significance at the 1% level, **denotes significance at the 5% level, *denotes significance at the 10% level. Panel C represents Prediction Errors (PRE) for the same measurements as in Panel A.

				SADR Matching using TNIC					
	Market*	GICS*	TNIC	TNIC Score	TNIC Score + ROE	TNIC Score + ROE + Growth	TNIC Score + ROE + Growth +Size	TNIC Score + ROE + Growth +Size + ND/EBIT	TNIC Score + ROE + Growth +Size + ND/EBIT + Beta
Panel A: APRE per selection method									
Mean of medians	0.296	0.248	0.240	0.285	0.288	0.276	0.273	0.270	0.267
Median	0.285	0.248	0.237	0.282	0.288	0.272	0.270	0.267	0.265
Mean	0.410	0.369	0.341	0.411	0.433	0.416	0.422	0.408	0.414
Interquartile range	0.419	0.358	0.345	0.383	0.396	0.395	0.390	0.385	0.385
Panel B: Statistical test on difference in central tendencies using Wilcoxon test									
GICS*	+***		-***	+***	+***	+***	+***	+***	+***
TNIC	+***	+***		+***	+***	+***	+***	+***	+***
SADR 1	+***	-***	-***		+	-***	-***	-***	-***
SADR 2	+***	-***	-***	-		-***	-***	-***	-***
SADR 3	+***	-***	-***	+***	+***		-*	-**	-*
SADR 4	+***	-***	-***	+***	+***	+*		-	-
SADR 5	+***	-***	-***	+***	+***	+**	+		+
SADR 6	+***	-***	-***	+***	+***	+*	+	-	
Panel C: PRE per selection method									
Mean of medians	-0.081	-0.049	-0.063	-0.050	-0.032	-0.032	-0.021	-0.017	-0.020
Median	-0.081	-0.505	-0.063	-0.050	-0.033	-0.031	-0.023	-0.016	-0.020
Mean	0.001	0.024	-0.003	0.053	0.097	0.084	0.102	0.091	0.095
Interquartile range	0.561	0.483	0.467	0.555	0.569	0.538	0.539	0.532	0.527

To summarise, selecting peers on product similarity seems to increase the performance of valuation estimates while the dynamics of its use is different from previously studied business profile factors such as industry. From a practical perspective, it could be problematic that large peer groups are required while the data availability is also lower than for traditional industry classifications.

In the following section, the methods investigated above will be further dissembled to understand prediction error distributions across methods, years, and industries.

4.5 Aggregated results and analysis

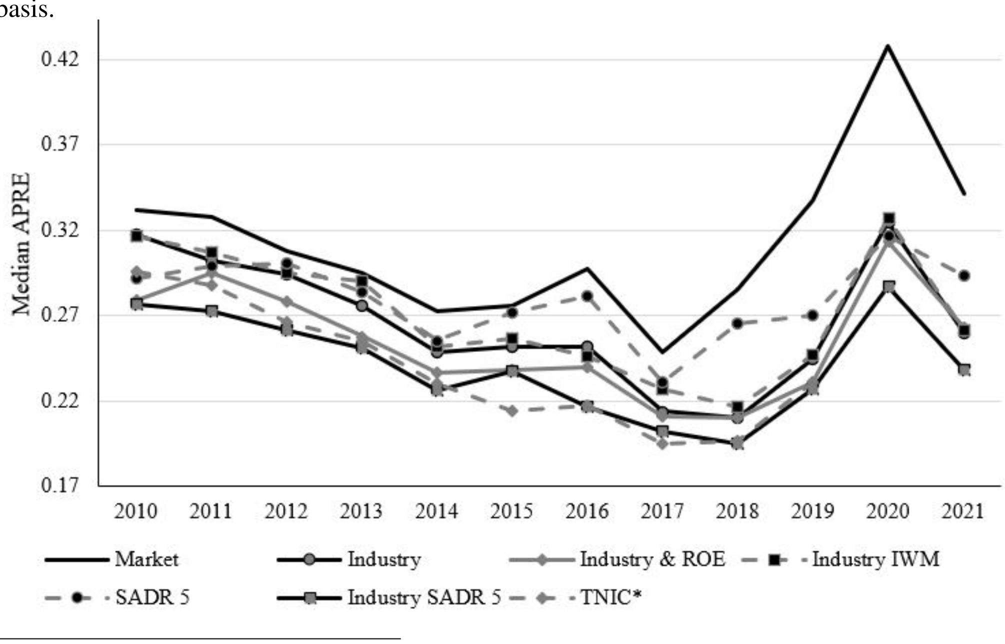
This section will finalise our analysis by presenting results from different methods on an aggregate level while also providing discussions on the aspects of time periods, different industries and the combined assessment of both precision and dispersion. The section will also aim to provide practical conclusions for the use of relative valuation.

4.5.1 Analysis across time

Figure 1 shows median APRE for the variations with the highest precision, spanning over the investigated time-period. Additionally, the year 2021 is included to facilitate discussion on external economic shocks.[26]

Figure 1: APRE prediction error over investigated period

Presents the development between 2010 – 2021 for the best variations of the non-linear univariate method, the linear multivariate method and the three SADR variations. Additionally, the industry classification GICS and the Market are included. TNIC* is only applied on the TNIC dataset, dissimilar to the other variations. The yearly prediction errors are measured on median Absolute Prediction Error basis.



[26]2021 displays recovery from impact of Covid-19, where S&P was up appx. 58% in march 2021, as compared to 2020.

The illustration confirms our overall results to also be representative over time. The Market consistently has the highest median APRE with the examined methods having a higher precision except for SADR 5. Furthermore, given its parsimoniousness, we note the relatively high accuracy of Industry & ROE.

Relating to the possible disadvantage of depending on market values, increases in median APRE during 2020 illustrates reduced accuracy during the external shock caused by Covid-19. Similar results was found by Knudsen et al. (2017) during the dot-com crisis. As can be seen, all prediction errors increases in line with the Market average, suggesting that despite historic low prediction errors, all methods are dependent on the market being stable. Hence, market risk will persist to have an appreciable impact on precision.

4.5.2 Analysis across industries

Following the discussion on industry belonging absorbing informational value provided by fundamental value drivers, it is of interest to investigate results across industries. Table 4.8 illustrates median APRE different industries based on GICS codes.

Table 4.8: Industry impact on prediction errors

Presents median APRE for the best variations of the non-linear univariate method, the linear multivariate method and the three SADR variations. TNIC* is only applied on the TNIC dataset, dissimilar to the other variations. Additionally, the Market is included. The industry sectors are derived from the GICS classifications system using the least granular hierarchical level.

GICS Sector	Sector	Total	Industry ROE	Industry IWM	Industry SADR 5	TNIC*
10	Energy	0.362	0.321	0.344	0.301	0.374
15	Materials	0.275	0.268	0.284	0.265	0.297
20	Industrials	0.263	0.258	0.266	0.244	0.215
25	Consumer Discretionary	0.307	0.292	0.300	0.282	0.242
30	Consumer Staples	0.290	0.270	0.275	0.257	0.221
35	Healthcare	0.315	0.290	0.296	0.256	0.282
40	Financials	0.257	0.183	0.211	0.172	0.235
45	Information Technology	0.345	0.335	0.342	0.318	0.335
50	Communication Services	0.402	0.382	0.395	0.373	0.392
55	Utilities	0.233	0.140	0.156	0.120	0.144
60	Real Estate	0.615	0.308	0.320	0.307	0.371

Comparing industry belonging to the different peer selection methods, we argue that fluctuations in prediction errors are driven primarily by the former. As can be seen by the results, prediction errors across methods are relatively similar for each industry. Comparing industries however, we argue that the lowest prediction errors are found in industries with reasonably lower dispersion in the nature of included firms. For example, the financial industry is characterised by heavy regulation while Utilities have low product differentiation, leading to relatively fewer differences in firm economics. On the contrary, for industries with reasonably high firm differentiation such as Communication services or Information technology, industry classifications fails to capture firm differences of certain business profile factors. An example of this is the Energy industry with higher prediction errors but also current ESG pressure. If assuming that different sustainability strategies across firms yields different market values, this would result in higher prediction errors since none of the methods consider similarity in ESG strategy.

Further interesting, when analysing across industries, results illustrates differences between GICS and TNIC. Looking at industries where TNIC* has lower prediction errors than GICS, these are Industrials, Consumer discretionary and Consumer staples. The common factor across these industries are that they are by nature product heavy, indicating that TNIC is superior compared to GICS for certain industries, while GICS is superior for others.

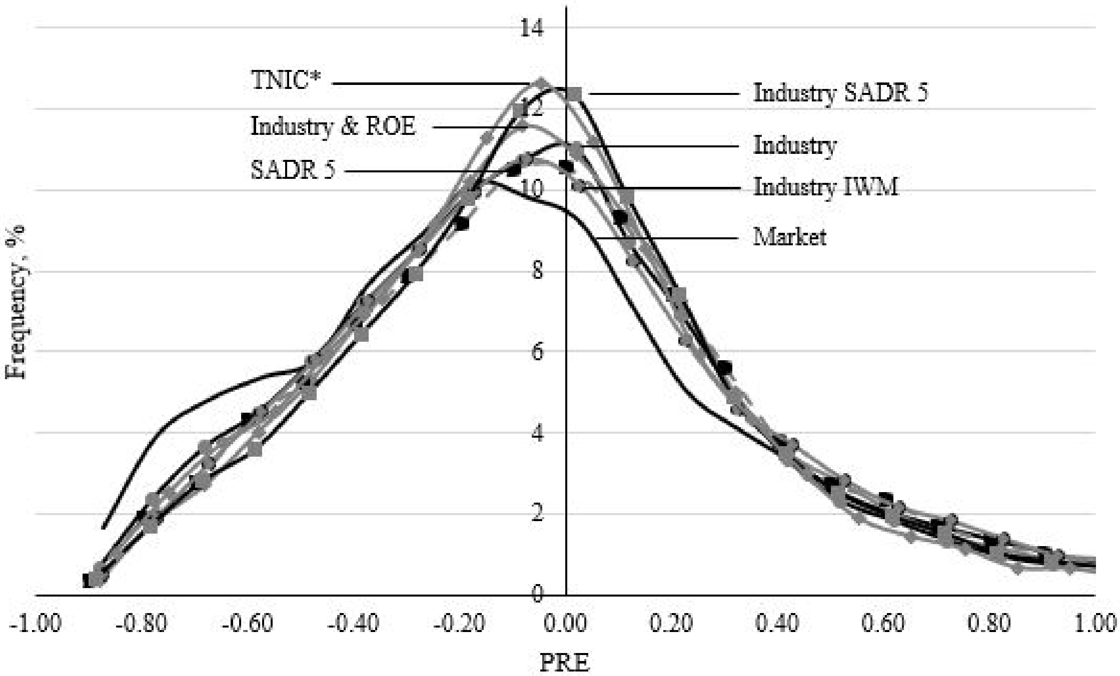
In essence, two conclusions adhere from this analysis. First, within industry firm differences seem to impact valuation accuracy since traditional industry classifications fails to capture all relevant business factors. Secondly, as a result, the improvement of relative valuation hinges on the ability of a combined use of different business profile factors, either simultaneously or adjusted after a certain industry. This is however challenging to consider if using current systematic peer selection methods, which we argue is a reasonable explanation of the the subjectively based peer selection within practice.

4.5.3 Analysis of overall accuracy

In Figure 2, distribution errors for the best variations in each method is presented.

Figure 2: Pooled distribution of prediction errors

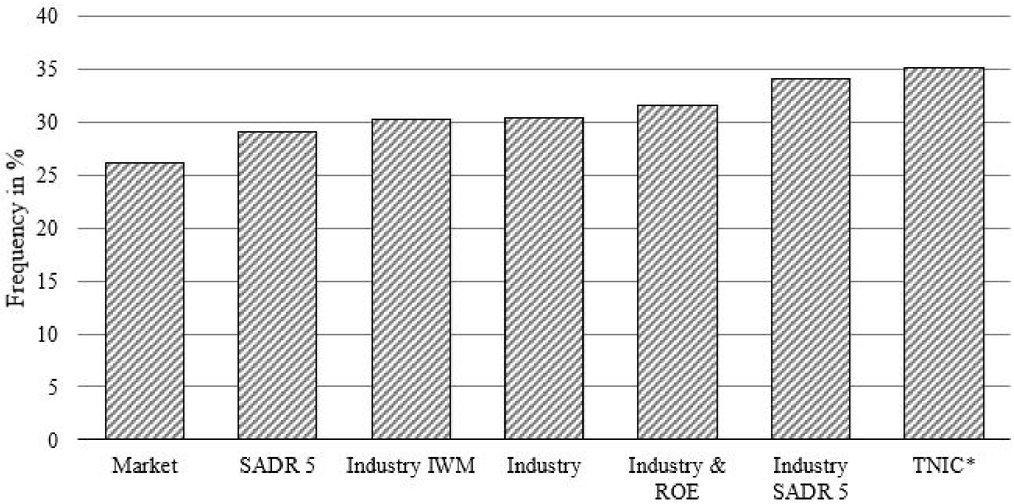
Presents pooled distribution of pricing errors for the best variations of the non-linear univariate method, the linear multivariate method and the three SADR variations. TNIC* is only applied on the TNIC dataset, dissimilar to the other variations. Additionally, the Market is included. The chart is derived from a histogram with columns of width -0.1.



When illustrating the distribution of PRE for all methods examined, it is more intuitive to assess the overall performance of a method considering both precision and dispersion. As can be seen in Figure 2, our previously presented results are confirmed with the Market average performing at the lowest level with added benefit from the different peer selection methods. Overall, all methods using industry classifications performs at the highest level, except for the linear multivariate method. Considering that stock prices are bound by zero, a certain skewness in the distribution can also be seen. When instead only considering the dispersion, our results are further confirmed as can be seen in Figure 3. Here, all methods using industry classification provides the best results. Noteworthy in both figures, is that TNIC results are using a subset of data which limits direct comparability to the other methods.

Figure 3: Frequency of observations within 15% prediction error

Presents the frequency of observations that are within a 15% prediction error range on both sides from zero. The best variations of the non-linear univariate method, the linear multivariate approach and the three SADR variations are included. TNIC* is only applied on the TNIC dataset, dissimilar to the other variations. Additionally, the the ma and the industry classification GICS are included.



5. Summary and conclusion

By using U.S. data while focusing on the forward P/E valuation multiple, we have analysed the pricing accuracy of different systematic peer selection methods presented in prior research. As peer selection to a large extent is performed on a subjective basis within practice, we argue that performing an analysis of systematic methods provides insights to the most accurate use of relative valuation. Furthermore, we contribute to an area of research with mixed results concerning the value provided by using industry classification and its relation to the use of fundamental factors to assess peer similarity.

Our results confirm the view that using industry classification (GICS) or product similarity (TNIC), significantly contributes to the accuracy of relative valuation. By using the TNIC classification, we present novel results showing that product similarity not only explains peer selection in practice, but also increases pricing accuracy. In regards to these results however, we acknowledge TNIC's lower practical relevance due to limited data availability and different peer selection dynamics. We also find that across industries, results indicate that accuracy increases for industries with lower differentiation between firms due to f.e. regulation or low differentiation of the product market. Conversely, accuracy decreases for industries with possible firm differences due to business profile factors not captured by current business profile classifications, such as ESG. On this note, we find that the performance of GICS and TNIC is seemingly contingent on industry, where their ability to distinguish firm differences depends on the nature of the industry. Considering these results, we thus encourage future research to further develop and examine methods on how to combine and expand business profile factors considered in peer selection. Also, considering the dynamics of the TNIC peer group, we identify the interesting opportunity to examine constructing synthetic peers for relative valuation.

When considering using fundamental factors to assess peer similarity, our results do not support the possibility for a multivariate method to replace the information provided by business profile similarity. We argue, which in our view is aligned with previous accounting research such as Runsten (1998), that our results confirms fundamental factors to be priced differently across industries.[27] This means that in theory, methods only considering fundamental factors should not be able to replace business profile similarity either.

When combining fundamental factors with industry classification however, we find support of non-linear methods to increase valuation accuracy. In line with the criticism from Sloan (2002), our results do not support the use of the linear multivariate method. Instead, which is more congruent with valuation theory, using a method that does not assume a relationship between variables, improves accuracy with each additional fundamental variable. With an adjustment to their method, we thus confirm the results of Knudsen et al. (2017) while suggesting that the marginal benefit compared to using industry is lower. Even if the non-linear method thus proves the ability to consider several factors and is relatively easy to use in practice (See Appendix 2), one could question its parsimoniousness compared to only considering ROE. The practical implications of our results are however that several fundamental factors indeed should be considered for peer selection.

Lastly, our results supports the previously indicated higher accuracy of relative valuation compared to current market prices, where the most theoretically congruent use is for private transactions with a short-term perspective. However, as our data include the external shock of Covid-19, our results also illustrate relative valuation's sometimes disadvantageous reliance on market prices, which is especially unreliable in uncertain markets.

As an ending note, since our results indicate the current inability for systematic methods to consider the several of the important business profile factors, we acknowledge the difficulty of persuading a subjective proponent of relative valuation to use systematic methods. However, we argue that our results provides insights to the practical use of relative valuation while also expanding an arguably under-served area of valuation research and pointing future research in new directions.

[27] Since accounting fundamental factors are used in these methods, the market price differences are reasonably a result of different accounting principles and business nature across industries.

6. References

6.1 Published articles

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

Anesten, S., Möller, N., Skogsvik, K., Skogsvik, S. (2020). The Pricing Accuracy of Alternative Equity Valuation Models: Scandinavian Evidence. Journal of International Financial Management Accounting, 31 (1): 5-34.

Asquith, P., Mikhail, M. B., Au, A. S. (2005). Information content of equity analyst reports. Journal of Financial Economics, 75, 245-282.

Baker, M.P., Ruback, R.S. (1999). Estimating Industry Multiples. Harvard Working Paper

Bhojraj, S., Lee, C.M. (2002). Why is My Peer? A Valuation-Based Approach to the Selection of Comparable Firms. Journal of Accounting Research, 40 (2): 407-439.

Bhojraj, S., Lee, C.M., Oler, D.K. (2003). Wha´s My Line? A comparison of Industry Classification Schemes for Capital Market Research. Journal of Accounting Research, 41(5): 745-774

Block, S.B. (1999). A Study of Financial Analysts: Practice and Theory. Journal of Analysts Journal, 55 (4): 89-95

Boatsman, J.R., Baskin, E.F. (1981). Asset Valuation with Incomplete Markets. The Accounting Review, 56 (1): 38-56.

Cheng, C.A., McNamara, R. (2000). The Valuation Accuracy of the Price-Earnings and Price-Book Benchmark Valuation Methods. Review of Quantitative Finance and Accounting, 15 (4): 349-370.

Chang, W-J., Landsman, W. R., Monahan, S. J. (2012). Selecting an Accounting-Based Valuation Model. Working paper (January), INSEAD/UNC Chapel Hill, VA.

Courteau, L., Kao, J. L., Richardson, G. D. (2001). Equity valuation employing the ideal versus ad hoc terminal value expressions. Contemporary Accounting Research, 18: 625–661.

Dang, C., Li, Z.F., Yang,C. (2018) Measuring Firm Size in Empirical Corporate Finance. Journal of Banking and Finance, 86: 159-176.

Demirakos, E.G., Strong, N.C., Walker, M. (2004). What Valuation Models Do Analysts Use?. Accounting Horizons, 18 (4): 221-240.

Dittmann, I., Maug, E.G. (2006). Biases and Error Measures: How to Compare Valuation Methods. Mannheim Finance Working Paper, 07.

Eaton, G.W., Guo, F., Liu, T., Officer, M.S. (2021). Peer Selection and Valuation in Mergers and Acquisitions. Journal of Financial Economics, Forthcoming.

Faber N.M. (1999). A Closer Look at the Bias–Variance Trade-Off in Multivariate Calibration. Journal of Chemometrics, 13 (1): 185-192.

Fama, E.F., French, K.R. (1997). Industry costs of equity. Journal of Finanical Economics, 43: 153-193

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

Francis, J., Olsson, P., Oswald, D.R. (2000). Comparing the Accuracy and Explainability of Dividend, Free Cash Flow, and Abnormal Earnings Equity Values Estimates. Journal of Accounting Research, 38, 45-70.

Herrmann, V., Richter, F. (2003). Pricing with Performance-Controlled Multiples. Schmalenbach Business Review, Vol.55.

Ho, K.-C., Lee, S.-C., Liu, C.-T., Min-Teh, Y. (2017). A comparative analysis of accounting-based valuation models. Journal of Accounting, Auditing Finance, 32: 561–575.

Hoberg, G., Phillips, G. (2008). Product Market Synergies and Competition in Mergers and Acquisitions. National Bureau of Economic Research, (No. 14289).

Hoberg, G., Phillips, G. (2016). Text-based Network Industries and Endogenous Product Differentiation. Journal of Political Economy, 124 (5): 1423-1465

Hoberg, G., Phillips, G., Prabhala, N. (2014). Product Market Threats, Payouts and Financial Flexibility. The Journal of Finance, 69(1): 293-324

Hrazdil, K., Trottier, K., Zhang, R. (2013) A Comparison of Industry Classification Schemes: A Large Sample Study. Economics Letters, 118(1): 77-80

Kaplan, S.N., Ruback, R.S., (1995). The Valuation of Cash Flow Forecasts: An Empirical Analysis. The Journal of Finance, 50 (4): 1059-1093.

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

Knudsen, J.O., Kold, S., Plenborg, T. (2017). Stick to the Fundamentals and Discover Your Peers. Financial Analysts Journal, 73(3): 85-105

Lee, C.M.C., Ma, P., Wang, C.C. (2015). Search-Based Peer Firms: Aggregating Investor Perceptions Through Internet Co-Searches. Journal of Financial Economics, 116 (2): 410-431.

Liu, J., Nissim, D., Thomas, J. (2002). Equity Valuation Using Multiples. Journal of Accounting Research, 40 (1): 135-172.

Liu, E., Lie, H.J. (2002). Multiples Used to Estimate Corporate Value. Financial Analysts Journal, 58 (2).

Nissim, D. (2013). Relative Valuation of U.S. Insurance Companies. Review of Accounting Studies, 18, 324-359.

Pinto, J.E., Robinson, T.R., Stowe, J.D. (2018). Equity Valuation: A survey of professional practice. Review of Financial Economics, 19 (2): 219-233.

Plenborg, T., Pimentel, R.C. (2016). Best Practices in Applying Multiples for Valuation Purposes. The Journal of Private Equity, 19 (3): 55-64.

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

Sharma, M., Prashar, E. (2013). A Conceptual Framework for Relative Valuation. The Journal of Private Equity, 16 (3): 29-32.

Skogsvik, K., Skogsvik, S. (2009). P/E-ratios in relative valuation – a mission impossible?. Investment Management and Financial Innovations, 5 (4): 237-248.

Roosenboom, P. (2007). How Do Underwriters Value Public Offerings? An Empirical Analysis of the French IPO Market. Contemporary Accounting Research, 24(4).

Runsten, M. (1998). The Associatoin Between Accounting Information and Stock Prices, Model development and empirical tests based on Swedish data. Stockholm School of Economics.

Sloan, R.G. (2002). Discussion of Who Is My Peer? A Valuation-Based Approach to the Selection of Comparable Firms. Journal of Accounting Research, 40(2):441-444

Williams, J,. (1939). The Theory of Investment Value. Journal of Political Economy, 47(2): 276-278.

Yee, K. (2004) Combining Value Estimates to Increase Accuracy. Financial Analysts Journal, 60 (4): 23-28.

Young, S., Zeng, Y. (2015). Accounting Comparability and the Accuracy of Peer-based Valuation Models. The Accounting Review, 90 (6): 2571-2601.

6.2 Books

Damodaran, A. (1994). Damodaran on Valuation: Security Analysis for Investment and Corporate Finance. United States: John Wiley Sons

Damodaran, A. (2002). Investment Valuation. United States: John Wiley Sons

Graham, B., Dodd, D. (1934). Security Analysis. United States: Whittlesey House, McGraw-Hill Book Co.

Penman, S.H. (2013). Financial Statement Analysis and Security Valuation. United States: McGraw Hill

Pratt, S.P., Reilly, L., Schweihs, S. (1999). Valuing Small Businesses and Professional Practices. United States: McGraw Hill

7. Appendix

Appendix 1: Example of constructed peer group using TNIC with the ten most similar firms in score

GVKEY1	GVEY2	Firm 1	Firm 2	Score
332115	33601	Armata Pharmaceuticals Inc	Translate Bio Inc	0.2624
332115	33274	Armata Pharmaceuticals Inc	Surface Oncology Inc	0.2593
332115	32899	Armata Pharmaceuticals Inc	Restorbio Inc	0.2507
332115	13835	Armata Pharmaceuticals Inc	Regulus Therapeutics Inc	0.2504
332115	18161	Armata Pharmaceuticals Inc	Conatus Pharmaceuticals Inc	0.25
332115	17350	Armata Pharmaceuticals Inc	Tetraphase Pharmaceuticals	0.2495
332115	33720	Armata Pharmaceuticals Inc	Rubius Therapeutics Inc	0.2475
332115	33664	Armata Pharmaceuticals Inc	Neon Therapeutics Inc	0.247
332115	26148	Armata Pharmaceuticals Inc	Nabriva Therapeutics Plc	0.2453
332115	21223	Armata Pharmaceuticals Inc	Contrafect Corp	0.245

Appendix 2: Comparison between peer selection in the SARD versus SADR method

The three following tables illustrates the difference peers selected when constructing peer groups using either SARD or the adjusted variation SADR. The following 10 firms across three variables has consciously been chosen to illustrate their difference where the most similar peers in absolute terms are the six firms below the considered firm. The column ”total” represents the score for each method, whereas firms included in the peer groups are denoted with * and are in bold.

Peers	ROE	Implied growth	Size
Peer 1	1%	1%	2
Peer 2	2%	3%	3
Peer 3	3%	2%	1
Considered firm	**10%**	**10%**	**10**
Peer 4	11%	11%	15
Peer 5	12%	13%	11
Peer 6	13%	14%	12
Peer 7	14%	12%	16
Peer 8	15%	16%	14
Peer 9	16%	15%	13

Appendix 2.1: Constructed peer groups from SARD ranking and subsequent SARD scores, where chosen peers are denoted with *

	SARD ranking			SARD scores			
Peers	**ROE**	**Implied growth**	**Size**	**ROE**	**Implied growth**	**Size**	**Total**
Peer 1	1	1	2	3	2	8	**8***
Peer 2	2	3	3	1	1	4	**4***
Peer 3	3	2	1	2	3	6	**6***
Considered firm	**4**	**4**	**4**	**-**	**-**		
Peer 4	5	5	9	1	5	7	**7***
Peer 5	6	7	5	3	1	6	**6***
Peer 6	7	8	6	4	2	9	**9***
Peer 7	8	6	10	2	6	12	12
Peer 8	9	10	8	6	4	15	15
Peer 9	10	9	7	5	3	14	14

Appendix 2.2: Constructed peer groups by instead using SADR ranking and subsequent SADR scores

Peers	**ROE**	**Implied growth**	**Size**	**Total**
Peer 1	9	9	8	26
Peer 2	8	7	7	22
Peer 3	7	8	9	24
Considered firm	**-**	**-**	**-**	
Peer 4	1	1	5	**7***
Peer 5	2	3	1	**6***
Peer 6	3	4	2	**9***
Peer 7	4	2	6	**12***
Peer 8	5	6	4	**15***
Peer 9	6	5	3	**14***

Appendix 3: Multicollinearity of independent variables from using the linear multivariate method

	P/E Ind. mean	**Ad. Profit Margin**	**Adj. Growth**	**Leverage**	**R&D Expense**	**ROE**	**Equity Beta**
P/E Ind. mean	1						
Adj. Profit Margin	0.0298	1					
Adj. Growth	0.0018	-0.1950	1				
Leverage	-0.0069	0.0313	0.0532	1			
R&D Expense	0.1137	0.0157	0.0605	-0.1272	1		
ROE	-0.0336	0.1275	0.0818	0.5424	0.0745	1	
Equity Beta	-0.1236	-0.0597	0.0773	0.0889	0.0400	0.0806	1