The role of comparable sets on market multiples' accuracy in European stock markets

Adriano Rafael Ferreira Lopes

Internship report
Master in Finance

Supervised by
Professor Jorge Bento Ribeiro Barbosa Farinha, PhD

2020

Acknowledgements

To Professor Jorge Farinha, for his guidance throughout this research. I would like to express my deep gratitude for his generous support, advice, patience, availability and readiness to clarify my doubts.

To my colleagues at Deloitte, for having welcomed and integrated me in the projects to the best of their ability, and for all the support given throughout the curricular internship.

To Catarina, for her encouragement and company in the pursuit of my academic achievements.

To my mother and father, for all the support, advice and for never letting me give up on my goals.

To my professors and colleagues, for sharing their experiences and wisdom, enabling me to complete this stage of my life.

Abstract

Valuation is core to most of what is done in Finance. Whether we are discussing market efficiency or behavioural finance, the notion of value, and its creation, is a persistent point of contention. The diversity in notions of value has led to the creation of various valuation methods, and respective adjustments, that aim to capture fundamentals perceived to be value creating.

This internship report discusses the methodology adjacent to Relative Valuation and aims at uncovering optimal precision practices. This study updates previous research and bridges the gap between theoretical and empirical applicability of the model. To achieve this, the impact of the comparable set criteria on the ability of the most commonly used multiples to explain value will be tested. To access precision, an error dispersion variable will be created using the market price, acting as a proxy for value, subtracted the results derived from the multiple.

We have found evidence that, in terms of relative accuracy, comparable sets designed around industry classifications outperform all other comparable criteria. Surprisingly, comparable sets built using the proximity of key financial ratios, namely the Return on Equity ratio, were outperformed by all other rules, even when comparables were based on all firms of the sample. Additionally, our evidence suggests that choosing comparables based on the level of economic integration of the respective country of main listing increases the precision of the Relative Valuation method when compared with a criterion based solely on country. Our results are robust for throughout our period of analysis, different error dispersion measures and statistical tests.

Keywords: Equity Valuation; Multiple; Comparable; Accuracy

JEL Classification: G13, G15, F15

Resumo

O trabalho de avaliação é essencial a todas as atividades compreendidas na área das Finanças. Quer se discuta a eficiência dos mercados financeiros quer finanças comportamentais, a noção de valor, e a respetiva criação, é constantemente um ponto de contenção. As diversas noções de valor levaram à criação de diferentes modelos de avaliação, e respetivos ajustes, que procuram capturar as mais diversas fontes de criação de valor.

O presente relatório de estágio discute a metodologia implícita à utilização da avaliação através de múltiplas e tem como objetivo descobrir as práticas maximizadoras de precisão. Este estudo atualiza resultados obtidos em trabalhos anteriores e aprofunda a conexão dos âmbitos académico e profissional. Para alcançar o anterior, será avaliado o impacto dos critérios de definição de empresas comparáveis na habilidade das múltiplas mais comuns em explicar valor. De modo a testar a precisão, calcularemos uma variável de dispersão do erro através da diferença entre o preço de mercado, que agirá como aproximação de valor, e a estimativa produzida pela múltipla.

Os resultados encontrados sugerem que, em relação à precisão relativa, os métodos de seleção de comparáveis construídos em torno das classificações de indústria são mais precisos do que todos os outros métodos testados. Surpreendentemente, grupos de comparáveis construídos através da proximidade de rácios financeiros chave, nomeadamente o Rendimento do Capital, verificaram o pior desempenho entre todos os métodos de seleção, mesmo quando comparados com o método baseado na seleção de todas as empresas na amostra. Adicionalmente, os resultados sugerem que comparáveis baseados no nível de integração do respetivo país aumenta a precisão da avaliação quando comparada com utilizar meramente o país. Os nossos resultados são robustos ao longo do período de análise, diferentes medidas de dispersão do erro e testes estatísticos.

Palavras-Chave: Avaliação Patrimonial; Múltipla; Comparável; Precisão

Classificação JEL: G13, G15, F15

Content Index

1. **Introduction** **1**
2. **Curricular Internship** **4**
    - 2.1. Overview of Deloitte's operations 4
    - 2.2. Description of activities developed 4
3. **Literature Review** **6**
    - 3.1. Valuation 6
    - 3.2. Relative Valuation 9
        - 3.2.1. Overview 9
        - 3.2.2. Comparable group 9
        - 3.2.3. Choice of multiple 11
        - 3.2.4. Adjustments 15
        - 3.2.5. Central dispersion measure 17
    - 3.3. Internship Report 18
4. **Hypothesis Building** **19**
5. **Methodology** **24**
6. **Data** **28**
    - 6.1. Sample Selection 28
    - 6.2. Descriptive statistics 30
7. **Results and Discussion** **31**
    - 7.1. Multiples employed 31
    - 7.2. Comparable set criteria 33
        - 7.2.1. Country consistency 38
        - 7.2.2. Valuation errors over time 40
        - 7.2.3. Comparable set in industry 43
    - 7.3. Robustness Tests 45

8. Conclusions **48**

9. References **51**

10. Appendices **55**

Appendix A 55

Appendix B 57

Appendix C 58

Index of Tables

Table 1 - Valuation methods sorted by perceived value source. 6

Table 2 – Popularity rate of multiples. 12

Table 3 – Fundamentals determining multiples. 15

Table 4 - Sample collection and data process 29

Table 5 - Number of observations by country. 29

Table 6 - Descriptive Statistics 30

Table 7 - Absolute prediction errors across multiples and comparable set criteria. 33

Table 8 - Comparison of different comparable set criteria 37

Table 9 - Distribution of the absolute median error by comparable set criteria and company of listing 39

Table 10 – Prediction error comparison of different comparable set criteria 47

Index of Figures

Figure 1 - Ranking of different comparable set criteria 38

Figure 2 - Error dispersion by comparable set criteria throughout time 42

Figure 3 – Interquartile dispersion of absolute error by SIC economic sector 43

Figure 4 – Interquartile dispersion of absolute error by TRBC economic sectors 45

Figure 5 – Economic Integration 57

Figure 6 – Interquartile dispersion of absolute error by comparable set 58

1. Introduction

Valuation is central to many of the activities comprised in the field of finance and is, therefore, part of the daily activities of both finance academics and practitioners. Academics, for once, rely on value estimates for the study of efficient corporate finance decisions, portfolio management and market efficiency, among others. Conversely, finance practitioners routinely estimate firm value for investment decisions, capital budgeting, merger and acquisition, and initial public offering (IPO) decisions (Bancel & Mittoo, 2014). Despite the fundamental role of valuation, there is still much contention regarding the sources which has spurred various valuation models and respective variants. Amongst all methods, Relative Valuation stands out due to its overwhelming popularity within finance professionals (Bancel & Mittoo, 2014; Bing, 1971; Pinto, Robinson, & Stowe, 2019). Relative to other valuation methods, the popularity this method enjoys is not accompanied by a solid and cohesive academic framework which leaves the empirical properties of multiples largely unexplored (Herrmann & Richter, 2003). This gap is especially prevalent for studies which focus on European samples (Dittmann & Weiner, 2005; Schreiner & Spremann, 2007). The current state of the literature surrounding Relative Valuation constitutes the first and main motivation for the elaboration of this report. Moreover, this report was developed concurrently with a four-month curricular internship at Deloitte's Portuguese branch, during which valuation methodology was essential for the execution of most tasks. Knowledge regarding the strengths and shortcomings of the different valuation approaches were essential to diligently deal with the responsibilities of working in Financial Advisory. Despite the importance of valuation, anecdotal evidence drawn from this internship points towards the importance given to "rules of thumb" by finance professionals. The testing of these tacit guidelines acts as the second and main motivation for the elaboration of this report.

To complement the literature framework surrounding Relative Valuation we aim at: (1) compiling literature surrounding Relative Valuation; (2) testing the impact of the comparable set criteria on the ability of multiples to explain value. Ultimately, our study aims at updating past literature regarding the impact of the comparable set in multiples performance in European markets. To achieve the previous, an error dispersion variable will be created using the market price, acting as a proxy for value, subtracted the results obtained from the multiple.

Several distinct groups should be interested in the results of this report. For instance, equity analysts and portfolio managers can use the results obtained to evaluate if and how their valuation practices could be improved. Additionally, researchers can benefit from the aggregation of the literature and its expansion.

We have found evidence that, in terms of relative accuracy, comparable sets designed around industry classifications outperform all other comparable criteria. The industry classification codes employed, TRBC and SIC, did not present statistically significant differences among them when the paired two-sample Kruskal-Wallis signed rank test was employed. TRBC did, however, present lower dispersion of prediction errors as well as lower median absolute error. The similarity of these two comparable set criteria is further illustrated when analysing the performances of the different criteria in the different European countries. Industry classifications provide the most precise valuations in 12 out of the 19 countries which constitute our sample, of which half is attributable to the TRBC rule and the other half to the SIC criteria. Surprisingly, comparable set rules based on the proximity of key financial ratios were outperformed by all other criteria, even the MARKET rule (using all firms of the sample). These results were obtained when either the Return on Equity or Return on Assets ratios were used. Moreover, financial ratio rules, when combined with industry classifications (INDROE), led to worst precision than solely using industry classification rules. When analysing region based comparable rules, we found that using the country of main listing or a combination of the previous and financial ratios to build comparable groups did not lead to improvements in precision when compared with using industry classifications. We did, however, find evidence suggesting that choosing comparables based on the level of economic integration of the respective country of main listing increases the precision of the Relative Valuation method when compared with a criterion based solely on country. These results point towards a trade-off between comparability and information introduced into the valuation model as integration comparables achieve an appropriate balance point between homogeneity of institutional backgrounds and information inputted. Finally, our results are robust for the period of analysis, different error dispersion measures and statistical tests.

In the next Chapter, the internship in which this report is inserted, and the tasks carried out will be described. Moreover, Chapter 3 and 4 details the state of the literature surrounding Relative Valuation and the hypotheses the present research aims to test. In Chapter 5 methodological considerations on the implementation of the model will be discussed. Chapter 6 will provide a brief description of the sample used while Chapter 7 will present the empirical results. Finally, Chapter 8 will conclude with a brief summary of the main findings and their respective limitations.

2. Curricular Internship

As discussed in the previous Chapter, the curricular internship performed at Deloitte Touche Tohmatsu Limited (hereinafter called Deloitte) established the main motivation for this report. This internship had the duration of four months, between October and January, and acted as a supplement to the Master's in Finance syllabus. This Chapter will briefly present the organizational context in which the internship was carried out and the nature of the work performed.

2.1. Overview of Deloitte's operations

Deloitte, founded in 1845 by William Welch Deloitte in London, is a multinational professional services network, a member of the "Big Four"[1] accounting organizations. This international network is composed by more than 312 000 professionals in more than 150 countries and territories. In the end of 2019 Deloitte's fiscal year, Deloitte registered a reported aggregate revenue of USD 46 200 million. The firm's operations can be grouped into the following five business areas[2]:

1. Consulting;
2. Audit and Assurance;
3. Financial Advisory;
4. Risk Advisory;
5. Tax & Legal.

2.2. Description of activities developed

The curricular internship was developed at Deloitte Portugal's Financial Advisory department, more precisely, the M&A advisory team. This team is responsible for the following tasks: (1) Identifying potential buyers/sellers, establishing contact and shortlisting on the basis of interest shown; (2) Creating reports with the client's financial data, valuation methodology and transaction price recommendations; (3) Aiding in the negotiations and driving the transaction process.

---

[1] The nickname "Big Four" is used to refer to the four largest professional services networks in the world. It comprises Deloitte, Ernst & Young, KPMG, and PricewaterhouseCoopers.

[2] Retrieved from Deloitte: 2019 Global Impact Report (Nov 16, 2019)

While the curricular internship consisted in performing the tasks formerly highlighted, it had as the major focus the development of valuation work. The previous consisted in, following Deloitte's guidelines, using valuation methods, such as Discounted Cash Flow and Relative Valuation, to arrive at transaction price recommendations. Within the context of the activities and the organization, it was concluded that a study regarding the impact of comparable sets on the error consistency of multiples in European Stock Markets was an appropriate and relevant research to conduct. The conclusions drawn by such research are key for the improvement of the transaction price estimates. For the previous, the main objective of this study is identifying top performing comparable set criteria in European Stock Markets.

On the next Chapter, the validity of the academically founded motivation is established by presenting the theoretical framework surrounding Relative Valuation, through an adequate review of the literature.

3. Literature Review

3.1. Valuation

Valuation activity has been subject to abundant attention from corporate finance textbooks such as Koller, Goedhart, and Wessels (2010), Damodaran (2012), and Penman (2007). While the different authors find diverse roles for valuation, they all seem to agree on its overwhelming importance. For once, Koller et al. (2010) claim that valuation activity is the main tool with which investors judge managerial performance. Investors can contrast the growth of their investment and the risk undertaken and, therefore, decide whether their activity was sufficiently compensated. More comprehensively, Damodaran (2007) sheds some light on the different roles for valuation activity, according to the different needs of the various fields in finance: (1) in **corporate finance**, investment, financing and dividend decisions are chosen based on the value they create; (2) in **portfolio management**, resources are expended in the hopes of capitalizing on divergences between price and value; and (3) in **market efficiency studies**, market prices and value are compared and, if in the presence of deviations, the rate at which these converge is analysed.

Despite the vast amounts of literature dedicated to the topic, the foundations upon which this work has been built are still a source of contention. In the following discussion, an overview of the debate on the notion of "value" will be provided and a variety of valuation models will aid in illustrating the different perceived value drivers. Generalizations regarding valuation models were made since the analysis of every methodology, and respective variants, would be unfeasible, allowing us, therefore, to have a generalized view of the rationale behind most methods (Damodaran, 2007). The following table, adapted from the conclusions presented by Fernandez (2004), provides several examples of valuation methods grouped by the characteristics perceived to be responsible for value creation:

***Table 1** - Valuation methods sorted by perceived value source. Source: Adapted from Fernandez (2004)*

Value Source	Methods			
Future returns	Discounted Cash Flow	Free Cash Flow to Equity	Adjusted Present Value	Discount Dividend Model
Flexibility	Contingent Claim Valuation	Expand Option	Abandon Option	Alternate Uses
Relative performance	Price Earnings Ratio	Dividends per Share	EV/EBITDA	EV/Sales
Book values	Book value	Adjusted Book Value	Substantial Value	Liquidation Value
Residual Income	Economic Profit	Economic Value Added	Cash Value Added	CFROI

Firstly, models such as Free Cash Flow to Equity (FCFE), Gordon's Growth (or Dividend Discount) and Adjusted Present Value (APV) have their roots in **discounting future cash flows**. By including a discount rate that incorporates the notions of risk and time-value of money, these assume that value comes from, aside from volume, the predictability and timing of an asset's cash flows[3]. Residual income models move away from the cash generating focus of the DCF models to **focus on book value of equity and net income**. Models such as Economic Value Added (or EVA) or Economic Profit aim to provide improved measures of the performance of managers or business units[4]. A distinct set of practitioners recognize the ability of a firm to react and adapt to its environment. These models, rooted in **real options**, incorporate flexibility as a potential value driver. Models such as Contingent Claim Valuation (CCV) are included[5] in this segment. Conversely, some methods highlight the importance of **book values and accounting estimates**. Methods such as Liquidation and Book Value compare the current value of the assets detained by the firm, excluding any future considerations from calculation, with the value derived from recurring operations[6]. Finally, the last set of methods are based, not only on fundamental metrics, but also on the **relative performance** of these with the values registered by peers. The rationale behind multiples valuation (also Relative Valuation or RV) is that instruments with similar characteristics such as growth opportunities or size should yield similar results.

Despite their widespread usage by finance practitioners, multiples have been subject of few academic studies (Herrmann & Richter, 2003). The popularity of the various valuation methods has been documented by several studies, some of which will be highlighted ahead. One of the earliest examples of this is the work performed by Bing (1971). By surveying several professionals in leading financial institutions, in-depth information regarding employed valuation techniques and respective implied theories was gathered. Nearly 75% of surveyed practitioners claim to prefer some variation of multiple valuation to other methods. This insight led the author to state that the "(…) results clearly indicate the wide current gap between a large body of theory and prevailing practice" (Bing, 1971, p. 57). These results are

[3] For further discussion see Parker (1968).

[4] For further discussion see Fernandez (2003).

[5] For further discussion see Quigg (1993).

[6] For further discussion see Berger, Ofek, and Swary (1996).

consistent with the work performed by Bancel and Mittoo (2014) which have surveyed finance practitioners in various European countries with CFAs or equivalent professional degrees. It has found that around 80% of practitioners use Relative Valuation when performing valuation work. Similarly, Damodaran (2012) notes that nearly 90% of equity research valuations and 50% of acquisition valuations employ some combination of comparables valuation methods. Most recently, Pinto et al. (2019) survey CFA Institute members with equity analysis jobs and found that about 93% of respondents claim to employ market multiples in their valuation work. This figure points to the overwhelming popularity of market multiples among equity professionals, especially when compared with the adoption rate of present value approaches of about 79%.

The overwhelming appeal of relative valuation can be attributed its market-based nature and simplicity, when compared with other valuation methodology (Dittmann & Weiner, 2005; Harbula, 2009). Imam, Barker, and Clubb (2008) investigate this conception by performing interviews to UK investment analysts regarding valuation methodology. Respondents strongly claimed that valuation services are client-oriented and, therefore, methodology is chosen depending on the client's preferences of valuation models. Additionally, client relationships strongly dictate methodology since, as is said by one interviewee, "[the] PE is the most effective short hand way of communicating the information that DCF can't convey. (…) The PE is a good single figure approximation". (p.521) Moreover, due to requiring less information and less assumptions than other models, RV is faster to perform and easily understood by clients and other practitioners (Damodaran, 2007). Finally, to use RV is to neglect the idea of fundamental value and assume, implicitly, that markets are efficient (Damodaran, 2007; Michaud, 1990).

While based on simple and sensible principles, the necessary inputs for Relative Valuation can easily lead to suboptimal results. In practice, these inputs are usually based on "rules of thumb", often learned through the experience of practitioners. The next section aims at exploring Relative Valuation methodology and compiling evidence, provided by the surrounding literature, on the optimal practices regarding this method.

3.2. Relative Valuation

3.2.1. Overview

The Relative Valuation approach consists of deriving the value of an asset from the market's valuation of a similar one. As seen previously, the rationale behind it is that assets with similar characteristics are expected to share similar returns. In summary, this equity valuation method requires practitioners to follow four main steps: (1) **Arrange a set of firms with similar characteristics** (often referred to as comparables); (2) **Standardize firm specific information into comparable figures** (usually named multiples) using performance measures; (3) For comparable groups which register a wide dispersion of a specific characteristic, **subjective adjustments** can be made towards uniformization; (4) **Summarize multiples** drawn from the comparable set by using a central dispersion measure.

In order to reach a share price prediction, the value extracted from the comparables is multiplied by the value driver of the company being valued. The process can be expressed as follows:

$$\hat{P}_{i,t} = VD_{i,t} \times CTM_{j \in y} \left\{ \frac{P_{j,t}}{VD_{j,t}} \right\}$$

Where $\hat{P}_{i,t}$ and $VD_{i,t}$ are the estimated price and value driver per share of firm i and $P_{j,t}$ and $VD_{j,t}$ are the actual price and value driver per share of firm j in period t. The multiple drawn from comparables is computed using a central tendency measure CTM over all firms j in y, the set of comparable firms for firm i.

The steps mentioned above will be analysed in further detail, respectively, in subsections 3.2.2, 3.2.3, 3.2.4 and 3.2.5.

3.2.2. Comparable group

Constructing a set of firms or assets with similar characteristics to the one under analysis constitutes the first stage of Relative Valuation. Ideally, comparables should share similar risks, growth prospects and cash-flows. Due to this loose definition, building a comparable set requires two key inputs. Firstly, the degree of rigorousness on what constitutes a "comparable" must be set. While being meticulous in this step might improve outputs, some advantages regarding the speed and simplicity of application could be lost (Damodaran, 2007). Secondly, some consideration must be given to the size of the comparable group. While defining a comparable group more broadly increases information, it is also expected for the diversity of firms to increase, introducing noise into the model (Cooper & Cordeiro, 2008; Damodaran, 2007).

For capital market research, and more specifically Relative Valuation, this need for firms to be divided into more homogeneous groups is commonly achieved through industry classifications. The most commonly discussed industry classification systems are the Standard Industrial Classification (SIC), which has been largely replaced by North American Industry Classification System (NAICS), and the Global Industry Classification Standard (GICS) (Bhojraj, Lee, & Oler, 2003). Academics have deliberated over the suitability of industry classification as proxies for industries criteria in valuation by analysing the impact of their usage on the precision of multiples. Bhojraj et al. (2003) have studied how using different industry classifications, as criteria to determine the comparable group, changes the predictive ability of multiples. The results indicate that GICS (Global Industry Classification Standard) codes more completely explain cross-sectional variations in valuation multiples than SIC, NAICS and Fama-French Industry Classifications. It argues that GICS "(…) provide superior industry classifications for most fundamental analysis and valuation studies that call for industry-based control samples" (Bhojraj et al., 2003, p.22). A different strain of academics debated the inadequacy of industry classifications and aimed at uncovering alternative proxies. For once, Alford (1992) studies the effect of the set of comparables on the accuracy of the price-earnings multiple and finds that selecting comparable firms by two-digit and three-digit SIC code yields similar results when using firms with similar risk and earnings growth prospect to build the comparable set. Similarly, Herrmann and Richter (2003) study multiple precision for several earnings and cash flow multiples, and the impact of comparables on the price predictions. For every multiple tested, considerably higher precision was achieved when comparable firms were based on relevant fundamentals, outlined using regression analysis, instead of SIC classifications. They do not find, however, that mixing industry classifications and financials to yield superior performance. These previous results are in disagreement with the Cheng and McNamara (2000) which study the precision of the PE Ratio, the MtB Ratio and a combination of the two. According to their results, picking comparables using a combination of industry categorization and fundamentals yields more precise valuations than solely using industry classifications. Finally, Boatsman and Baskin (1981) compared the effect on the precision of the PER when using a comparable set comprised of a random sample of firms from the same industry to firms with similar 10-year average growth rate in earnings and found the previous yields the most accurate estimates. They do, however, examine only 80 firms from a single year, 1976, and use a single comparable firm which, according to the literature, will lead to higher standard errors than if several equally comparable firms were identified (Cooper & Cordeiro, 2008). Regarding comparable set size, Cooper and Cordeiro (2008) study the effect of the size of comparables by testing the predictive ability of the forward earnings value driver using as comparables the nearest firms in terms of growth rates. They find that adding more comparables to the valuation brings net benefits until the set reaches a sample size of five. From that point on, adding more comparables has the benefit of adding more information, but at the cost of adding more noise.

3.2.3. Choice of multiple

Once the comparable set is built, the appropriate multiple must be chosen to ensure that firm-specific financial data is comparable to its peers. As opposed to real assets, firms differ, even if slightly, in one aspect or other, requiring them to be scaled to a common variable. Following the work of Penman (2007), we define a market multiple as a ratio of a market price variable (such as market capitalization or enterprise value) to a firm value driver (such as earnings or revenues). There is an extensive body of literature dedicated at surveying financial practitioners on their preferences regarding valuation methodology. For Relative Valuation specifically, Bancel and Mittoo (2014) find that the most popular multiples amongst European practitioners are Firm Value/EBITDA, PE Ratio and Firm Value/EBIT which were used by 83%, 70% and 68% of multiple users, respectively. Similarly, the results presented by Mukhlynina and Nyborg (2016) claim that investment professionals based in Western Europe give preference to cash-flow based multiples. Most recently and opposing previous surveys, Pinto et al. (2019) survey CFA Institute members with equity analysis job responsibilities and find that these prefer primordially the PE Ratio. The following table summarizes the popularity rate of the different multiples amongst surveyed, as reported by the different studies:

Table 2 *– Popularity rate of multiples. Source: Adapted from Bancel and Mittoo (2014), Mukhlynina and Nyborg (2016) and Pinto et al. (2019). *Pinto et al. (2019) includes both EBITDA and EBIT ratios in the same classification*

Multiple	Bancel and Mittoo (2014)	Mukhlynina and Nyborg (2016)	Pinto et al. (2019)
Firm Value/EBITDA	83%	95%	77%*
PE Ratio	68%	85%	88%
Firm Value/EBIT	45%	88%	77%
Price-Book Ratio	45%	72%	59%
Firm Value/Sales	45%	60%	40%

The preferences of practitioners, regarding multiple employment, are relatively aligned with the conclusions drawn by academics. For once, J. Liu, Nissim, and Thomas (2002) test a plethora of multiples, such as MtB Ratio, PE Ratio and PEG Ratio, in explaining stock prices across a comprehensive sample of North American firms. They concluded that earnings multiples outperformed multiples based on cash flows, book values, and sales. Their results also indicate that forward looking figures, greatly increase multiple precision. Following a similar methodology, Cheng and McNamara (2000) find that, for most definitions of comparable firms, the PE Ratio performs better than the MtB Ratio. Moreover, Kim and Ritter (1999) examine the pricing of IPOs using Relative Valuation and show that forward looking PE Ratio multiples outperform all other multiples in accuracy. However, when using historical data to build multiples, the study found that their predictive ability was severely hampered. Finally, Jing Liu, Nissim, and Thomas (2007) updated the findings in J. Liu et al. (2002) by extending their analysis to Africa, Asia, Europe and Oceania. They found that earnings forecasts represented substantially better measures of value than forecasted operational cash-flows in all five countries and most industries examined.

Despite the generalized support for earnings multiples, practitioners should be aware of some of their main flaws. For once, Easton (2004) developed a model to build earnings and earnings growth estimates and compared these with the values implied in PEG Ratio and PE Ratio. He notes that the PEG Ratio approach over relies on short term growth since it assumes short-run growth forecasts to also capture the long-run. An opinion shared with Beaver and Morse (1978) which discuss the determinants of the PEG Ratio, its theoretical foundations and limitations. Koller et al. (2010) report two additional flaws of the PE Ratio.

Firstly, the ratio is affected by a company's capital structure, making it so that an all equity company can artificially increase its PE Ratio by swapping debt for equity. Secondly, due to net income being calculated after nonoperating items, the PE Ratio can increase or decrease on the account of non-recurring gains and losses or amortization of intangibles.

As an alternative, EBITDA multiples can be used to overcome some of PE Ratio's shortcomings. Studies such as Baker and Ruback (1999) calculate multiples based on EBITDA, EBIT, and sales, and find that industry-adjusted EBITDA outperforms the remaining measures. Additionally, Gilson, Hotchkiss, and Ruback (2000) compare valuation estimates for firms emerging from bankruptcy with different valuation methods including multiples. They conclude that there are no significant differences between DCF and comparables valuation when using EBITDA as a value driver. Similarly, Kaplan and Ruback compare the accuracy of the results provided by EBITDA multiples and DCF in the context of market prices realized in leveraged transactions. While the DCF appears to produce slightly better results than RV, their findings suggest that the usage of both methods improves price forecasting. With the previous in mind, it is noteworthy that the EV/EBITDA suffers from one major limitation. The exclusion of changes in the working capital requirements and capital investments, for many industries, overstates cash-flows since the depreciation of existing assets is the equivalent of setting aside capital required to posteriorly replace the assets (Fernandez, 2002; Koller et al., 2010).

Moving on to book value drivers, Lie and Lie (2002) test the accuracy of various multiples for the firms constituting the Compustat North America database. They found that multiples using balance sheet drivers tend to outperform multiples based on sales and earnings drivers. The usage of accounting estimates of equity value has its foundation on literature which draws a strong connection between the MtB Ratio and return on equity (Wilcox, 1984). This method suffers, however, from the fact that accounting standards are victims of a great deal of subjectivity (Fernandez, 2002). The precision of this method deteriorates further when greater heterogeneity of accounting and tax regulations is presented, as is the case for European comparables (Schreiner & Spremann, 2007).

When compared with the evidence provided by the multiple precision literature against the usage of sales multiples (Baker & Ruback, 1999; J. Liu et al., 2002; Jing Liu et al., 2007), these multiples enjoys, according to various surveys (Bancel & Mittoo, 2014; Mukherjee, Kiymaz, & Baker, 2004; Pinto et al., 2019), disproportionate popularity throughout the years. Negative earnings and cash flows figures are often the main motive for the usage of sales for RV since these render the application of this method impossible. Additionally, when working with certain emerging industries, the sales multiple can be particularly useful if earnings and cash flow are perceived to be uninformative (J. Liu et al., 2002). While useful at these times, sales multiples are accompanied by some caveats. For once, these imposes the restriction of similar operating margins on the company's existing business (Koller et al., 2010). Additionally, sales multiples should only be used with Enterprise Value since they are attributable not only to shareholders but all stakeholders (Fernandez, 2002).

Academics (Cheng & McNamara, 2000; Lie & Lie, 2002) have conjectured that the combination of multiples would provide more precise results than the ones obtained individually since estimates derived from earnings-based multiples were proved to be positively biased and sales and asset multiples negatively biased. Lie and Lie (2002) estimated a new set of values by weighing equally the values generated by the accounting values and EBITDA multiples which performed better than the individual multiples. Similarly, Cheng and McNamara (2000) tested the ability of the PE Ratio and MtB Ratio multiples to predict market prices and found evidence that combining them yielded the most accurate valuation results. Yoo (2006) investigates the prospect of combining various multiples valuations to enhance the accuracy of the simple valuation technique. His findings show that a combination of historical multiples reduces valuation errors. However, this combination should not include the forward PER since historical multiples do not increment information to a forward PER. Nevertheless, since that combination improves historical multiples, it should be performed when forward-looking information is unavailable.

Although escaping the scope of the present internship report, a brief discussion on the role of industry-specific multiples is warranted at this point. By industry-specific multiples it is meant multiples which, by understanding the inner workings of a specific industry, adapt existing metrics or create new ones to produce "more informative" figures. For instance, Amir and Lev (1996) find that, for the wireless communications industry, nonfinancial indicators such as market penetration and franchise value are highly informative while conventional financials (e.g., earnings, book values, and cash flows) do not explain acquisition prices. Similarly, Schreiner and Spremann (2007) argue that knowledge-related multiples outperform traditional multiples in science-based industries. By knowledge-related multiples, they mean multiples which add back amortization and R&D expenditures to EBIT and net income. Finally, a word of caution when using nonfinancial multiples from Koller et al. (2010) which advises that "If a company cannot translate visitors, page views, or subscribers into profits and cash flow, the nonfinancial metric is meaningless." (p.353).

The key takeaway of this segment is that there is no "rule of thumb" for choosing the multiple to be employed. As Kim and Ritter (1999) put it, "there is no clear-cut answer for which multiples should be used" (p. 416). There are, however, several considerations that can be taken in this step as to maximize the efficiency of this valuation method.

3.2.4. Adjustments

Once the multiple is chosen and the comparable set is draw, we can perform subjective adjustments to the multiple as to control for disparity in a certain asset characteristic. The multiple chosen can be modified to incorporate the outlined characteristic into the valuation with the aid of a companion variable, a variable considered to explain much of the multiple's behaviour. Damodaran (2007) presents the following table with fundamentals that could be considered companion variables:

Table 3 *– Fundamentals determining multiples. Source: Damodaran (2007)*

Multiple	Fundamental determinants
Price Earnings Ratio	Expected growth, Payout, Risk
Price to Book Equity ratio	Expected growth, Payout, Risk, ROE
Price to Sales ratio	Expected growth, Payout, Risk, Net margin
EV to EBITDA	Expected growth, Reinvestment rate, Risk, ROC, Tax rate
EV to Capital ratio	Expected growth, Reinvestment rate, Risk, ROC
EV to Sales	Expected growth, Reinvestment rate, Risk, Operating margin

Damodaran (2002) deduces analytically the determinants of various multiples and promotes the use of regression analysis to determine a firm's value. Schreiner and Spremann (2007) criticize the approach encouraged by this study since it faces multicollinearity and non-normal distribution of regression residuals. With similar aims, Herrmann and Richter (2003) and Schreiner and Spremann (2007) establish theoretical links between multiples and various value drivers, drawing inspiration from valuation principles established in the literature. By analysing the strength of correlations between certain value drivers and multiples, they draw strong relationships between multiples and growth prospects, profitability and leverage.

These findings are corroborated by Cerqueira, Brito, and Couto (2017) who study the correlations between various financial ratios and several multiples to study optimal Relative Valuation inputs. They find that the RoA and RoE ratios drive most multiples as strong correlations between these were found. Moreover, they find that comparable sets should be drawn using one of these ratios since they minimize absolute mean error of multiples.

We will dedicate the next segment highlighting the debate of whether to choose historical (also known as trailing) or forwards-looking value drivers. Trailing multiples, for once, use the latest numbers in the financial statements for the recent fiscal quarter or year of a specific value driver. On the contrary, if the value driver of a multiple refers to a forecast figure, it is termed forward-looking. The literature shows overwhelming support for the usage of forecasts. For once, LeClair (1990) tests the ability of multiples adjusted for (1) current-period earnings, (2) average earnings over two years and (3) earnings attributable to tangible and intangible assets to value closely held firms. The evidence, while no tests amongst methods were conducted, suggests that average earnings perform best. Similarly, J. Liu et al. (2002), as seen above, find that forward-looking earnings forecasts reflect value better than historical accounting information. This analysis is extended to nine additional countries in Jing Liu et al. (2007) where they, once again, find that moving from trailing numbers to forecasts improves the valuation accuracy. This evidence is once more reinforced by the findings of Schreiner and Spremann (2007). For all the multiples analysed, they found that forward-looking figures outperformed trailing ones. This improvement of precision was especially noticeable for the PE Ratio multiple. Finally, Koller et al. (2010) argues that forwards looking multiples, as opposed by historical multiples, are consistent with the principles of valuation, that a company's value equals the present value of future cash flows.

The role of adjustments regarding a firm's level of leverage has also been a strong point of contention among academics. On one side, Harbula (2009), by analysing correlations between multiples and leverage, find that enterprise value multiples and equity value multiples are strongly influenced by this ratio. Enterprise value multiples, on one hand, have a mildly positive correlation with the level of financial leverage up to the 60% leverage ratio threshold (debt over assets), after which the relationship turns significantly negative. Equity multiples follow the opposite trajectory as they have a negative correlation until the same inflexion point and then turn strongly positive for strong financial leverage levels. Conversely, work performed by J. Liu et al. (2002) finds that adjusting for leverage does not improve the valuation properties of multiples such as EBITDA and sales. These results, as claimed by the authors, indicate a trade-off between signal and noise when more complexity is incorporated in RV.

3.2.5. Central dispersion measure

To summarize the values obtained by the comparable set, a decision is to be made regarding the **central dispersion measure**. The mean is commonly rejected as an optimal choice due to the weight given to outliers, leading to a constant overestimation of value (Herrmann & Richter, 2003). If, for instance, a firm has near zero earnings, its PE Ratio can sharply increase and bring biases to the whole set (Damodaran, 2007). In this regard, Beatty, Riffe, and Thompson (1999) examine different linear combinations of value drivers and their behaviour when faced with different central dispersion measures. They show that multiples' precision is maximized when using the harmonic mean, when compared with alternative simple estimators such as the simple mean, median, and value-weighted mean. These results are shared with subsequent literature. For once, Beatty et al. (1999), by testing the relative performance of multiples, compare the performance of four central tendency measures: the simple mean, the harmonic mean, the value-weighted mean, and the median. They find that the harmonic mean dominates the remaining estimators. Following similar methodologies, J. Liu et al. (2002) examine the proximity of stock prices generated by value drivers and the actual stock price and confirm that the harmonic mean produces the most precise forecasts when compared with the arithmetic mean or the median. Finally, Herrmann and Richter (2003) test (1) the precision of the different multiple, (2) the selection criteria of comparables and (3) the choice of a suitable statistical estimator. Regarding the former, poor performance was obtained when using the arithmetic mean. Contrary to the results seen in Baker and Ruback (1999) and J. Liu et al. (2002), the harmonic mean demonstrated a regular underestimation of potential market price. The authors claim that their results are a product of a heterogeneous sample and claim that "(…) in a heterogeneous sample like the one presented here, the median represents by far the best estimator of potential market price." (Herrmann & Richter, 2003, p. 19).

3.3. Internship Report

To advance research on the topic I aim to, keeping in mind the recommendations revised in the present Chapter, test the impact of the different comparable sets in the precision of the Relative Valuation method in estimating asset value. This study will update past valuation literature focused on European stock markets. To achieve this, the current price of an asset will act as a proxy for value and, by subtracting from it the Relative Valuation output, an error dispersion variable will be created. The comparable criteria considered will be ranked from "best performer" to "worst performer" according to their relative and absolute performance in every industry and in overall terms.

4. Hypothesis Building

Relative valuation is lauded, by both practitioners and academics, for its simplicity and relative quickness of performance when compared with other valuation methods (Dittmann & Weiner, 2005; Harbula, 2009). All inputs necessary for the usage of this valuation method, including the criteria for the comparable set, trade-off simplicity for accuracy and unbiasedness. If these principles were to be taken to the extreme, the comparable set should use all firms from the market as comparables. This would come at a great cost however, since the rigour of similarity between firms would significantly suffer. We would be introducing an immense amount of firms into the comparable set, which would, although increasing information, introduce a great deal of noise (Cooper & Cordeiro, 2008). Analysing the accuracy of multiples using the whole market as a comparable set has the benefit of acting as a benchmark for subsequent studies. Various empirical research has found that this criteria of comparables is inefficient when compared with the results obtained by SIC codes (Cheng & McNamara, 2000; Dittmann & Weiner, 2005; J. Liu et al., 2002). Industry classifications, by drawing industry lines, aim to group firms facing similar risks and growth prospects. We expect that SIC codes will reduce the noise introduced into the model by increasing comparability of firms and, therefore, hypothesise:

H1: *The usage of Standard Industry Classification codes when defining the comparable set improves precision and reduces biasedness of multiples when compared with the usage of all the firms in the cross-section.*

While an improvement over using the market as a whole, the problems regarding SIC codes have been extensively documented by literature (Clarke, Ford, & Saren, 1989; Fan & Lang, 2000; Kahle & Walkling, 1996). For once, Guenther and Rosman (1994) compare the Compustat and CRSP databases for SIC industry classifications and the impact of their usage in financial research. They tested the degree of intra-industry economic relatedness when using both databases and found that correlations of intra-industry monthly stock returns are larger, and variances of intra-industry financial ratios are smaller for industries based on the COMPUSTAT database. Confirming the previous results, Kahle and Walkling (1996) further show that samples based on Compustat codes are more likely to detect abnormal performance than CRSP based samples. These studies conjecture that the source of SIC codes is a driver for the conclusions drawn by financial research (Kahle & Walkling, 1996).

Finally, Bhojraj et al. (2003) presents alternatives to the SIC system in regards to Relative Valuation. They compare the ability of the SIC system, the NAICS, the Fama-French industry groupings (FF) and the Global Industry Classifications Standard (GICS) in a variety of applications to capital market research. They find that GICS classification are significantly better at explaining cross-sectional variations in valuation multiples and recommend that "(…) the GICS classification system will provide a better technique for identifying industrial peers." (p.23). As to devise an alternative to SIC codes, we use the Thomson Reuters Business Classification (TRBC). Due to data restrictions, it was impossible to use GICS classifications as initially intended. We found, however, TRBC to share various characteristics with GICS classifications and found it to be, therefore, a suitable substitute. Like GICS, the TRBC is a market-based system which judges a company's sources of revenue and earnings, as well as market perception[7]. These systems sharply differ from SIC and NAICS, which rely on production in delineating industry categories. This has led us to hypothesise the following:

H2: *The usage of TRBC codes when defining the comparable set improves precision and reduces biasedness of multiples when compared with the Standard Industry Classification codes.*

The usage of industry classifications when performing Relative Valuation assumes that firms within one industry share similar underlying economics. This assumption is contested by Herrmann and Richter (2003) which compare the impact on the precision of multiples of comparable set criteria based on industry classifications and financial ratios. This study shows that predictions of considerably higher accuracy can be achieved if comparable firms are selected based on fundamentals such as earnings growth or ROE (when using the PE ratio) rather than SIC classifications. In a similar method, Dittmann and Weiner (2005), when studying the effect of comparable set criteria on the valuation of European listed firms, found that choosing comparables from the same industry, when using the SIC codes, was suboptimal for all European countries. They suggest, as an alternative, that comparables should be drawn in terms of proximity of return on asset (RoA). These findings are further elaborated by Cerqueira et al. (2017) who analyse all Relative Valuation input stages in a comprehensive and global sample of 7.590 firms. They recommend the usage of RoA and RoE to build comparable sets since strong correlations between these financial ratios and

[7] Retrieved from Refinitiv: TRBC Sector Classification (Aug 16, 2012)

multiples were found. Additionally, they found that comparable set criteria revolving around financial ratios were optimal since multiple error dispersion was minimized when compared with industry classifications. This study will focus on a comparable set rule using RoE as it has strong connections with the literature regarding value creation (Wilcox, 1984). The stated above and to ensure comparability between this report and Dittmann and Weiner (2005) has led us to hypothesise the following:

H3: *The usage of similar firms in terms of Return on Equity ratio (RoE) when defining the comparable set improves precision and reduces biasedness of multiples when compared with the industry classification codes.*

The previous results have led some academics to conjecture that multiple valuation can be enhanced if these methods were to be used conjunctly. One example of this is Cheng and McNamara (2000) which study the valuation accuracy of the PE Ratio, the MtB Ratio and a combination of these multiples. Their results indicate that picking comparables using a combination of industry classifications and ROE ratio yields more precise valuations than solely using industry classifications. Similarly, Bhojraj and Lee (2002) demonstrate that a combination of industry membership with total assets results in improvements over the use of industry membership alone. Following these results, Herrmann and Richter (2003) devised an algorithm with the aim of delineating comparable pools. Through regression analysis, comparables are selected on the basis of similarity of financial fundamentals deemed essential for the firm being valued. They find that using the previous algorithm, with the aid of industry classifications, leads to the more efficient multiples valuation than when solely using industry classifications. Following the previous evidence, we hypothesise that:

H4: *The usage of a mixture of financial data and industry classification codes when defining the comparable set improves precision and reduces biasedness of multiples when compared with the outputs obtained individually.*

When compared with the U.S., the European market observes greater heterogeneity of accounting and tax regulations which translate into less comparability of financial data (Harbula, 2009; Schreiner & Spremann, 2007). When multiples valuation requires the usage of inter-country comparables, the potential effect on the effectiveness of the valuation method cannot be oversighted. For once, while studying the drivers of the PE ratio, Beaver and Morse (1978) find that risk and growth explain approximately 50% of the variance of the PE ratio around the mean. The remaining 50% they attribute to, among others, differences in accounting methods. Anecdotally, the authors claim that firms which use conservative accounting methods would tend to have higher P/E ratios than firms that use less conservative methods, if all else remained constant. Empirical evidence of heterogeneity in multiples among European firms has been found by Dittmann and Weiner (2005). They study the effects of the different comparable set criteria in multiple precision and find that Relative Valuation precision is enhanced when comparable pools are based on the same country of listing and a financial rule, when compared with solely using industry classifications for firms in the United Kingdom, Denmark and Greece. We expect that employing industry classifications will be inefficient when compared with selecting comparables from the same country of listing due to heterogeneity in institutional backgrounds and accounting standards that persist intra-industry. Additionally, we expect the latter results to be improved when a mixture of country of listing comparable pool and financial ratios comparable rule is employed. The previous led us to hypothesise:

H5.1: *The usage of the country of listing data when defining the comparable set improves precision and reduces biasedness of multiples when compared with the outputs obtained with industry classifications.*

H5.2: *The usage of a mixture of financial data and country when defining the comparable set improves precision and reduces biasedness of multiples when compared with the outputs obtained from comparables set by country solely.*

Dittmann and Weiner (2005) find that, for the remaining countries, comparable pools should be set from countries with similar economic integration such as the 15 European union member states (as of 2003) or the countries which compose the OECD. The OECD comparable pool was optimal for four countries while the European Union membership (as of 2003) was found optimal for eight (Belgium, Finland, France, Germany, Italy, the Netherlands, Portugal, and Spain). Indeed, a large body of evidence regarding the impact of the European efforts of economic integration in financial markets has been compiled over recent years. For once Gikas A. Hardouvelis, Malliaropulos, and Priestley (2006) and G. A. Hardouvelis, Priestley, and Malliaropulos (2004) have studied the period of implementation of the Euro and the respective impact on the differential between country specific risk and

EU specific risk factors. These papers hypothesize that efforts for increased integration would lead to convergence of inflation and interest rates of all EU countries. The evidence found suggests that the inflation and interest rates of all EU members and its three best performing countries have converged as a product of integration efforts. By having these rates converge, as is further elaborated by the literature, increased opportunities for risk sharing appeared, which in turn decreased European firms' cost of capital. In short, integration led to the harmonization of discount rates among European firms leading to more homogeneous valuation of equities. The importance of the adoption of the euro is further elaborated by the anecdotal evidence of the United Kingdom. By choosing not to join the Eurozone, UK's market showed no signs of increased integration with the EU and, therefore, deviations in valuations persisted. Following these results, Bekaert, Harvey, Lundblad, and Siegel (2013) aimed to study the impact of the EU between 1990 and 2007 on the absolute differences of earnings yields (the inverse of PE ratios), across industries in different countries. Intra-industry earnings yield differentials were used as proxies for valuation harmonization. It was hypothesized that both discount rate differentials and expected earnings growth differentials would decrease as Europe's integration level increased. The paper finds that the decrease in equity market segmentation was motivated by the various integration efforts of European countries. We expect, therefore, that similarity in financial and economic integration, due to the convergence of institutional backgrounds, will improve comparability among firms and, therefore, improve multiple estimation precision when compared with comparable pools from the same country of primary listing. Additionally, it is hypothesized that this method will be improved by using a mixture of financial ratios and the degree of economic integration to delineate the comparable set.

H6.1: *The usage of the country's degree of economic integration data when defining the comparable set improves precision and reduces biasedness of multiples when compared with the outputs obtained with same country of listing data.*

H6.2: *The usage of a mixture of the country's degree of economic integration and financial ratios when defining the comparable set improves precision and reduces biasedness of multiples when compared with solely the country's degree of economic integration.*

5. Methodology

The methodology employed in this report will follow studies such as Alford (1992), Schreiner and Spremann (2007) and Dittmann and Weiner (2005) due to their similarities in research objectives. This ensures comparability among results, allowing these studies to act as benchmarks. With regards to the **multiples chosen**, the Literature Review Chapter establishes that earnings multiples result in more accurate forecasts than multiples based on book values or sales. Moreover, there is a large body of evidence pointing towards the efficiency gains of using multiples calculated from analysts' forecasts over multiples based on historical data. As to maximize the efficiency of RV methodology, we follow the research performed by J. Liu et al. (2002), Herrmann and Richter (2003) and Schreiner and Spremann (2007) by using as multiple the EPS ratio, which we define as follows:

$$EPS_{i,t} = \frac{{P^{eq}}_{i,t}}{NIF_{i,t}}$$

Where ${P^{eq}}_{i,t}$ is the market capitalization (or the number of shares outstanding times the price per share) and $NIF_{i,t}$ is the I/B/E/S estimation of 1-year forwards net income for firm i at the end of fiscal year t.

Additionally, to further research the topic, we also use MtB Ratio and a combination of both methods. Our motivation resides in the efficiency gains in multiples valuation presented by Cheng and McNamara (2000) and Lie and Lie (2002). Our multiples focus on equity values since equity value multiples have been found to outperform entity value multiples in European listed firms (Schreiner & Spremann, 2007). We define MtB Ratio as:

$$PBR_{i,t} = \frac{{P^{eq}}_{i,t}}{BVE_{i,t}}$$

Where ${P^{eq}}_{i,t}$ is the market capitalization and $BVE_{i,t}$ is the Thomson Reuters Eikon Datastream (henceforth Datastream) book value of equity for firm i at the end of fiscal year t.

As to ponder both multiples in the estimate, we average their respective outputs as follows:

$$MIX_{i,t} = \frac{\left(\frac{{P^{eq}}_{i,t}}{BVE_{i,t}} + \frac{{P^{eq}}_{i,t}}{NIF_{i,t}}\right)}{2}$$

To condense the multiples drawn from the comparable set into a single estimator, we use the harmonic mean. Our choice of **central dispersion measure** is motivated by the results obtained by Baker and Ruback (1999), Beatty, Riffe and Thompson (1999) and Liu, Nissim and Thomas (2002). These claim that harmonic mean results yield more precise forecasts than the arithmetic mean or the median. For instance, the forward PE Ratio valuation method can be, therefore, expressed as follows:

$$\widehat{P^{eq}}_{i,t} = NI_{i,t} \times HM_{j \in y} \left\{ \frac{P^{eq}{}_{j,t}}{NIF_{j,t}} \right\}$$

In order to finalize the relative valuation process, we examine several competing **criteria of selecting comparable companies**. All methods utilized exclude from the pool of comparables the firm that is being valued. The comparable set criteria used are the following:

(1) **MARKET**: This method selects all observations in our sample from the same year as the valuation target. This will provide a useful benchmark since its results are expected to be significantly outperformed by the remaining methods. The results produced will also work as an indicator for the underlying dispersion of the total sample.

(2) **SIC**: This method uses Standard Industrial Classification codes as a classification for industry membership. We start by considering firm-years where industry is defined with the most digits available of the target firm's designated industry classification. If the subsample obtained is composed by less than 5 observations, we remove the last digit until our condition is met. Finally, if the 1-digit code is shared with less than 5 firm-years, we use all that are available. The minimum number of comparable firms is based on the results of Cooper and Cordeiro (2008) which state that the presence of five comparable firms is optimal since it minimized noise in the model and information requirements.

(3) **TRBC**: This method uses Thomson Reuters Business Classification industry codes as a classification for industry membership. This comparable set criteria shares the methodology presented above regarding the digit code used.

(4) **ROE**: This method considers as comparables the 5 firms closest to the target firm in terms of Return on Equity ratio (ROE). Once again, the number of firms considered is based on the results of Cooper and Cordeiro (2008). We have decided to use ROE above other methods due to its theoretical foundations (Wilcox, 1984) and empirical support (Herrmann & Richter, 2003).

(4) **INDROE**: This method chooses as comparables the 5 firms closest to the target firm in terms of ROE that belong to the same 2-digit TRBC industry classification.

(5) **COUNTRY**: In this set, we consider firm-years with the same country of primary listing as the comparable pool. In the instance of a country having less than 6 observations in any given year, all firm-years available will be used.

(6) **INTEG:** This method selects as comparables companies that belong to a country which enjoys similar degrees of economic integration. The INTEG variable considers three degrees of integration, these are: the European Union, the European Single Market and the Euro Zone. The countries comprising the sample which belonged to the EU during the period of 2010 until 2019 were characterized as belonging to the European Union and are the following: Austria, Belgium, Czech Republic, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Poland, Portugal, Spain, Sweden and United Kingdom. The European Single Market includes European countries which enjoy free movement of goods and services among themselves and are the following: Austria, Belgium, Czech Republic, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Norway, Poland, Portugal, Spain, Sweden, Switzerland and United Kingdom. Finally, the countries which have entered the European monetary union and adopted the euro as national currency are characterized as belonging to the Euro Zone and are the following: Austria, Belgium, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal and Spain. See appendix B for more details.

Following the work performed by Alford (1992), the **accuracy of the valuation method** is assessed by comparing each firm's predicted stock price with its actual price. Comparisons of the various multiples and comparable set criteria are based on absolute prediction errors, $e_{i,t}$, which can be expressed as follows:

$$e_{i,t} = \left|\widehat{P^{eq}}_{i,t} - P^{eq}{}_{i,t}\right|$$

Where $\widehat{P^{eq}}_{i,t}$ is the estimated share price by the Relative Valuation method for firm i at the end of fiscal year t and $P_{i,t}$ is the actual share price at this same period, extracted from Datastream.

Annual data is used since it is not affected by seasonality. Since we use actual share price as a proxy for value, we are assuming that, on average, market prices correctly reflect fundamentals. We scale pricing errors by actual price to control for the size effect, similarly to the commonly employed methodology in similar studies (Cheng & McNamara, 2000). We then summarize the accuracy of a comparable criteria by extracting the median and the absolute interquartile range, scaled by actual stock price over all years of the sample.

6. Data

6.1. Sample Selection

Table 4 reports the sample selection criteria. Financial data regarding the year end results of all constituents of STOXX Europe 600 (henceforth Stoxx 600) from the year of 2003 until 2019 were collected from Datastream database, commercialized by REFINITIV. The STOXX Europe 600 Index represents large, mid and small capitalization companies across 17 European countries[8] and represent approximately 85 percent of the total market capitalization in Western European Markets[9]. Due to this, the index will act as a proxy for the whole of the European Market. This led to our initial sample being composed of 958 firms and 9,320 firm-years. The unavailability of financial information from Thomson Reuters Eikon Datastream has led to the cutting of several observations from our sample. More precisely, the lack of data regarding shares outstanding (item WC05301 in Datastream), market capitalization (item WC08001), book value of equity (item WC05301) and analysts prediction of one year forwards net income (item DINI) have led to the exclusion of 843, 355, 0 and 201 observations, respectively (to 7,921 firm-years). Additionally, we have chosen to cut observations which presented negative value drivers (to 7,591) since it avoids negative predicted prices. To classify firms into different industries and subindustries, we use the SIC (WC07021) industry codes as well as the system provided by Datastream, the TRBC (TR4) classification. This is due to the evidence pointing towards the consistent misclassification of firms by SIC codes (Clarke et al., 1989; Fan & Lang, 2000; Kahle & Walkling, 1996). Firm-years whose either industry classifications were missing have been excluded. Moreover, observations whose country of main listing is unknown were also excluded. Finally, we proceeded to the exclusion of 125 observations for which there is a mismatch between the country of incorporation (Worldscope item 6027) and the currency of the market data (Datastream item ISOCUR). Such a mismatch occurs when a firm is not listed on a domestic but only on a foreign stock exchange. We arrive at a final sample which consists of 7,466 observations or 886 firms.

---

[8] The countries represented in Stoxx Europe 600 are Austria, Belgium, Denmark, Finland, France, Germany, Ireland, Italy, Luxembourg, the Netherlands, Norway, Poland, Portugal, Spain, Sweden, Switzerland and the United Kingdom.

9 Retrieved from Refinitiv: Stoxx Index Methodology Guide (Portfolio based indices) (Jul 1, 2020).

Table 4 - *Sample collection and data process*
This table illustrates the sample size for the different stages of data collection.
* By value drivers is meant the book value of equity and 1-year analyst prediction of net income.

Data	Operator	Descriptions	Number of Observations
Stoxx600 constituents	Include	Firms which have belonged to the Stoxx600 index between 01/01/2010 to 01/01/2020	9,320
Financial data	Availability	Financial data available in Thomson Reuters Eikon Datastream	7,921
Industry Classification	Availability	SIC and TRBC codes available in Datastream	7,921
Negative Value Drivers	Exclude	Firms demonstrate negative values in variables deemed as value drivers*	7,591
Country of main listing	Exclude	Firms which register a mismatch between the country of incorporation and currency of market.	7,466

Table 5 displays the number of observations in our sample for each country and each year. It shows the difference in size between the individual countries' capital markets and how mature financial markets such as the UK dominate the sample (over one quarter of the total number of firms). Moreover, it is noteworthy that the number of UK firm-years are steadily declining over time. Our sample is distinct from other Europe focused literature such as Herrmann and Richter (2003), Dittmann and Weiner (2005) and Schreiner and Spremann (2007) due to its temporal scope.

Table 5 - *Number of observations by country.*
This table displays the number of observations for each European country between the fiscal year of 2010 to 2019. The different firm-years are allocated to the different countries by their country of primarily listing.

Country	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Austria	12	12	12	12	12	13	13	13	13	12
Belgium	23	20	22	21	21	22	19	20	19	19
Czech Republic	2	2	2	2	2	2	2	2	2	2
Denmark	22	22	22	23	24	23	25	26	23	24
Finland	21	21	20	21	20	21	21	21	20	21
France	95	96	96	94	95	96	95	95	96	94
Germany	85	87	86	89	89	93	92	93	94	92
Greece	10	7	5	5	7	7	8	8	7	6
Ireland	12	14	15	13	15	15	14	15	15	15
Italy	41	42	42	40	42	43	43	42	43	43
Luxembourg	5	6	6	7	8	7	7	7	7	6
Netherlands	34	34	36	36	35	34	35	34	35	35
Norway	16	17	16	15	17	15	14	14	14	17
Poland	8	9	9	9	9	9	9	9	9	8
Portugal	10	10	10	8	8	7	6	6	6	6
Spain	37	38	37	38	39	41	39	39	37	37
Sweden	50	50	50	51	51	50	50	51	52	51
Switzerland	57	58	58	58	58	59	60	61	59	60
United Kingdom	208	211	206	202	203	202	195	188	188	176

6.2. Descriptive statistics

Table 6 reports some descriptive statistics for the main variables used in our study. The values have been pooled across all countries and years, and all currencies have been converted to EUR€. Our sample shows an average Book Value of Equity of 0,732 million of euros and an average Book Value of Assets of 5,692 million of euros. The median Book Value of Equity is €0,233m and the median Book Value of Assets is €0,677m. The average Market Capitalization value is €1,245m. and the median Market Capitalization value is €0,477m. The mean and the median for the analyst prediction of one-year forward net income is 96 thousand euros and 31 thousand euros, respectively. The arithmetic mean of the Return on Equity ratio is 22,7% whereas the median value is 14,4%. The P^{eq} to BVE ratio shows an average value of 371% and a median value of 114%. This ratio presents similar results to the study performed by Dittmann and Weiner (2005) where the Market-to-Book ratio poses a median value of 123%. Lastly, the P^{eq} to NIF ratio presents a mean and median value of 184% and 145%, respectively.

Table 6 - *Descriptive Statistics*

This table presents the mean, standard deviation, minimum, 25th percentile, median, 75th percentile and maximum of five main financial variables and two multiples in our analysis. **BVE** is defined as the Book Value of Equity in millions of euros at the end of the fiscal year t. **BVA** is defined as the Book Value of Assets in millions of euros at the end of the fiscal year t. **Market Cap** stands for Market Capitalization in millions of euros at the end of the fiscal year t and is computed by the multiplication between common shares outstanding and the market price per share. **NIF** is the I/B/E/S estimation of 1-year forwards net income in millions at the end of the fiscal year t+1 for firm i. **ROE** reports the Return on Equity Ratio where the numerator is Net Income at the end of the fiscal year t+1 and the denominator is the Book Value of Equity at the end of the fiscal year t. **P^{eq} to BVE** reports the ratio between the Market Capitalization and Book Value of Equity. **P^{eq} to NIF** reports the ratio between the Market Capitalization and forward Net Income.

		Mean	**St.Dev**	**Min**	**p25**	**Median**	**p75**	**Max**
Financials	BE	0.732	1.491	0	0.102	0.233	0.636	18.352
	BVA	5.692	19.675	0.004	0.261	0.677	2.581	233.178
	Market Cap	1.245	2.259	0.002	0.217	0.477	1.189	27.709
	NIF	0.096	0.193	0	0.015	0.031	0.081	3.441
	ROE	0.227	0.974	0	0.093	0.144	0.215	41.83
Multiples	P^{eq} to BVE	3.711	19.866	0.037	1.147	1.972	3.429	895.276
	P^{rq} to NIF	18.458	49.21	0.029	10.836	14.502	19.4	3021.32
Number of observations: 7,466								

7. Results and Discussion

7.1. Multiples employed

Indicators of valuation accuracy for equity value multiples are reported in Table 7. This table presents the median absolute valuation accuracy for the different multiples used and comparable set criteria. The table underlines the significant ability of RV to correctly explain equity market values as the median absolute valuation error lies below thirty five percent for more than half the methodology combinations employed. In other words, half of the value predictions are 30% below or above the actual market value of equity. The median errors range from 23,1 (the NIF_{med} multiple with the TRBC comparable set criteria) to 52,3 percent (the BVE_{med} multiple with the MARKET comparable set criteria). Table 7 also displays the results obtained from the paired two-sample Wilcoxon signed rank test of the differences in the median absolute errors between the first and second-best predictors. In this regard, the NIF multiple, when the median dispersion measure is used, shows consistency in being the best predictor. For seven out of the nine comparable set criteria employed, the NIF multiple presents the most precise valuation results when median absolute prediction errors are compared. The difference between the first best and second-best predictors is statistically significant for a 90% confidence interval for the comparable set criteria TRBC, SIC, COUNTRYROE, INTEG and MARKET. Solely COUNTRY and INTEGROE present a different multiple as best predictor. In the former, the MIX multiple, when using the harmonic mean, presents the lowest median absolute prediction error statistically significant for a 90% confidence interval. Whereas in the latter, the MIX_{med} multiple, presents the lowest median absolute prediction error although without statistical significance when compared with the NIF_{hm} multiple. The similarity in median absolute error of the NIF and BVE multiples, when the ROE and INDROE comparable sets are used, was discussed in Cheng and McNamara (2000). They claim that the major benefit of using RoE to build a comparable rule comes from the earnings information it introduces in the model. When introducing earnings based multiple, such as forward EPS, this benefit is essentially lost.

The results obtained regarding the optimal multiple were unexpected. The efficiency gains proposed by Cheng and McNamara (2000) and J. Liu et al. (2002) did not materialize since, while constantly outperforming the BVE multiple, the multiple seldom proved more efficient over the NIF multiple when the median absolute errors were compared. These results do, however, seem to be in line with the research performed by Yoo (2006) on the topic of multiple combinations. The research claims that, while combining several multiples improves the accuracy over the simple valuation technique, the gain is null when one component of the valuation is a forward net income multiple. Our results show that, as described by Yoo (2006), the introduction of a historical multiple did not increment any information when combined with the forward PE ratio. Further, the generalized efficiency loss when comparing the MIX and NIF multiples suggests that noise was introduced into the valuation model. Similarly, our results surrounding the optimal multiple dispersion measure also contradict the established literature on the subject. Contrary to academics which have explored the US financial market, (Beatty et al., 1999; J. Liu et al., 2002), the median constantly outperformed the harmonic mean (with the exception of the COUNTRY comparable rule). Shifting the focus to studies aimed at European firms provides a more complete perspective on our own results. For once, Herrmann and Richter (2003), when studying a comprehensive sample of European and US firms, find that the harmonic mean constantly underestimates market price. As an alternative, the authors suggest the usage of the median since it provides a sharp improvement on estimates. The authors further argue that, while in homogeneous samples both estimators lead to similar results, in heterogeneous samples the harmonic mean regularly underestimates the company's value. More expectedly, the BVE multiple constantly underperformed when compared with the forward PE ratio multiple. BVE multiples are better suited for predicting the value of firms in which goodwill plays a major role (Penman, 2007), which tend to decrease in importance the more mature and competitive an industry becomes. Due to our sample being composed of mostly large and mature European firms, the performance of the multiple has been eroded.

To simplify the evidence presented below, and with regards to the evidence analysed so far, we focus solely on the NIF multiples for subsequent analysis. We do not, therefore, report the results of the remaining multiples.

Table 7 **-** *Absolute prediction errors across multiples and comparable set criteria.*
This table displays the median absolute prediction error for 9 comparables selection methods for each different multiple. A selection method consists of a comparable pool (TRBC, SIC, COUNTRY, INTEG or MARKET) and a selection rule (ROE, INDROE, COUNTRYROE or INTEGROE). **'TRBC'** refers to Thomson Reuters Business Classification. **'SIC'** refers to Standard Industrial Classification. **'ROE'** refers to a selection rule that considers as comparables the 5 firms closest to the target firm in terms of book return on equity. **'INDROE'** refers to a selection rule that chooses as comparables the 5 firms closest to the target firm in terms of ROE that belong to the same TRBC industry classification. **'COUNTRY'** refers to comparables from the same country of primary listing. **'COUNTRYROE'** refers to a selection rule that chooses as comparables the 5 firms closest to the target firm in terms of ROE that belong to the same country of primary listing. **INTEG'** refers to comparables that belong to a country which enjoys similar degrees of integration. The degree of integration are as follows: the European Union, the European Single Market and the Euro Zone. **'INTEGROE'** refers to a selection rule that chooses as comparables the 5 firms closest to the target firm in terms of ROE that belong to a country which enjoy similar degrees of integration. **'MARKET'** refers to all available firms in one specific year. Absolute prediction errors were calculated independently for each of the ten years from 2010 to 2019. Then absolute prediction errors were pooled across these ten years and the mean (HM) and median (Med) shown in the table were calculated. **'NIF'** refers to Market Capitalization over analysts' prediction of 1-year forward Net Income, **'BVE'** refers Market Capitalization over Book Value of Equity and **'MIX'** refers to the mean of the two previous multiples. * indicates that the difference between absolute error of the best (in bold) and second-best predictors is statistically significant at the 10% level when the the paired two-sample Kruskal-Wallis signed rank test is employed.

Comparable Criteria:	BVE (HM)	BVE (Med)	MIX (HM)	MIX (Med)	NIF (HM)	NIF (Med)
TRBC	0,373	0,388	0,265	0,263	0,242	**0,231***
SIC	0,386	0,4	0,275	0,278	0,245	**0,234***
ROE	0,316	0,292	0,315	0,292	0,316	**0,291**
INDROE	0,265	0,258	0,26	0,255	0,26	**0,254**
COUNTRY	0,283	0,285	**0,274***	0,279	0,275	0,282
COUNTRYROE	0,487	0,506	0,334	0,346	0,27	**0,266***
INTEG	0,494	0,503	0,353	0,347	0,3	**0,268***
INTEGROE	0,283	0,282	0,282	**0,28**	0,283	0,281
MARKET	0,495	0,523	0,357	0,355	0,316	**0,284***

7.2. Comparable set criteria

Moving on to the analysis of the comparable selection criteria, Table 8 provides the mean and median price estimation error for the different comparable set criteria (SIC, TRBC, ROE, INDROE, COUNTRY, COUNTRYROE, INTEG, INTEGROE and MARKET) when the NIF_{med} multiple is utilized. Further insight on the relative performance of the different criteria is obtained from the table in the form of the results regarding the Kruskal-Wallis signed rank test.

An overview of the results obtained proves the capability of Relative Valuation to accurately predict equity prices. When using the NIF_{med} multiple, the median absolute error ranges from 23,1% to 28,4%, values similar to the ones reported by studies which employ similar methodologies. Herrmann and Richter (2003), for once, find that the PE ratio, when firm relevant financial ratios are used to build comparable pools, reports a median absolute error of 28,7%. Similarly, Schreiner and Spremann (2007) find that the one-year forward PE ratio multiple reports median absolute error of 24,4% when comparables are drawn from the SIC codes.

Starting off with the hypothesis discussion, the lacklustre precision of the MARKET comparable set criteria is far from unexpected. This rule is outperformed by seven of the other methods presented when median absolute errors are compared. Of the previous seven, six present statistically significant differences when the paired two-sample Kruskal-Wallis signed rank test is employed. The sharpest increases in precision over the MARKET rule are the industry classification comparable set criteria (from 28,4% when using MARKET to 23,1% and 23,4% when using SIC and TRBC respectively). These results are unsurprising since most US and Europe oriented literature, which test the efficacy of different comparable sets (Alford, 1992; Dittmann & Weiner, 2005; Herrmann & Richter, 2003; J. Liu et al., 2002), uncover similar results. The novelty in our study comes from the degree of the efficiency gain caused by the methodology shift. For once, Herrmann and Richter (2003) find that transitioning from using all firms in the year to build comparable sets to using SIC code decreases median error by 3,1 percentage points, which is modest when compared with our decline of 5,2 p.p. Due to the differences between MARKET, SIC and TRBC being statistically significant at a 1% level, we have enough evidence to reject the null hypothesis and, therefore, accept **H1**.

While an uncommon occurrence in the literature, both industry classifications proved to be the most reliable methodologies in terms of precision which goes along with commonplace professional practice. The outstanding performance of these comparable rules might be attributable to the multiple chosen, the 1-year forward PE ratio. Net Income based financials enjoy a direct relationship with the amount of debt incurred, due to the increase in interest paid, and are, therefore, highly sensitive to changes in leverage. Industry classifications are better suited to highlight comparables with similar levels of leverage, due to intra-industry trends, which leads to an increase in the precision of industry classification rules as a result (Dittmann & Weiner, 2005). While no statistically significant differences were found between the two, TRBC leads to both lower mean (23,1% against 23,4%) and median (42,5% against 43,7%) absolute errors. With the previous, we can infer that the TRBC rule is less susceptible to the estimation of extreme values than the SIC rule. Despite these differences, the lack of statistically significant evidence regarding the increase in precision does not allows us to reject H0.

Moving on to financial ratio rules, it was a surprise to find ROE as the worst performing value predictor. Selecting comparables on the basis of proximity of the Return on Equity ratio performed even worse than our benchmark method, MARKET, in terms of median (29,1% against 28,4%) and mean (58,4% against 55,3%) error dispersion. Additionally, every other methodology provides more accurate estimates both in terms of mean and median with statistical significance level of 5%. Due to the evidence claiming the increased accuracy in value prediction by the TRBC and SIC rules over the ROE, H0 is not rejected and, in turn, **H3** is not accepted. In other words, evidence found opposes the notion that the RoE ratio is more suitable than industry classifications to select comparables. These results are similar to the ones found by Herrmann and Richter (2003), simple financial ratios did not prove to reduce errors when compared with industry classifications based comparable rules. Contrary to these researchers, we did find statistical differences between the two methods, except the signal of the difference in precision was opposite of what was originally hypothesized. These results will be discussed in further detail in subsection 7.2.2, when the temporal performance of comparable set criteria will be described.

Following the results obtained by ROE, introducing the comparable rule of firms requiring the same 2-digit TRBC code, vastly improves the precision (29,1% to 25,4%) of the method, leading it to be statistically significant at a 1% level. The evidence collected allows us to reject the null hypothesis and, therefore, accept **H4**. Despite this improvement, the method did not surpass the simple TRBC rule in terms of precision. These results contrast with the ones obtained Herrmann and Richter (2003). We attribute these differences to our usage of a simple financial ratio when compared with a more advanced algorithm employed in their methodology. For other studies in the literature, such as Alford (1992) and Cheng and McNamara (2000), the differences in results might be explained by differences in the methodology. When employing industry classifications to build comparable set rules, we remove one digit of the code until our "5 comparable firms" rule is met. Alford (1992) and Cheng and McNamara (2000) take a more simplistic approach as they consider only a single specific number of digits for industry codes and drop observations when the criteria is not met.

Finally, we will observe the performance of comparable sets built around the firm's country and how these rules fare against financial ratios and industry classifications. Against what we have hypothesized, comparables selected from the same country do not improve value estimation precision over industry classification methods. Indeed, when compared with previous European literature (Dittmann & Weiner, 2005; Herrmann & Richter, 2003), heterogeneity in European financial markets might not be as pronounced as it was in the past decade. Due to the statistical significance of the difference in predictive ability, we cannot accept **H5.1** since evidence claims that industry classifications lead to more precise valuation output than the COUNTRY rule. This rule does, however, greatly benefit from the introduction of an added rule regarding the financial ratio RoE. When compared with COUNTRY, COUNTRYROE sees its median and mean absolute error decrease from 28,2% to 26,6% and from 50,8% to 49,2%, respectively. The previous improvement registers statistically significance at the 1% level, leading us to reject the null hypothesis and, therefore, accept **H5.2**. When shifting our attention to comparable set criteria based around the degree of economic integration of a firm's country, we see clear improvements over the COUNTRY rule. Once more, these results lead us to believe that differences of institutional backgrounds among firms from different countries have seen a steady convergence. Additionally, we are remitted towards the idea of a trade-off between comparability and information introduced into the valuation model as integration comparables achieve an appropriate balance point between homogeneity of institutional backgrounds and information inputted. The improvement registered is significant at the 1% level, which allows us to accept **H6.1**. Opposing what was hypothesized, the previous comparable rule did not see any improvements in added precision by the inclusion of a financial ratio rule. On the contrary, the mixture of methods saw a decline in its predictive ability as seen by the decrease in mean and median absolute errors (from 26,8% to 28,1% and 50,7% to 56,5%, respectively). The previous evidence, added the statistical significance of the difference reported in Table 8, leads us to fail to reject H0.

Table 8 **-** *Comparison of different comparable set criteria*

This table displays p-values of the paired two-sample Kruskal-Wallis signed rank test, and the sign test for thirty-six comparisons. **'TRBC'** refers to Thomson Reuters Business Classification. **'SIC'** refers to Standard Industrial Classification. **'ROE'** refers to a selection rule that considers as comparables the 5 firms closest to the target firm in terms of book return on equity. **'INDROE'** refers to a selection rule that chooses as comparables the 5 firms closest to the target firm in terms of ROE that belong to the same TRBC industry classification. **'COUNTRY'** refers to comparables from the same country of primary listing. **'COUNTRYROE'** refers to a selection rule that chooses as comparables the 5 firms closest to the target firm in terms of ROE that belong to the same country of primary listing. **INTEG'** refers to comparables that belong to a country which enjoys similar degrees of integration. The degree of integration are as follows: the European Union, the European Single Market and the Euro Zone. **'INTEGROE'** refers to a selection rule that chooses as comparables the 5 firms closest to the target firm in terms of ROE that belong to a country which enjoy similar degrees of integration. **'MARKET'** refers to all available firms in one specific year. These results are analysed for median values of the 'NIF' (refers to Net Income Forward) multiple.

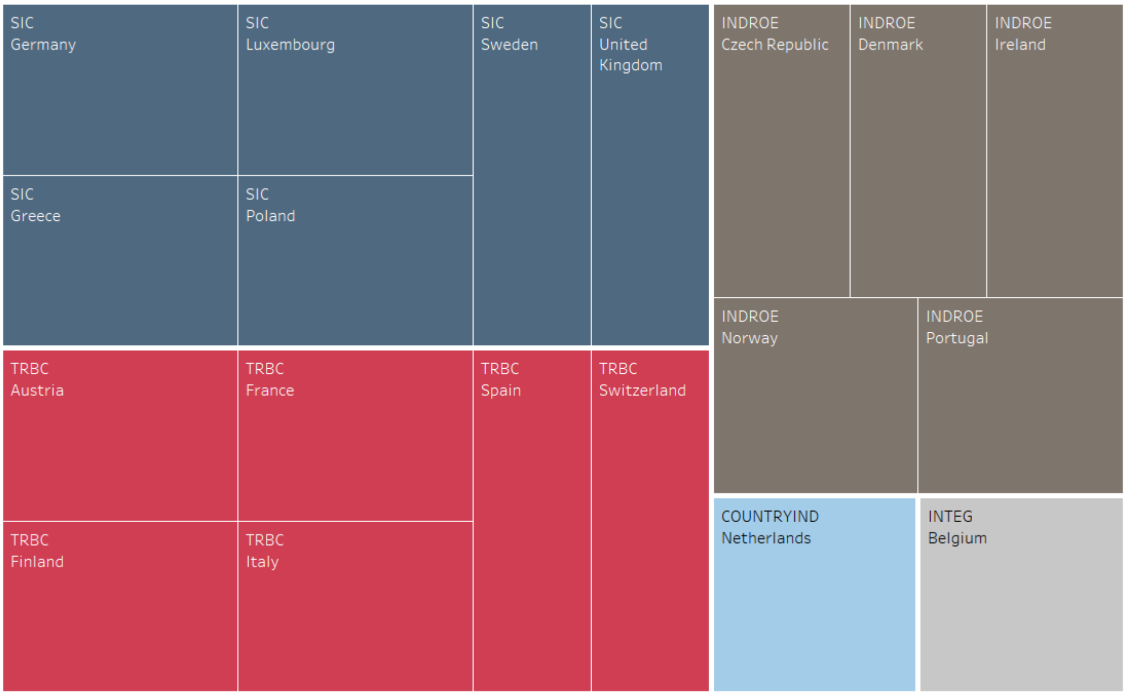
	TRBC	**SIC**	**ROE**	**INDROE**	**COUNTRY**	**COUNTRYROE**	**INTEG**	**INTEGROE**	**MARKET**
Median absolute prediction error	0,2311	0,2337	0,2914	0,2543	0,2817	0,2659	0,2683	0,2808	0,2835
Mean absolute prediction error	0,4254	0,4369	0,5842	0,5439	0,5082	0,4922	0,5071	0,5649	0,5536
									Kruskal-Wallis values
SIC	0,13								
ROE	183,00***	176,30***							
INDROE	34,87***	31,25***	57,00***						
COUNTRY	121,28***	115,39***	7,51***	24,33***					
COUNTRYROE	56,20***	51,74***	37,58***	2,22	11,98***				
INTEG	68,55***	63,50***	29,88***	4,94**	7,70***	11,98***			
INTEGROE	136,22***	129,85***	3,52*	32,40***	0,7	18,09***	12,69***		
MARKET	131,44***	125,21***	4,77**	29,58***	0,32	15,97***	10,88***	0,08	

In Subsection 7.2.1, we analyse the prediction errors across countries, in order to identify the optimal comparable selection method for each individual country. In Subsections 7.2.2 and 7.2.3, we pool the prediction errors across countries and industries. The depth to which these results will be analysed has great practical relevance as it highlights error minimizing practices.

7.2.1. Country consistency

Due to the heterogeneity of institutional backgrounds of European countries, an optimal multiple for all countries was not expected (Herrmann & Richter, 2003). Figure 1 describes the comparable set criteria that minimizes median absolute error for each country. When analysing it we see that industry classifications are the best predictors as they occupy 12 out of the 19 "best predictor rankings". While the SIC rule is optimal for Germany, Luxembourg, Greece, Poland, Sweden and United Kingdom, the TRBC is for Austria, France, Finland, Italy, Spain and Switzerland. Moving on we have INDROE with five countries in which it is optimal and then COUNTRYIND and INTEG with one country each.

Figure 1 *- Ranking of different comparable set criteria*
This figure displays the best performing comparable set criteria for each country.



Moving on to Table 9, we realize that the TRBC is the second-best predictor for 8 countries (Denmark, Germany, Greece, Ireland, Netherlands, Norway, Sweden and United Kingdom), followed by SIC which is the second-best predictor for only 4 countries (Czech Republic, France, Italy and Switzerland). Unsurprisingly, Table 9 displays that ROE is the worst predictor for 8 countries, which are Belgium, Finland, France, Greece, Italy, Netherlands, Poland, and United Kingdom. This comparable method behaves even more poorly than our benchmark method, MARKET. The ROE comparable criterion is followed by COUNTRY, MARKET and COUNTRYIND which are the worst predictors for 5, 3 and 2 countries, respectively. Our results are opposed to the findings of Dittmann and Weiner (2005), which find that financial ratios, either simple ratios or a mixture, were found to be optimal for all countries.

Table 9 **-** *Distribution of the absolute median error by comparable set criteria and company of listing*

This table displays the median absolute error of the different comparable set criteria by country of listing. **'TRBC'** (1) refers to Thomson Reuters Business Classification. **'SIC'** (2) refers to Standard Industrial Classification. **'ROE'** (3) refers to a selection rule that considers as comparables the 5 firms closest to the target firm in terms of book return on equity. **'INDROE'** (4) refers to a selection rule that chooses as comparables the 5 firms closest to the target firm in terms of ROE that belong to the same TRBC industry classification. **'COUNTRY'** (5) refers to comparables from the same country of primary listing. **'COUNTRYROE'** (6) refers to a selection rule that chooses as comparables the 5 firms closest to the target firm in terms of ROE that belong to the same country of primary listing. **INTEG'** (7) refers to comparables that belong to a country which enjoys similar degrees of integration. The degree of integration are as follows: the European Union, the European Single Market and the Euro Zone. **'INTEGROE'** (8) refers to a selection rule that chooses as comparables the 5 firms closest to the target firm in terms of ROE that belong to a country which enjoy similar degrees of integration. **'MARKET'** (9) refers to all available firms in one specific year.

Country	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
Austria	**0,2553**	0,2816	0,2587	0,3036	0,3096	0,2859	0,2886	0,2997	0,2714
Belgium	0,2259	0,2226	0,2686	0,2082	0,2424	0,2130	**0,2022**	0,2509	0,2092
Czech Republic	0,1942	0,1527	0,2108	**0,1285**	0,3030	0,3030	0,2160	0,2442	0,2320
Denmark	0,2521	0,2799	0,3147	**0,2377**	0,3243	0,2768	0,3144	0,2889	0,3175
Finland	**0,1866**	0,2084	0,2467	0,2057	0,2188	0,2122	0,2006	0,2155	0,2207
France	**0,2260**	0,2517	0,2995	0,2667	0,2938	0,2722	0,2603	0,2850	0,2713
Germany	0,2378	**0,2239**	0,3203	0,2821	0,3235	0,2825	0,2706	0,2990	0,2838
Greece	0,4244	**0,4085**	0,9353	0,6879	0,4664	0,5497	0,4814	0,6612	0,5111
Ireland	0,2439	0,2770	0,3020	**0,2398**	0,2827	0,3219	0,2750	0,3218	0,2975
Italy	**0,2027**	0,2085	0,3690	0,3277	0,3198	0,3090	0,3402	0,3688	0,3455
Luxembourg	0,2838	**0,2154**	0,2491	0,3033	0,3253	0,2912	0,2775	0,3404	0,2747
Netherlands	0,2339	0,2375	0,3000	0,2530	0,2770	**0,2252**	0,2483	0,2834	0,2850
Norway	0,2676	0,2821	0,3577	**0,2605**	0,3634	0,3295	0,4006	0,3968	0,3229
Poland	0,2375	**0,2278**	0,3473	0,2764	0,2374	0,2362	0,2657	0,2892	0,2875
Portugal	0,3645	0,3877	0,3623	**0,3134**	0,4729	0,4178	0,3951	0,3788	0,3805
Spain	**0,2203**	0,2596	0,3071	0,2555	0,2841	0,2777	0,2869	0,3110	0,3280
Sweden	0,2068	**0,2038**	0,2406	0,2174	0,2373	0,2274	0,2366	0,2358	0,2545
Switzerland	**0,1954**	0,1972	0,2522	0,2528	0,2368	0,2476	0,2521	0,2387	0,2771
United Kingdom	0,2452	**0,2363**	0,2852	0,2528	0,2689	0,2558	0,2634	0,2774	0,2773

7.2.2. Valuation errors over time

Figure 2 contains the median absolute errors for each year from 2010 to 2019 for the 9 comparable selection methods employed in our study. In order to keep this analysis simple, we do not report results for the other multiples or dispersion measures described in section 7.1.

Panel A displays the results for the TRBC, SIC and MARKET and Panel B for the TRBC, SIC, ROE and INDROE. Panel C shows the results for the comparable criteria set of COUNTRY, COUNTRYROE, INTEG and INTEGROE and Panel D displays the results for TRBC, SIC, COUNTRY and INTEG. The plots reveal that the valuation accuracy varies markedly over time. All plots share a peak in median absolute error in 2011, during the European sovereign debt crisis of 2008 to 2012. After the shock, valuation errors reverted to 2010 levels but seeing a steady growth in subsequent years until the end of the analysis period.

Panel A shows the precision of the TRBC and SIC methods over time. The figures demonstrate that TBRC and SIC codes lead to the lowest valuation errors over the period under analysis. After 2016, there is a clear advantage of TRBC over SIC codes. In contrast, MARKET's performance is comparatively poor – especially during the interval of 2010 to 2013. Hence, our result that comparables should be selected according to industry classifications (SIC and TRBC), hypothesis 1, is robust over time. Similarly to Dittmann and Weiner (2005), we find that when shocks occur, regarding valuation errors, industry classification based comparable rules lead to lower deviations from the mean.

Panel B adds ROE and INDROE to the analysis which provide some interesting insight into the "financial rule versus industry classification" debate. When compared with TRBC and SIC, these financial criteria were suboptimal for every year of the analysis period, with INDROE performing slightly better than ROE. INDROE, during the 2011 shock, saw a much smaller increase in valuation error, highlighting the benefits drawn from mixing comparable rules.

Moving on to panel C and as expected, we find that INTEG, COUNTRY, INTEGROE and COUNTRYROE, show similar median absolute error patterns through the years. The precision of COUNTRY and INTEGROE are almost indistinguishable throughout all the period of analysis. These comparable set criteria were outperformed for most years under observation by the mixture of COUNTRY and the ROE rule (COUNTRYROE) when the median absolute errors are compared. Unlike COUNTRY, INTEG loses precision when financial ratios rules are introduced for every year of the sample, except 2019.

Finally, we can see in Panel D, the outperformance of industry classification rules, TRBC and SIC, when compared with region and country-based rules for every year of the sample. We find the disparity between region and industry comparables to have increased steadily through most years under analysis. This could be a side effect of the homogenization of institutional background of European firms, leading to a convergence of European firm's financials, making industry classifications increasingly relevant.

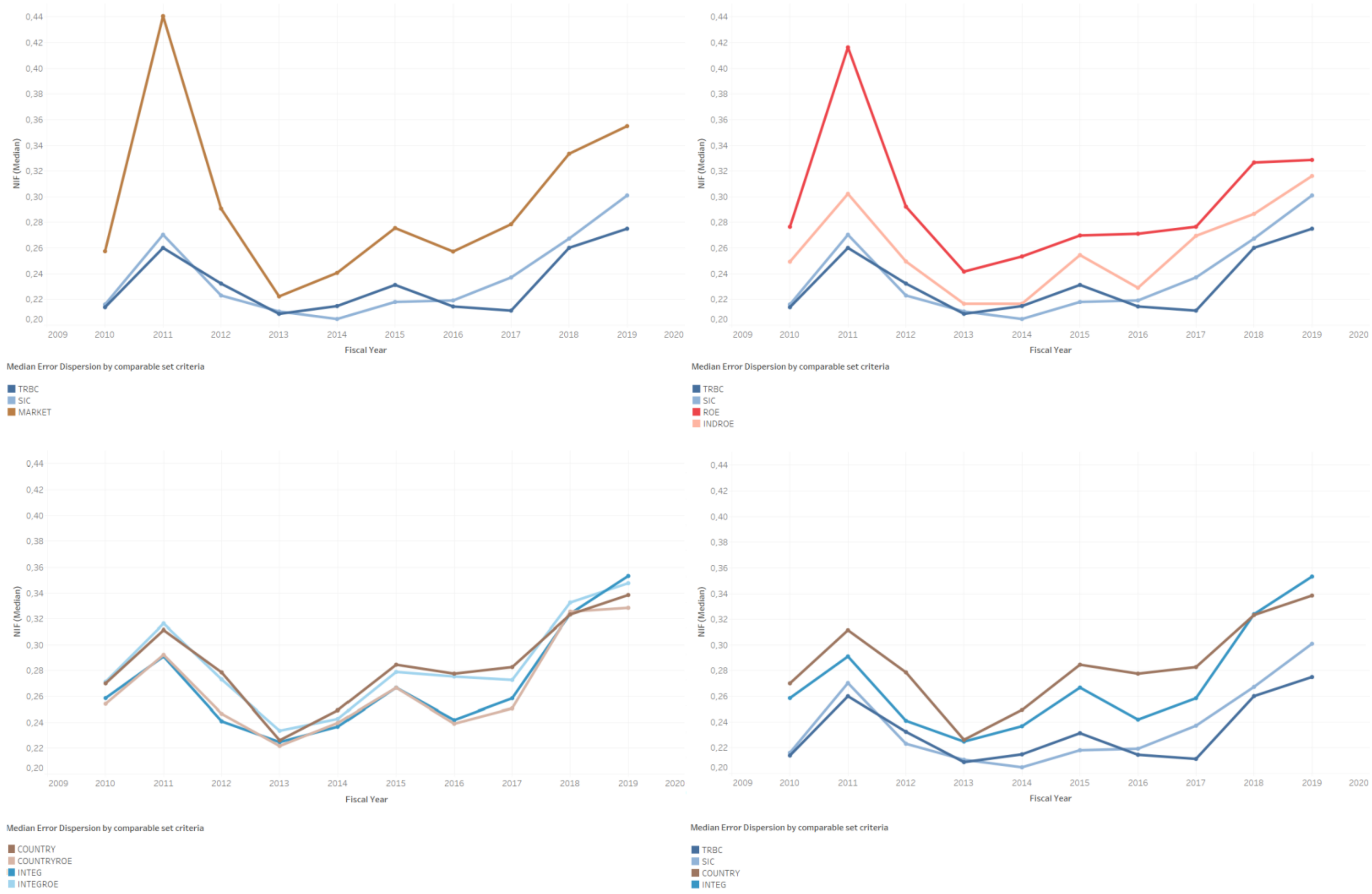
Figure 2 - *Error dispersion by comparable set criteria throughout time*

(A)

(B)

(C)

(D)



7.2.3. Comparable set in industry

To further understand the drivers of Relative Valuation error, let us dedicate this subsection to understanding which industries are most prone to relative valuation errors. The results by SIC and TRBC economic sector are shown in Figures 3 and 4, respectively. Figure 3 represents the median absolute error of the NIF_{med} multiple when the comparable rule SIC is used while Figure 5 uses the comparable rule TRBC.

As presented in Figure 3, the sector which presents the lowest absolute error dispersion is the Wholesale Trade economic sector. The interquartile range of this economic sector is 0,2486 and the median is 0,2221. Following Wholesale Trade, the sectors which have the lowest median error dispersion are the economic sector of Manufacturing, Transportation, Communications, Electric, Gas and Sanitary Service and Finance, Insurance and Real Estate with interquartile ranges of 0,2811, 0,2971 and 0,3116, respectively. The economic sector that presents the biggest median error dispersion is the Mining sector which presents a median value of 0,3885 and an interquartile range of 0,5135.

***Figure 3** – Interquartile dispersion of absolute error by SIC economic sector*

This table displays the median error dispersion by SIC economic sectors. The SIC economic sectors are as follows: **'Construction'** refers to firms whose SIC code is between 1500-1799. '**Finance, Insurance and Real Estate'** refers to firms whose SIC code is between 6000-6799. **'Manufacturing'** refers to firms whose SIC code is between 2000-3999. **'Mining'** refers to firms whose SIC code is between 1000-1499. **'Retail Trade'** refers to firms whose SIC code is between 5200-5999. **'Services'** refers to firms whose SIC code is between 7000-8999. **'Transportation, Communications, Electric, Gas and Sanitary service'** refers to firms whose SIC code is between 4000-4999. **'Wholesale Trade'** refers to firms whose SIC code is between 5000-5199.

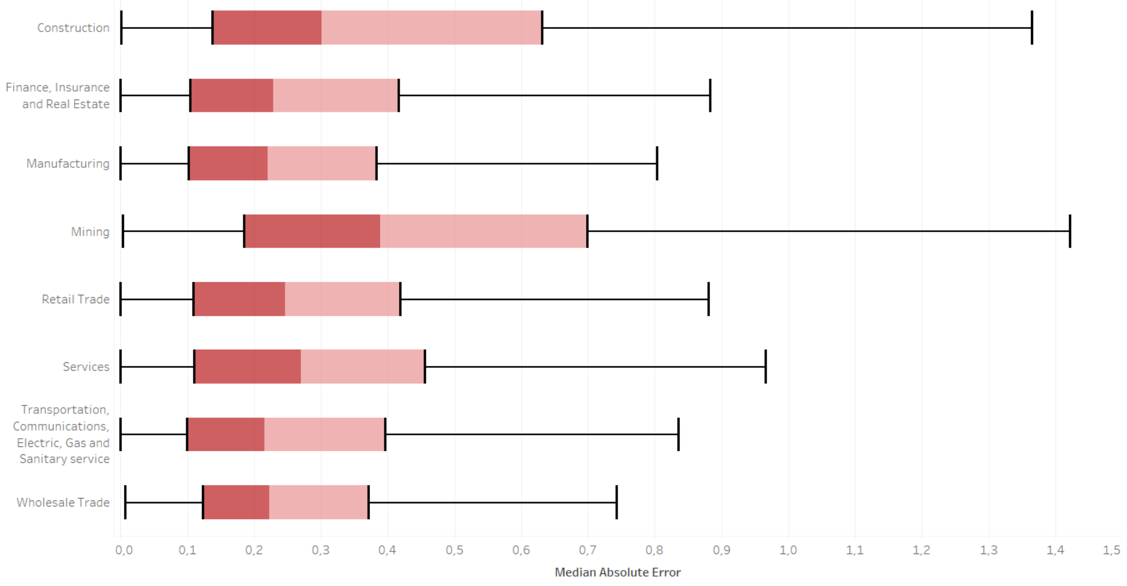
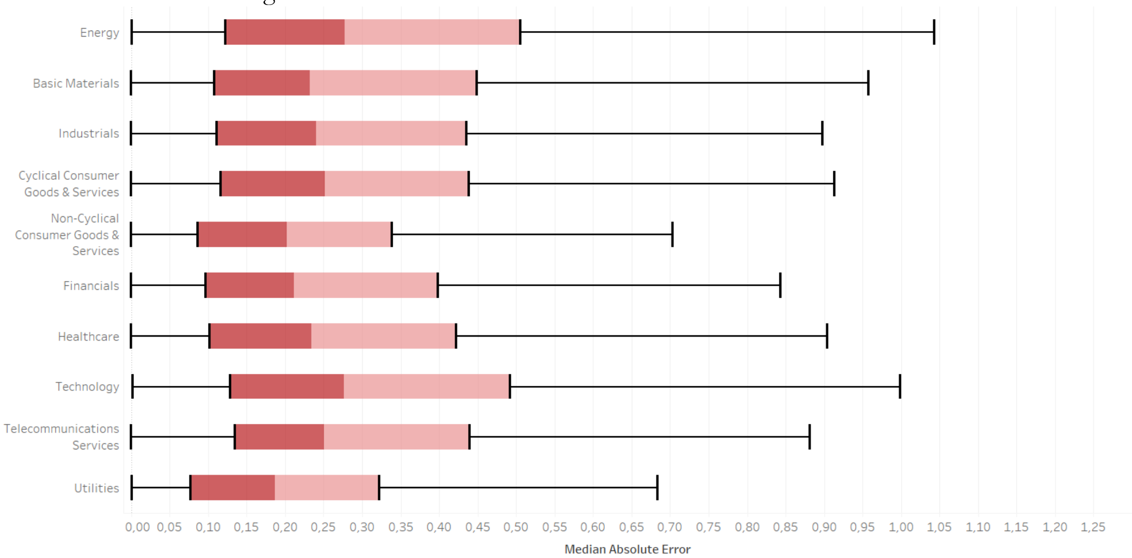


Figure 4 shows the median error dispersion by TRBC economic sectors. It shows that the Utilities sector has the lowest absolute error dispersion with an interquartile range of 0,2455 and a median value of 0,1862. The second lowest median error dispersion is presented by the Non-Cyclical Consumer Goods and Services. with a median value of 0,203 and an interquartile range of 0,2525. The highest absolute error dispersion is presented by the Energy sector with a median value of 0,2755 and an interquartile range of 0,3823.

The differences in error dispersions among industries can be mostly attributed to: (1) Industry classifications being inadequate proxies for growth prospects and risk. The previous is especially relevant for industries such as "Technology" whose lines are constantly being redrawn. As Dittmann and Weiner (2005) claim, "(…) SIC industry classification is not able to separate new economy firms from old economy firms." (p.13). (2) Industries which have less presence in the overall sample tend to have higher dispersion of financials and, therefore, lead to higher errors in multiples valuation (Cheng & McNamara, 2000). For instance, when discussing SIC codes, the "Construction" and "Mining" industries characterize 247 and 300 firm-years in the sample. When compared with the 2814 firm-years comprised in the "Manufacturing" sector, divergences become apparent. With regards to TRBC codes, these differences aren't as obvious. The "Energy" industry, the worst in terms of error dispersion, is comprised by 422 firm-years while "Non-Cyclical Consumer Goods & Services", the best performer, is comprised by 537.

***Figure 4 –** Interquartile dispersion of absolute error by TRBC economic sectors*

This table displays the median error dispersion by TRBC economic sectors. The TRBC economic sectors are as follows: **'Energy'** refers to firms whose 2-digit TRBC code is 50. **'Basic Materials'** refers to firms whose 2-digit TRBC code is 51. **'Industrials'** refers to firms whose 2-digit TRBC code is 52. **'Cyclical Consumer Goods & Services'** refers to firms whose 2-digit TRBC code is 53. **'Non-Cyclical Consumer Goods & Services'** refers to firms whose 2-digit TRBC code is 54. **'Financials'** refers to firms whose 2-digit TRBC code is 55. **'Healthcare'** refers to firms whose 2-digit TRBC code is 56. **'Technology'** refers to firms whose 2-digit TRBC code is 57. **'Telecommunications Services'** refers to firms whose 2-digit TRBC code is 58. **'Utilities'** refers to firms whose 2-digit TRBC code is 59.



7.3. Robustness Tests

The results obtained by our study rely on several proxies and assumptions described in Chapter 5. This section is dedicated to discussing the adequacy of these inputs and how results sway as these change.

Firstly, we verify the robustness of our conclusions when using a different precision measure, the interquartile error dispersion as suggest by some academics (Alford, 1992; Cheng & McNamara, 2000). Table 10 describes the quartile distribution obtained by the different comparable set and the respective interquartile range. According to this measure, the list of criteria ranked from best to worst predictor is the following: TRBC, SIC, INDROE COUNTRYROE, INTEG, COUNTRY, MARKET, INTEGROE, ROE. When compared with the performance list obtained from the median absolute error, the only change is the shift of INDROE from 6th best predictor to 3rd.

Table 10 – *Quartile distribution*

This table displays the quartile distribution and interquartile range for the different comparable set criteria.

	P25	P50	P75	Interquartile
TRBC	-0,2587	0,0021	0,2123	0,4710
SIC	-0,2541	0,0014	0,2171	0,4712
ROE	-0,3600	-0,0011	0,2578	0,6178
INDROE	-0,3098	0,0018	0,2264	0,5361
COUNTRY	-0,3362	0,0017	0,2469	0,5831
COUNTRYROE	-0,3104	0,0000	0,2309	0,5412
INTEG	-0,3150	-0,0004	0,2390	0,5539
INTEGROE	-0,3557	-0,0070	0,2430	0,5987
MARKET	-0,3383	-0,0070	0,2430	0,5814

Additionally, we confirm the soundness of the statistical tests employed in section 7.1 by employing the paired two-sample Wilcoxon signed rank test. Table 11 agglomerates the results obtained for the thirty-six sign tests. We have found the same evidence as when the Kruskal-Wallis signed rank test was employed in either signal or statistical significance.

Table 11 – *Prediction error comparison of different comparable set criteria*

This table displays p-values of the paired two-sample Wilcoxon signed rank test, and the sign test for thirty-six comparisons. **'TRBC'** refers to Thomson Reuters Business Classification. **'SIC'** refers to Standard Industrial Classification. **'ROE'** refers to a selection rule that considers as comparables the 5 firms closest to the target firm in terms of book return on equity. **'INDROE'** refers to a selection rule that chooses as comparables the 5 firms closest to the target firm in terms of ROE that belong to the same TRBC industry classification. **'COUNTRY'** refers to comparables from the same country of primary listing. **'COUNTRYROE'** refers to a selection rule that chooses as comparables the 5 firms closest to the target firm in terms of ROE that belong to the same country of primary listing. **INTEG'** refers to comparables that belong to a country which enjoys similar degrees of integration. The degree of integration are as follows: the European Union, the European Single Market and the Euro Zone. **'INTEGROE'** refers to a selection rule that chooses as comparables the 5 firms closest to the target firm in terms of ROE that belong to a country which enjoy similar degrees of integration. **'MARKET'** refers to all available firms in one specific year. These results are analysed for median values of the 'NIF' (refers to Net Income Forward) multiple.

	TRBC	**SIC**	**ROE**	**INDROE**	**COUNTRY**	**COUNTRYROE**	**INTEG**	**INTEGROE**	**MARKET**
Median absolute prediction error	0,2311	0,2337	0,2914	0,2543	0,2817	0,2659	0,2683	0,2808	0,2835
Mean absolute prediction error	0,4254	0,4369	0,5842	0,5439	0,5082	0,4922	0,5071	0,5649	0,5536
	Wilcoxon values								
SIC	-0.360								
ROE	-13.528***	-13.278***							
INDROE	-14.314***	-14.058***	-0.760						
COUNTRY	-5.905***	-5.590***	7.550***	8.316***					
COUNTRYROE	-11.013***	-10.742***	2.740***	3.529***	-4.933***				
INTEG	-7.497***	-7.193***	6.130***	6.909***	-1.492	3.461***			
INTEGROE	-8,279***	-7.969***	5.466***	6.238***	-2.223**	2.775***	-0.726		
MARKET	-11.671***	-11.395***	1.875**	2.638***	-5.692***	-0.839	-4.253***	-3.562***	

8. Conclusions

The main purpose of this work was to compile existent literature surrounding Relative Valuation and test the impact of the comparable set criteria on the ability of multiples to explain stock value.

We have found evidence that, in terms of relative accuracy, comparable sets designed around industry classifications outperform all other comparable criteria. The industry classification codes employed, TRBC and SIC, did not present statistically significant differences among themselves when the paired two-sample Kruskal-Wallis signed rank test was employed. TRBC did, however, present lower dispersion of prediction errors as well as lower median absolute error. This evidence, while not statistically significant, does agree with the notion that industry codes built by specialists and of a financial-oriented nature improve intra-industry comparability (Bhojraj et al., 2003). To further illustrate the similarity between the two industry classifications, when discussing the best performing methodology among countries that constitute our sample, we find that 12 out of the 19 countries see precision maximized when industry codes are used, of which half is attributable to the TRBC rule and the remaining half to the SIC criteria.

Surprisingly, comparable set rules based on the proximity of key financial ratios were outperformed by all other criteria, even the MARKET rule (criteria based on all other firms of the sample being assumed comparable). These results are robust when using either the Return on Equity or Return on Assets ratios. While not to such an extreme degree, the related literature find similar results. Herrmann and Richter (2003) and Cheng and McNamara (2000), for instance, find that proximity in simple financial ratios or metrics such Total Assets or Return on Equity are constantly outperformed by industry classifications.

These authors do not, however, share our conclusions with regards to the performance of the mixture of financial ratio rules and industry classifications (INDROE). We have found INDROE to be the best performing method for 5 out of the 19 countries constituting our sample while the sixth best value predictor in terms of median absolute error. The evidence collected by studies such as Herrmann and Richter (2003), Cheng and McNamara (2000), Alford (1992) clashes with our own in the sense that they either find improvements when introducing financial ratios or no statistically significant differences between them. We believe our conclusions to derive from our simplistic approach in combining the two comparable set rules.

When analysing region based comparable rules, we found that using the country of main listing or a combination of the previous and financial ratios to build comparable groups was suboptimal when compared with using industry classifications. We did, however, find evidence suggesting that choosing comparables based on the level of economic integration of the respective country of main listing increases the precision of Relative Valuation method when compared with a criterion based solely on country. These results point towards a trade-off between comparability and information introduced into the valuation model as integration comparables achieve an optimal point between homogeneity of institutional backgrounds and information inputted (Bekaert et al., 2013).

Our results are robust for different multiples, central tendency measures and error dispersion measures. Additionally, the performance of the various methodologies employed is robust throughout the period of analysis and when alternative statistical tests are employed.

The main limitations our study relate to the sample used. The usage of the Stoxx 600 index has led to some representativeness biases regarding the size of the firms used, regions of main listing and industry segments. For once, firms in countries which enjoy a higher degree of financially development such as the United Kingdom, Germany and France represent around 50% of our sample. The same occurs for large cap firms, which dominate our sample. According to previous literature, the size of the firm is directly linked to the precision of the Relative Valuation model (Cheng & McNamara, 2000), which may lead our conclusions to not be representative of all European firms. Finally, due to some industry segments being underrepresented, prediction errors may be skewed in favour of industries which enjoy more presence (Cheng & McNamara, 2000).

To mitigate these adverse effects, our suggestion for further research would be the introduction of small and medium cap European firms to the sample. While this would level the playing field for all the representation biases mentioned above, some caveats must be kept in mind. For once, traditional multiples would lose some explicative ability in favour of more industry specific multiples. Anecdotally, technology "start-ups" whose constant negative earnings would translate into null value according to multiple valuation, would more suitably be evaluated by industry-oriented metrics such as subscriber count or market penetration.

Finally, we would like to highlight how the period of analysis employed limited our analysis. The constituents of the Stoxx 600 index prior to 2010 were unable to be collected from Thomson Reuter Eikon database which made the analysis of previous years impossible. The extension of the period of analysis would prove desirable to further the discussion regarding the connection between the homogeneity of institutions and the precision of multiples valuation. If extended enough, the study of certain events which promoted integration, such as the introduction of the Euro, or which deteriorated integration, such as "Brexit", could lead to further insight on this relationship.

9. References

Alford, A. W. (1992). THE EFFECT OF THE SET OF COMPARABLE FIRMS ON THE ACCURACY OF THE PRICE EARNINGS VALUATION METHOD. *Journal of Accounting Research, 30*(1), 94-108. doi:10.2307/2491093

Amir, E., & Lev, B. (1996). Value-Relevance of Nonfinancial Information: The Wirless Communications Industry. *Journal of Accounting and Economics, 22*, 3-30. doi:10.1016/S0165-4101(96)00430-2

Baker, M., & Ruback, R. (1999). Estimating industry multiples. *Harvard University*.

Bancel, F., & Mittoo, U. R. (2014). The Gap between the Theory and Practice of Corporate Valuation: Survey of European Experts. *Journal of Applied Corporate Finance, 26*(4), 106-117. doi:10.1111/jacf.12095

Beatty, R., Riffe, S., & Thompson, R. (1999). The Method of Comparables and Tax Court Valuations of Private Firms: An Empirical Investigation. *Accounting Horizons - ACCOUNT HORIZ, 13*, 177-199. doi:10.2308/acch.1999.13.3.177

Beaver, W., & Morse, D. (1978). What Determines Price-Earnings Ratios? *Financial Analysts Journal, 34*(4), 65-76. doi:10.2469/faj.v34.n4.65

Bekaert, G., Harvey, C. R., Lundblad, C. T., & Siegel, S. (2013). The European Union, the Euro, and equity market integration. *Journal of Financial Economics, 109*(3), 583-603. doi:https://doi.org/10.1016/j.jfineco.2013.03.008

Berger, P. G., Ofek, E., & Swary, I. (1996). Investor valuation of the abandonment option. *Journal of Financial Economics, 42*(2), 259-287. doi:https://doi.org/10.1016/0304-405X(96)00877-X

Bhojraj, S., Lee, C., & Oler, D. (2003). What's My Line? A Comparison of Industry Classification Schemes for Capital Market Research. *Journal of Accounting Research, 41*, 745-774. doi:10.2139/ssrn.356840

Bing, R. A. (1971). Survey of Practitioners' Stock Evaluation Methods. *Financial Analysts Journal, 27*(3), 55-60.

Boatsman, J. R., & Baskin, E. F. (1981). Asset valuation with incomplete markets. *The accounting review : a publication of the American Accounting Association, 56*(1), 38-53.

Cerqueira, A., Brito, P., & Couto, J. (2017). The Method of Market Multiples on the Valuation of Companies: A Multivariate Approach. In.

Cheng, C. s., & McNamara, R. (2000). The Valuation Accuracy of the Price-Earnings and Price-Book Benchmark Valuation Methods. *Review of Quantitative Finance and Accounting, 15*, 349-370. doi:10.1023/A:1012050524545

Clarke, K., Ford, D., & Saren, M. (1989). Company technology strategy. *R&D Management, 19*(3), 215-229. doi:10.1111/j.1467-9310.1989.tb00643.x

Cooper, I., & Cordeiro, L. (2008). Optimal Equity Valuation Using Multiples: The Number of Comparable Firms. *SSRN Electronic Journal*. doi:10.2139/ssrn.1272349

Damodaran, A. (2007). Valuation Approaches and Metrics: A Survey of the Theory and Evidence. *Foundations and Trends® in Finance, 1*(8), 693-784. doi:10.1561/0500000013

Damodaran, A. (2012). *Investment valuation: Tools and techniques for determining the value of any asset* (Vol. 666): John Wiley & Sons.

Dittmann, I., & Weiner, C. (2005). Selecting comparables for the valuation of European firms. *Available at SSRN 644101*.

Easton, P. (2004). Discussion—Earnings Surprises and the Cost of Equity Capital. *Journal of Accounting, Auditing & Finance, 19*(4), 515-521. doi:10.1177/0148558X0401900410

Fan, J. P. H., & Lang, L. (2000). The Measurement of Relatedness: An Application to Corporate Diversification. *The Journal of Business, 73*(4), 629-660.

Fernandez, P. (2002). Valuation Using Multiples. How Do Analysts Reach their Conclusions? *SSRN Electronic Journal*. doi:10.2139/ssrn.274972

Fernandez, P. (2003). *Three residual income valuation methods and discounted cash flow valuation*. Retrieved from https://EconPapers.repec.org/RePEc:ebg:iesewp:d-0487

Fernandez, P. (2004). Most Common Errors in Company Valuation. *SSRN Electronic Journal, 2*. doi:10.2139/ssrn.545546

Gilson, S. C., Hotchkiss, E. S., & Ruback, R. S. (2000). Valuation of bankrupt firms. *Review of Financial Studies, 13*(1), 43-74. doi:10.1093/rfs/13.1.43

Guenther, D. A., & Rosman, A. J. (1994). Differences between COMPUSTAT and CRSP SIC codes and related effects on research. *Journal of Accounting and Economics, 18*(1), 115-128.

Harbula, P. (2009). Valuation multiples: Accuracy and drivers evidence from the european stock market. *Business Valuation Review, 28*(4), 186-200.

Hardouvelis, Gikas A., Malliaropulos, D., & Priestley, R. (2006). EMU and European Stock Market Integration. *The Journal of Business, 79*(1), 365-392. doi:10.1086/497414

Hardouvelis, G. A., Priestley, R., & Malliaropulos, D. (2004). The impact of globalization on the equity cost of capital.

Herrmann, V., & Richter, F. (2003). Pricing with Performance-Controlled Multiples. *Schmalenbach Business Review, 55*(3), 194-219. doi:10.1007/BF03396674

Imam, S., Barker, R., & Clubb, C. (2008). The Use of Valuation Models by UK Investment Analysts. *European Accounting Review, 17*(3), 503-535. doi:10.1080/09638180802016650

Kahle, K. M., & Walkling, R. A. (1996). The Impact of Industry Classifications on Financial Research. *The Journal of Financial and Quantitative Analysis, 31*(3), 309-335. doi:10.2307/2331394

Kim, M., & Ritter, J. R. (1999). Valuing IPOs. *Journal of Financial Economics, 53*(3), 409-437. doi:10.1016/s0304-405x(99)00027-6

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

LeClair, M. S. (1990). Valuing the Closely-Held Corporation: The Validity and Performance of Established Valuation Procedures. *Accounting Horizons, 4*(3), 31.

Lie, E., & Lie, H. J. (2002). Multiples Used to Estimate Corporate Value. *Financial Analysts Journal, 58*(2), 44-54. doi:10.2469/faj.v58.n2.2522

Liu, J., Nissim, D., & Thomas, J. (2002). Equity valuation using multiples. *Journal of Accounting Research, 40*(1), 135-172. doi:10.1111/1475-679x.00042

Liu, J., Nissim, D., & Thomas, J. (2007). Is Cash Flow King in Valuations? *Financial Analysts Journal, 63*(2), 56-68. doi:10.2469/faj.v63.n2.4522

Michaud, R. O. (1990). Demystifying Multiple Valuation Models. *Financial Analysts Journal, 46*(1), 6-8.

Mukherjee, T. K., Kiymaz, H., & Baker, H. K. (2004). Merger motives and target valuation: A survey of evidence from CFOs. *Journal of Applied Finance, 14*(2).

Mukhlynina, L., & Nyborg, K. G. (2016). The choice of valuation techniques in practice: education versus profession.

Parker, R. H. (1968). Discounted Cash Flow in Historical Perspective. *Journal of Accounting Research, 6*(1), 58-71. doi:10.2307/2490123

Penman, S. H. (2007). *Financial statement analysis and security valuation* (Vol. 3): McGraw-Hill New York.

Pinto, J. E., Robinson, T. R., & Stowe, J. D. (2019). Equity valuation: A survey of professional practice. *Review of Financial Economics, 37*(2), 219-233. doi:10.1002/rfe.1040

Quigg, L. (1993). Empirical Testing of Real Option-Pricing Models. *The Journal of Finance, 48*(2), 621-640. doi:10.2307/2328915

Schreiner, A., & Spremann, K. (2007). Multiples and their valuation accuracy in European equity markets. *Available at SSRN 957352*.

Wilcox, J. W. (1984). The P/B-ROE Valuation Model. *Financial Analysts Journal, 40*(1), 58-66. doi:10.2469/faj.v40.n1.58

Yoo, Y. (2006). The valuation accuracy of equity valuation using a combination of multiples. *Review of Accounting and Finance, 5*, 108-123. doi:10.1108/14757700610668958

10. Appendices

Appendix A

Glossary

Price and Value:

APV: Adjusted Present Value.

BVA: Book Value of Assets.

BVE: Book Value of Equity.

EV: Market capitalization plus total debt minus cash and short-term investments.

P: Share price.

P^{eq}: Price of equity.

MtB Ratio: Market-to-Book ratio.

PE Ratio: Price Earnings ratio.

PEG Ratio: Price/Earnings-to-Growth ratio.

Variables:

DCF: Discounted Cash Flows.

EPS: Earnings per share.

FCFE: Free Cash Flow to Equity.

Market Cap: The product of the number of shares outstanding and the share price.

NI: Net Income.

NIF: I/B/E/S estimation of 1-year forwards net income.

ROA: Return on Assets.

ROE: Return on Equity.

TA: Total Assets.

Additional Abbreviations:

Big Four: refer to the four largest professional services networks in the world. It comprises Deloitte, Ernst & Young, KPMG, and PricewaterhouseCoopers.

CCV: Contingent Claim Valuation.

CFA: Chartered Financial Analyst.

FF: Fama-French industry groupings.

GICS: Global Industry Classification Standard.

HM: Harmonic Mean.

IPO: Initial Public Offering.

M&A: Mergers and Acquisitions.

MED: Median.

NAICS: North American Industry Classification System.

OECD: Organisation for Economic Co-operation and Development.

R&D: Research and development.

RV: Relative Valuation.

SIC: Standard Industrial Classification.

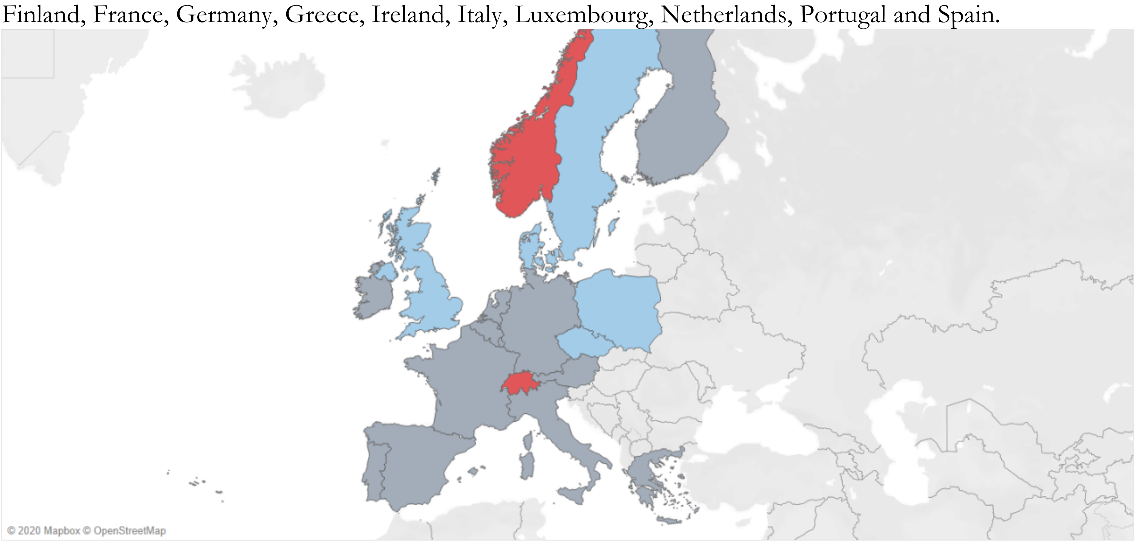
TRBC: Thomson Reuters Business Classification.

Appendix B

INTEG Definition

Figure 5 – *Economic Integration*

This figure represents the countries which enjoys similar degrees of economic integration considered for the comparable criteria **INTEG**. The INTEG variable considers three degrees of integration, these are: the European Union, the European Single Market and the Euro Zone. The **European Union** comprises the following countries: Austria, Belgium, Czech Republic, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Poland, Portugal, Spain, Sweden and United Kingdom. The **European Single Market** includes the following countries: Austria, Belgium, Czech Republic, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Norway, Poland, Portugal, Spain, Sweden, Switzerland and United Kingdom. The **Euro Zone** is composed by the following countries: Austria, Belgium, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal and Spain.



Map based on Longitude (generated) and Latitude (generated). Colour shows details about integration.

Degree of Economic Integration of countries constituing the sample
- Member of the Single Market
- Member of the European Union and Single Market
- Member of the European Union and Single Market and adopted the Euro

Appendix C

Multiple Biasedness

Figure 6 *– Interquartile dispersion of absolute error by comparable set*

This table displays the Relative Valuation errors interquartile dispersion by comparable criteria. This figure presents the bias of the signed prediction error variable. The distribution is centred around zero and slightly skewed to the right.