UNIVERSIDADE CATÓLICA PORTUGUESA

Equity valuation using accounting numbers

An assessment of earnings value relevance

Final Work on Dissertation mode
presented to Universidade Católica Portuguesa
for the obtainment of the master degree in Business Economics

by

Diogo Filipe de Abreu Marques

supervised by
PhD Paulo Alves

Católica Porto Business School
September 2018

Acknowledgments

First, I would like to thank my thesis advisor, Professor Paulo Alves, for all the knowledge shared, patience and support provided.

I also want to thank all my master's Professors both in Católica Business School and Lancaster University Management Shcool for the insights which made this project possible.

I would also like to particularly thank Mariana for all the relentless support and for unconditionally believing me throughout the last years, Dr. Manuel Firmino for helping me when I needed the most and my godmother Paula, for her tenderness and patience.

To my good friend Dimitrios Fountas, for all the availability for the endless hours of discussion and exchange of ideas.

Last but not least, my parents, this is end of a journey and the start of a new one, thank you for trusting me unconditionally.

Finally, to all those who have supported and inspired me directly or indirectly through my academic path, a big thank for being part of this, you made the journey much more pleasant.

Abstract

Asset valuation is not a straightforward task, whether we are trying to value an entire firm or its equity, assumptions are required, and in some circumstances they are not more than just an educated guess. Equity valuation has been more used and criticized than ever, with the examples from Enron, Comcast and recently Facebook reminding us that market prices should not be taken as absolute. Notwithstanding, the recent openness for seed, venture and public capital raises, reinforces the positive outlook on the both private and public equity markets. Consequently, equity valuation arises as necessary tool to assess current and potential value.

Therefore, the purpose of this study is to assess if fundamentals drive equity value, namely earnings, aiming the answer the research question: "Do earnings determine market prices?". This is done through the application of well-known equity valuation models, such as the dividend discount model, residual income valuation model, price to forward earnings and price to book value multiples to a large sample of U.S. firms in order to assess which model performs the best in determining market prices. Additionally a small sample (two firms) is chosen to exemplify the connection between the extent to which earnings determine market prices and the models' value estimates errors. Hence, through the analysis of the results from the previous methods, it was possible to assess that firms in which earnings have a higher/lower explanatory power yield lower/higher errors in model pricing using earnings as the value driver.

Keywords: Equity valuation; Value driver; Value relevance

Table of contents

Table of Figures

Table of tables

Chapter 1: Introduction

1.1. Motivation

Towards the end of my academic path and after deepen my studies in finance, I felt instigated to better understand equity value drivers, compelled by the challenge to perform a study which is fundamentally quantitative, yet sensitive to qualitative valuation matters.

Be that as it may, the fact that equity valuation is "as much an art as it is a science" has attracted me to understand how my perspective of firms as assets yielding returns can be different after understanding the widely-held implications in valuation. In addition, while performing a preliminary study on this topic, I have realised that reconciling my perspective with the academics' and practitioners' one, has brought me an increase sense of awareness of valuation techniques and, therefore, the need to keep studying this topic, in which I hope to have a career.

1.2. Investigation challenge – Objectives and Scope

The purpose of this report is to address the question on how " Do earnings determine market prices?", discussing two approaches, namely static investment decisions models and relative valuation models, commonly designated as, discounted cash-flow and comparable assets methods. Particularly, within the discounted valuation models the scope will tighten and the focus will be on equity valuation models both based in cash and accrual perspectives. The same follows for the comparable assets method, with the use of equity multiples.

1.3. Methodology

The method used to answer the proposed research question will be quantitative and can be split into two sub-sections. Namely, a large sample analysis which consists in a statistical analysis of accounting-based valuation models for a large sample of U.S. firms and a small sample analysis, where a smaller sample of brokers' firm reports will be analysed to provide evidence on the practitioners' valuation assumptions with the purpose to reconcile these insights with the ones from the theoretical framework and large sample analysis.

In more detail, issues to be explored include industry contingencies in the choice of the valuation models, peer choices to which value and performance are benchmarked, inclusion of qualitative and non-financial information in the models and how the information available in the financial statements is used.

Thereafter, regressions are made with the purpose of assessing fundamentals' impact on market prices, namely earnings. These are to be compared to the value estimates yielded by the models in order to answer the research question.

1.4. Dissertation structure

The present dissertation report is mainly divided in five major chapters. Firstly, there is the present section, which relates to the introduction.

Secondly, follows the literature review chapter, materialized as theoretical framework of the topic.

The third chapter corresponds to the presentation of the adopted methodology, in which entails the methodology design, the investigation strategy as well as the data recovery and analysis for both large and small sample.

The results yielded from the large and the small sample analysis are described in the fourth chapter and discussed within chapter five, claiming conclusion to the study and, after all, an answer to the research question.

Chapter 2: Literature review

At the outset of this report on equity valuation using accounting numbers is the purpose to assess earnings value relevance.

In an earlier stage, it is intended to provide the reader a conceptual basis on which the valuation literature is built upon. Succeeding, follows a literature review on the main streams of thought on equity valuation models, thereafter easing the task for both reader and writer to portray the answer to the research topic. Further, the text dives into the fundamentals of equity valuation models applying these to both a large and small set of data. At last, from the former analysis my intent is to yield a conclusion on whether earnings determine market prices.

2.1. Contextualization

Before diving into further detail of valuation models, it is important to understand the core concepts in evaluating equity and their implications. First, there is an extent discussion where intrinsic value is recurrently mentioned. So what is intrinsic value? According to Damodaran (2005; p. 5) "Consider it the value that would be attached to an asset by an all-knowing analyst with access to all information available right now and a perfect valuation model." Undoubtedly, intrinsic value itself is useless as it is a theoretical construction, however it can help defining a benchmark comparison to market pricing.

Second, valuation is inherently prospective and interdisciplinary, relying on an expectation of future cash flow generating events and spans knowledge from accounting and finance to marketing, economics and corporate strategy. The fact that is prospective also entails that is tainted by uncertainty.

Third, accounting systems are crucial to valuation, as they deeply affect financial statements. Accounting systems can change both in time and across firms, and as systems often include options. Understanding these is crucial to assess the validity of earnings forecasts post to their announcement and, more broadly speaking, to understand all the financial dimension of the firm.

Fourth, valuation models are merely pro-forma financial statements which guide the user through the information to be forecasted and its conversion into a value estimate. According to Lee, Myers & Swaminathan (1999) "It is the forecast that breathes life into valuation models." and "In making this forecast, we specify a valuation model, and we predict its parameters with the aid of fundamental analysis."

Fifth, fundamental analysts are driven by the purpose to perform better forecasts, through the use of financial, non-financial, industry-wide, macroeconomic, competitive trends, input and output markets information.

Overall there are four main approaches to valuation: first, discounted cash flow valuation, which links the current value of an asset to the present value of expected future cash flows. Second, liquidation and accounting valuation are underpinned by accounting estimates or book value as a base line. Third, relative valuation approaches compute the value of an asset by looking at trading prices of assets similar in growth, risk and cash flows. Fourth, contingent claim valuation techniques use option pricing models to measure the value of assets that share option characteristics, commonly named as real options. As a note, the latter will not be discussed in this paper, due to topic constraints.

2.2. Absolute valuation

2.2.1. Discounted cash flow valuation

Behind the name, the bottom line in discounted cash flow methods is that an asset with higher and less volatile cash flow should yield a greater value than an asset with lower and more volatile cash flows. Bohm-Bawerk (1903) and Marshall (1907; *cit in* Damadoran, 2005) laid the bedrock for the discounted cash flow valuation with the concept of present value as exemplified in a household purchase with 20 annual instalments (Bohm-Bawerk, 1903; *cit in* Damodaran, 2005).

Withal, modern valuation principles were developed by Fisher & Barber (1903) and Fisher (1930), which contributed with four alternative investment decision methods – first, chose the one with the highest present value at the market interest rate; second, the investment with the present value of benefits which exceeds the present value of the costs the most; third, the rate of return on sacrifice that most exceeds the market interest rate or, fourth, that when compared to the most costly investment, yields a rate of return over cost that exceeds the market interest rate. However, other authors such as Boulding (1935), who computed the internal rate of return for an investment as a function of its cash flows and initial investment; Keynes (1937), who computed the rate of return that would equalize the investment return to its initial price; or Samuelson (1937), who assessed the differences amongst Fisher's and Keynes's theories, to conclude that an investor should maximize the present value and not the internal rate of return.

According to Damodaran (2005), there are four variants of the discounted cash flow models. In the first place, consider a cash flow generating asset, which is discounted using a risk adjusted rate, followed by an approach where the cash flow themselves are risk adjusted and then discounted at the risk-free rate to yield the value of the risky asset. Additionally, one can value a business, in an initial stage not considering the effects of debt, but later incorporating them into the model. Finally, an asset can be valued considering the returns it generates over or under what is normal, generally named as excess returns.

2.2.1.1. Discount rate adjustment models

Equity discounted cash flow models

This section, is entirely focused on the investor's stake value in a business, recurring to risk adjusted rates to discount cash flows. The following sub-sections will approach different perspectives of what can be considered as cash flows. In the first place we only consider dividends to be cash flows, followed by the incorporation of stock buybacks as cash flows to equity and finally other potential dividends or free cash flows to equity (Damodaran, 2005).

Dividend discount model

The dividend discount model's characteristics are easily understood and although throughout the years, less analysts have been using it, the principles that guide this approach still lay the bedrock for other free cash flow to equity models. Durand (1957), Dodd & Graham (1934; *cit in* Damodaran, 2005) and Williams (1938; *cit in* Damodaran, 2005) have agreed on the proposition that expected price is determined by future dividends. Accordingly, the investor can receive two types of cash flows from an asset capital gains from increases in price and dividends received during the holding period. While the original model only considers dividends to be the only cash flow generated by the firm that investors can lay claim, other authors have argued that this approach is too narrow and unrealistic, yielding conservative estimates. As stated in the previous section, forecasting is a core task in valuation, which the original dividend discount model over-simplifies, only demanding an estimate of dividend growth to compute the model. This model was intensively used by Gordon (1962; *cit in* Damodaran, 2005), nonetheless the former, acknowledged its applicability is constrained to constant growth firms. Consequently, variants were developed, which consider different growth stages, and therefore overcome the previous caveat. In detail, these variants consider different growth rates in the initial years, which later converge to a stable growth rate. Fuller & Hsia (1984) went further and developed a model which includes risk and dividend payout to derive simpler high growth models based on the assumption that high growth rates decline over the extraordinary growth phase in a linear way, while cost of equity and payout ratios remain constant. However, this approach is not flawless, and some authors argue that when growth rates tend to decrease, payout ratios tend to increase, limiting this model's applicability.

Irrevocably, the dividend discount model has its advantages and caveats, namely, the model is blind to cash balances and the investors' claim on its value. Leading to an undervaluation of firms with large and increasing cash balances. The opposite is also true, in the sense that some firms finance their dividends through equity and/or debt issues.

Extended equity valuation models

Dividends vs. Potential dividends

To state the difference between dividends or potential dividends it is necessary that one understands the use of such mechanisms. Whether is due to smaller firms being uninterested in paying dividends (DeAngelo, 2004), an increasing number of investors not interested in receiving dividends (Baker & Wurgler, 2004a; 2004b) or increase firm specific risk (Hoberg & Prabhala, 2005; *cit in* Damodaran, 2005), there were authors who stated that the gap between potential dividends and actual dividends is increasing over time. Such acknowledgment brought difficulties to those using the dividend discount model. Thereafter, solutions considering other cash flows to equity other than just dividends were developed.

Buybacks as dividends

According to the dividend discount model section before, the investor can receive two types of cash flows. Likewise, firms found stock buybacks as an instrument to return cash to the stockholders, which originally would be missing out from the dividend discount model. These can be incorporated into the model by adding an average four to five year stock buybacks to smooth onetime events and subtracting long term debt issues potentially used to finance the buyback.

Free cash-flow to equity

Free cash flow to equity is the cash flow left over after all reinvestment needs and debt payments (Damodaran, 2005). Although assuming all the free cash flow to equity is paid to investors implies two things, first, that there will be not cash build up in the future, and that the expected growth in the free cash flow to equity is directly related to the growth in income from operating assets and not from marketable securities.

Earnings models

Both situations previously mentioned - firms who do not pay out their free cash flow to equity as dividends and difficulties in estimating cash flows -, induced some authors to defend that a firm is best valued through discounting earnings or variants of earnings (Feltham & Ohlson, 1995). Penman & Sougiannis (1997) argue that GAAP earnings can be replaced by dividends given that adjustments are made to prospective earnings and book value to reflect the latter.

Potential dividends vs. Dividend discount models

The difference from using potential or actual dividends is the value for corporate control. On one hand, if the management is likely to be changed, then the free cash flow to equity might better reflect the equity value of the firm. On the other hand, if it is not, and there are legal and market constraints to corporate takeovers then dividend might yield a fair estimate of value. Moreover, the value the model yields can be used as a base line for the valuation (assuming the cash is not being wasted in value destructing projects). It is a fair estimate of value when the firm pays out all their free cash flows to equity as dividends, and in sectors where cash flows are difficult or impossible to estimate (such as banks), dividends are a simpler measure to be forecasted.

2.2.1.2. Excess return models

The previous models are the most commonly used, but there are variants such as the excess returns approach, where we separate the cash flows into excess return and normal return cash flows. Normal cash flows earn a rate of return equal to the cost of equity or capital, accordingly to equity or firm valuation. Any additional cash flows are considered to be excess cash flows and can be either positive or negative. From the former we can write the business value yielding equation as value of a business to be equal to its capital invested today plus the present value of its excess return cash flows from existing and future projects.

Using book value of capital as a proxy for the capital invested in assets today implies that excess cash flows will trade at market values higher than their book values, and vice-versa for companies who yield negative excess cash flows.

Basis for the models

The basis of these models relies on principles of capital budgeting and net present value, and the assumption that a project only adds value to a business if it has a positive NPV is always present.

In line with the above, it should be borne in mind that earnings and cash flow growth are only valuable when they are accompanied by excess returns - a situation in which shareholders' equity exceeds the cost of equity. Thus, the excess return models are based on this conclusion to support the next calculation step for a company, in function of expected surplus returns.

Notwithstanding the fact that there are several possible versions to adopt when referring to models of excess return, one of the most widely considered and used is the economic value added, from now on EVA, also adopted in this context. Thus, EVA measures the added value created by an investment or by an investment portfolio, being calculated through the product between the excess return obtained in one or more investments and the capital invested in them.

Aggregate economic value = (Return on invested capital - Cost of capital) (Invested capital) = Operating revenue after taxes - (Cost of capital) (Capital invested).

Finally, EVA can be seen as an extension of the net present value rule, once the net present value of a project corresponds to the present value of the economic value added by the project over its useful life. In this sense, EVAt is the economic value added by the project in year t, and this same project has a useful life of n years and kc is the cost of capital.

Measuring Economic Value Added

EVA requires 3 inputs: 1) return on capital earned on investments; 2) the cost of capital for those investments; 3) the capital invested in them.

Notwithstanding, one has to account for distortions and mis-categorizations.

To assess how much capital is invested in existing assets given the difficulty in measuring all assets at market value, and that the market capitalization considers not only the assets in place but expected growth as well, we turn instead to book value of capital as a proxy of capital invested. However, as previously mentioned, the book value of capital incorporates present and past accounting choices, which can distort the true value. This would imply an endless number of adjustments, on which accounting literature has focused and this text will not. Nonetheless, the most efficient solution that was found in cases where the number of adjustments is too high, is to estimate the capital invested from scratch. Starting with the assets owned by firm, estimating the value of those assets and cumulating their value over time. To measure the return on capital invested, an estimate of the after-tax operating income is necessary along with adjustments for R&D expenses, operating leases and one-time charges to yield a true measure of the firm's operating earnings. The last component is the cost of capital. Ultimately this should be calculated using the market values of debt and equity. Using book value cost of capital tends to understate the WACC for most firms and with greater impact for firms with higher than for those with lower leverage. The understatement of the cost of capital will cause the overstatement of the EVA.

2.2.2. Liquidation and accounting valuation

Going concern valuations methods differs from asset-based valuations, and for firms with profitable growth opportunities, leading the second to yield lower valuations than the first.

2.2.2.1. Book value based valuation

Some authors argue that accounting-based valuation methods are more reliable than methods which rely on shaky assumptions about the future. In the following section, an analysis of book value as a measure of going concern is extended to forecasted earnings based on book value valuation methods.

Book value

Daniels (1934 *cit in* Damodaran, 2005; p. 54) wrote that "In short the lay reader of financial statements usually believes that the total asset figure of the balance sheet is indicative and is intended to be so, of the value of the company. He probably understands this value as what the business could be sold for, the market value the classic meeting of the minds between a willing buyer and seller."

Evidence can be found that book value is a good proxy of market value, when Graham (1949; *cit in* Damodaran, 2005) proposed using book value as benchmark to find undervalued stocks. Academics reiterate this evidence by stating that stocks with low price to book value earn in average higher returns than the rest of the market. On one hand book value yields a reasonable measure for the true value of a business if a firm is mature, has little or no growth opportunities and no potential for excess returns. On the other hand, for firms with significant growth opportunities in businesses where they can generate excess returns, book values will be very different from true values.

Book value plus earnings

Ohlson (1995) states that the true value of equity is a function of its book value of equity and the excess equity returns that the firm can generate in the future. As a consequence this is termed a residual income model and can be derived from a simple dividend discount model:

Value of equity = $SUM(E(Dividends_t)/(1 + Cost\ of\ Equity_t)^t$

Now substitute in the full equation for book value of equity as a function of the starting book equity and earnings and dividends during a period (Clean Surplus Relationship):

Book value of $equity_t$ = BV of $equity_{t-1}$ + Net $income_t$ – $Dividends_t$

Substituting back into the DDM, we get

Value of $equity_0$ = BV of $equity_0$

$$+\sum_{t=1}^{t=\infty} \frac{(\text{Net Income}_t - \text{Cost of Equity}_t * \text{BV of Equity}_{t-1})}{(1+\text{Cost of Equity})^t}$$

Dechow *et al.* (1999), Frankel & Lee (1998) and Hand & Landsman (1999; *cit in* Damodaran, 2005) all find that the residual income model explains 70-80% of variation in prices across stocks. The test whether changes in equity value are correlated with changes in book value of equity and net income show that the model performs as well as the other established models.

2.2.2.2. Liquidation valuation

Liquidation valuation is likely to be the most accurate market value of a business. Once, it represents the current state of the economy, the hurry to dispose the asset and its potential buyers. Consequently, it underestimates its value, since its does not consider future growth prospects, opposite to the discounted cash flow methods. Additionally, the assets tend to be alienated at a discount from their book value, representing the eagerness of the seller to dispose it.

Damodaran (2005) states that liquidation valuation is likely to yield more realistic estimates of value for firms that are distressed, where the going concern assumption underlying conventional discounted cash flow valuation is clearly violated. For healthy firms with significant growth opportunities, it will provide estimates of value that are far too conservative.

2.3. Relative valuation

In relative valuation methods, an asset is valued by comparison with other similar assets in the market. This description embeds three fundamental steps in valuation. First, finding comparable assets that are priced by the market, a task that is easier to accomplish with real assets such as real estate than with stocks. Analysts tend to use companies in the same sector as comparable, however, this method raises the question whether they are truly comparable. The second step is calling market prices to a common variable to generate standardized prices that are comparable. While this may not be necessary when comparing identical assets, it is necessary when comparing assets that vary in size or units. *Ceteris paribus*, a smaller house or apartment should trade at a lower price than a larger residence. In the context of stocks, this equalization usually requires converting the market value of equity or the firm into multiples of earnings, book value or revenues. The third and last step in the process is adjusting for differences across assets when comparing their standardized values. Again, using the example of a house, a newer house with more updated amenities should be priced higher than a similar sized older house that needs refurbishments. With stocks, differences in pricing stocks can be attributed to all of the fundamentals that we talked about in discounted cash flow valuation. Higher growth companies, for instance, should trade at higher multiples than lower growth companies in the same sector. Many analysts adjust for these differences qualitatively, making every relative valuation a story telling experience; analysts with better and more believable stories are given credit for better valuations (Damodaran (2005)).

2.3.1. Standardized values and multiples

Comparing assets that are not exactly the same can be a challenge, as the value of the firm is a function of both the value of the equity and the number of shares outstanding. Thus, something as a stock split 1/2 will half the price of the stock. To compare firms, it is necessary to establish some common variable, or standardize a measure of value.

There are several ways this can be performed: first, is to consider the value of the asset as a multiple of its earnings related with price (PE). Still, earnings used in the denominator do not necessarily have to be present, and one can chose whether to use past earnings (trailing PE) or forecasted analyst consensus (forward PE). Second, accountants tend to compare the market price with its book value, in order to assess over or under valuations (value investing). For those which do not consider the book value to represent the true value of the assets, it can also be used the replacement cost of assets or Tobin's Q. Third, some researchers argue that sales are not affected by changes in accounting policies, rules and principles and so are a fairer basis of comparison between firms, especially if not within the same sector. Thus, the price to sales ratio is widely used to compare firms from different industries, however it is necessary to pay attention that depending on benchmark profit margins from each sectors, the multiple varies widely across sectors, causing increased sensitiveness to peer choices. Fourth, earnings, book value and revenue multiples are those that can be computed for firms in any sector and across the entire market, however there are some multiples that are specific to a sector. Although these can be biased and yield consistent over or under valuations for an entire sector, also they add complexity in their interpretation making the investor more reluctant to pay the price. Additionally, these multiples are also difficult to connect to the firm's fundamentals, which is essential to the good use of multiples.

There are relatively few studies that compare the performance of multiples, although one can assess trends in sectors or practitioners, such as the use of EV/EBITDA dominance of valuations on heavy infrastructures businesses such as cable and telecom, Morgan Stanley research arm refers as the most popular both PE and EV/EBITDA (Damodaran, 2005). Liu *et al*. (2002) compared the performance of multiples in pricing 19879 firm-year observations between 1982 and 1999 and suggest that multiples of forecasted earnings per share do best in explaining pricing differences, that multiples of sales and operating cash flows do worst and that multiples of book value and EBITDA fall in the middle. Lie & Lie (2002) examine 10 different multiples across 8621 companies between 1998 and 1999 and arrive to similar conclusions.

2.3.2. Determinants of multiples

The determinants of the multiples are growth, risk, potential cash flow generation. Similarly, to the discounted flows models firms with higher growth rates, lower risk and higher cash flow generating potential should trade at higher multiples than firms with lower growth, higher risk and lower cash flow generating potential

2.3.3. Comparable firms

Value standardization is performed for a group of comparable firms, but what is deemed comparable? The three factors previously discussed underpin the comparable concept. Given that, two firms with similar growth, risk and cash flow generating capacities are deemed comparable, additionally one would want to compare two exactly identical firms in terms of these factors. On one hand, nowhere in the definition of comparable is a suggestion that industry or sector of firm placement legitimates the multiple analysis. On the other hand, practitioners tend to use firms in the same business or businesses, due to similarities in these factors, to perform the comparison. Bhojraj & Lee (2002), Bhojraj *et al*. (2003) and Cheng & McNamara (2000) argue that picking comparables using a combination of industry categorization and fundamentals such as total assets yields more precise valuations than using just industry classification.

2.3.4. Controlling for differences across firms

It does not matter how much carefully we build our list of comparable firms, we will end up with firms that are different from the firm we are valuing. The differences may be small on some variables and large on others and we will have to control for these differences in a relative valuation.

There are three ways of controlling these differences, described in the next subsections.

2.3.4.1. Subjective adjustments

Relative valuation usage implies two choices: the multiple used and the group of comparable firms. In the majority of relative valuations, the multiple is calculated for each of the firms and then averaged. However, the researcher should question what is the best way to compute this average. Beatty *et al*. (1999) examine multiples of earnings, book value and total assets and conclude that the harmonic mean provides better estimates of value than the arithmetic mean. To evaluate an individual firm, the analyst then compares the multiple it trades at to the average computed. If there are significant differences, the analyst uses its judgement to assess if that is due to differences in the previously stated factors. Notwithstanding, in case the difference cannot be explained by the fundamentals, the firm is considered to be overvalued (undervalued) if the multiple is above (below) the average. Subsequently, the result of this subjective approach to explain the difference between the individual and group multiple is found to be biased by the analyst's cognitive matrix.

2.3.4.2. Modified multiples

This approach implies the identification of the variable that has the greatest impact on the multiple – the companion variable. An example of such adjustment is dividing the PE by the growth rate, yielding a growth-adjusted PE ratio or PEG ratio. Nonetheless, the use of modified multiples implies two things. First, assuming that these firms are comparable on all the other measures of value, other than the one being controlled for. For instance, when comparing PEG ratios across companies, we are assuming that they are of equivalent risk. If some firms are riskier than others, you would expect them to trade at lower PEG ratios. Secondly, that the relationship between multiples and fundamentals is linear. Given that growth doubles would imply the PEG ratio to double, if this condition does not hold, then PE ratios do not increase proportional to growth, and companies with high growth rates will look cheap on a PEG ratio basis.

2.3.4.3. Stastical techniques

Sector regressions

Both subjective adjustments and modified multiples are difficult to implement given the complex nature of the relationship between multiples and fundamentals. However, there are statistical techniques that offer alternatives.

In a regression, it is attempted to explain a dependent variable, reflecting what we believe influences the dependent variable. Similarly to what we are attempting to do in relative valuation, where we try to explain differences across firms on a multiple (P/E, EV/EBITDA) using fundamental variables.

The main advantages of regressions over subjective approaches are:

- The regression output yields a measure of how strong the relationship is between the multiple and the variable being used.
- If the relationship between a multiple and the fundamental we are using to explain it is nonlinear the regression can be modified for the relationship.
- Differently from the modified multiple approach where we were able to control for differences on only one variable, a regression can be

extended to allow for more than one variable and even cross effects across these variables.

Market regressions

Searching for comparable firms within the sector in which a firm operates is fairly restrictive, especially when there are relatively few firms in the sector or when a firm operates in more than one sector. Since comparable and industry peer are different concepts, we do not need to restrict our choice of comparable firms to those in the same industry. The regression introduced in the previous section controls for differences on those variables that we believe cause multiples to vary across firms. The fundamentals that drive PE, PBV and PS ratios are summarized in figure 1, yet proxies to these variables, such expected growth in eps (growth), beta (risk) and payout ratio (cash flow) are necessary to perform the regression and imply the assumption of a non-linear relationship between the former and the multiples. To overcome this limitation, the researcher can add further variables to the model, for instance firm size as a proxy for risk.

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

Figure 1: Fundaments determing equity multiples

Source: Damodaran, 2005

The advantages of the market wide approach over the subjective comparison across firms within a sector, is that the relationship between fundamentals and the multiples is quantified. Nonetheless it contains errors, which are a part of reality that the analysts chose to ignore in their approach. Moreover, the difficulty in evaluating firms in sectors with few firms is then overcome using market wide approaches. Lastly, it allows a comparison of overvaluation/undervaluation of all firms within the market.

Evidence from Damodaran (2005) tells us that the explanatory power of a regression of fundamentals on the PE ratio decreases over time, as a result of the earnings volatility. Furthermore, the author states that the PBV and PS ratio have yielded higher R^2 than the PE ratios.

The limitations of statistical techniques are distributional properties of the multiples that we talked about earlier in the chapter and the relationship among and with the independent variables used in the regression.

For instance, the distribution of multiples across a population is not normal for a clear reason, most multiples are restricted from taking values below zero but can assume very large positive values. Using standard regression techniques pose a problem in these non-normal distributions which are accentuated using small samples, consequently magnifying the asymmetry of the distribution due to a few large outliers.

In a multivariate regression model, we implicitly assume that independent variables are as such, not related to one another. However, once we are using growth, risk and cash flows, it is highly likely that they are correlated. Given this we will have a multicollinearity problem which further undercuts the explanatory power of the regression.

In addition to non-normal distribution of the multiples, we also have that their distribution change over time, making it difficult or pointless to compare firms over time, due to decreases in the explanatory power of the regression.

As a final note of caution, the R^2 on relative valuation regression will almost never be higher than 70% and it is common to see it drop to 30% or 35%. Rather than ask the question of how high an R^2 has to be to be meaningful, we should focus on the predictive power of the regression. When the R^2 decreases, the ranges on the forecasts from the regression will increase.

2.4. Value Relevance

Value relevance studies are often conducted with the purpose of providing insights into questions of interest to standard setters and practioners. Accordingly, one of the primary focus of standard setters is equity investment, as such the use of valuation models to assess value relevance is admissible despite the models' simplifying assumptions. The topic of value relevance in terms of research may encompass and state the differences between conservatism and optimist forecasts. Additionally, studying this theme may also contribute to understand the relation between accounting amounts and equity values. Lastly, value relevance studies are designed to evaluate the relation between information used by equity analysts and particular accounting amounts used to value firms' equity.

"In the extant literature, an accounting amount is defined as value relevant if it has a predicted association with equity market values" (Barth, Beaver and Landsman, 2001; p. 79) The main users of value relevance literature are academic researchers, although these may be useful to analysts' intending to understand the significance and trustworthiness of accounting figures reflected in equity values.

The FASB's conceptual framework uses relevance and reliability as the baseline criteria to choose between accouting alternatives. As such it is deemed necessary to address as a bottom line in this study the search for the value relevance of earnings in explaining stock returns and their relation with equity valuation using earnings, models' performance.

2.5. Research question and assumptions

2.5.1. Research question

The literature review and the discussion of the possible approaches to its gaps opens the way and guide a research essay addressing the question on *How do earnings determine market prices?*

2.5.2. Assumptions

The main assumption deduced from this literature review is a general one and states that market is efficient (Fama (1991)). Moreover, there is the fact that differences in the value estimates are driven by fundamentals or unknown factors.

Secondly, there are some conceptual definitions that help to break down down the research question and guide the reader towards the purpose of this study, as made below:

Equity valuation:

Process of converting accounting forecasts into an estimate of the value of the firm. Implies the forecasting of financial statement data into perpetuity, the choice of the accounting parameter to be discounted and the appropriate discount rate (Mazzariol & Thomas, 2016).

Extension:

Even though there are several definitions for the "extension" concept, due to the fact that this single word has different meanings according to the context in which it is used;,the one which deserves attentions in this scope is the logical one. So, in this context, the word "extension" must be seen as "the range of a term or concept as measured by the object which it denotes or contains" (English Oxford Living Dictionaries).

Earnings:

A company's profit divided by its number of common outstanding shares. If a company earning $2 million in one year had 2 million common shares of stock outstanding, its EPS would be $1 per share. In calculating EPS, the company often uses a weighted average of shares outstanding over the reporting term. The one-year (historical or trailing) EPS growth rate is calculated as the percentage change in earnings per share. The prospective EPS growth rate is calculated as the percentage change in this year's earnings and the consensus forecast earnings for next year (NASDAQ).

Determine:

Once determine is a verb, the meaning of this word in this context must be assumed as the influence that something has in other variable, specifically meaning the influence that earnings have in market prices – defined below.

This assumption of meaning is supported in the one presented in the dictionary, cited "Cause (something) to occur in a particular way or to have a particular nature" (English Oxford Living Dictionaries).

Market prices:

The amount of money that a willing buyer pays to acquire something from a willing seller, when a buyer and seller are independent and when such an exchange is motivated by only commercial consideration (NASDAQ).

Chapter 3: Methodology

3.1. Methodology design

The current section the purpose is to present the methodology design adopted to explore and test the research question, formulated according to the literature review, core to the study, once the main purpose is to answer it.

The research adopted an empirical quantitative observation of what is stated in the literature review, through the use of a large sample of data. Furthermore, to complement the empirical study, follows a narrower sample of data – small sample – to portray the connection between earnings value relevance and models' value estimates pricing errors.

From the analysis and relation between the large and small sample results the answer to the research question.

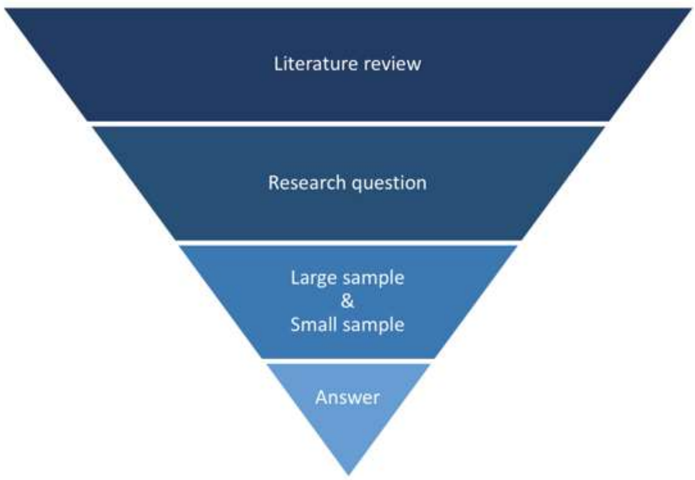

Figure 2: Methodology design

3.2. Research strategy

Conferring the above comparison between a quantitative and a qualitative strategy of investigation, it becomes clear that in this case the adopted strategy is mostly quantitative, nonetheless requiring some interpretation of analysts reports on a final stage of the study. However, this qualitative part has no expression on the global investigation, when compared with all the rest of the study which is quantitative.

In terms of assumptions - namely because *variables are identifiable and relationships measurable* - purpose - once its based on *casual explanations* and in a *deductive experimental* style of testing hypothesis -, approach -since it *reduces data to numerical indices* – and researcher role which aims *detachment and impartiality* – all points out to a quantitative strategy of investigation in terms of methodology.

	Quantitative	✔	Qualitative	✔
Assumptions	Reality s objectively given (objective)	✔	Reality is socially constructed (subjective)	
	One truth	✔	Multiple truths	
	Usually associated with positivism	✔	Usually associated with interpretivism	
	The method is the focus	✔	The subject is the focus	
	Variables are identifiable and relationships measurable	✔	Variables are complex and often not measurable	
Purpose	Generalisation	✔	Contextualisation	
	Prediction	✔	Interpretation	
	Test a theory or hypothesis	✔	Derives hypothisis and ground theory	
	Casual explanations	✔	Understanding different perspectives	
	Deductive / Experimental	✔	Inductive / Descriptive	
	Explanatory: What are the relationships between variables?	✔	Exploratory: What are the variables involved?	
Approach	Starts with hypothesis and theories	✔	Starts with research questions	
	Ends with casual laws	✔	Ends in hypothesis and ground theory	
	Uses formal instruments	✔	Researcher is the primary instrument	
	Component analysis	✔	Searches for patterns	
	Seeks consensus	✔	Seeks pluralism	
	Reduces data to numerical indices	✔	Makes minor use of numerical indices	
Researcher Role	Detachment and impartiality	✔	Personal involvement and partiality	
	Objective portrayal	✔	Empathic understanding	

Figure 3: Research strategy: Quantitative vs. Qualitative

Source: Glesne & Peshkin (1992); Healy & Perry (2000); Robson (2002); Hancock *et al.*, (1998).

3.3. Large sample

According to Richardson & Tinaikar (2004), the best model is the one that minimizes the pricing error (price – intrinsic value), which implies that the market is pricing the asset correctly (efficient market hypothesis). However, the forecasting branch of research has been focusing on the matter that prices can deviate from the fundamentals in the short term, moving towards them in the long-term. As such, valuation models have been a way to yield a measure of the intrinsic value and consequently of help to identify pricing errors.

3.3.1. Data selection

First of all, the current subsection's purpose is to portrait the sample chosen and levels of information to be gathered. Accordingly, due to a large amount of available information, the sample chosen comprises exclusively U.S. public firms across a wide range of industries including financial firms. The firm-level information is organized in four broad categories:

1) General information describing firm names and industry classification.
2) Firm-level accounting data on sales, earnings, assets and common stakeholders' equity.
3) Analysts' forecasts
4) Market pricing data, such as stock price four months after the fiscal year end and betas.

Category (1) and (2) information were retrieved from Compustat®, (3) from I/B/E/S and (4) from CRSP.

3.3.2. Data analysis methods

Following the former subsection, the present will entail the description of the variables and procedures implemented to the sample. The methods can be divided into two stages: preliminary analysis, with the purpose to assess different model properties and factor sensitivities. Lastly, the performance measurement analysis, will use the value estimate produced by the models to create a proxy for performance, in this manner allowing the ranking of firms by model performance.

As per the previous paragraph, the variables used throughout the modelling process are ticker symbol (ticker), company name (comnam), fiscal year (fyear), firm specific adjustment factor, cumulative by ex-date (ajex), book value per share (bkvlps), total common equity (ceq), common shares outstanding (csho), common dividends (dvc) stock price four months after fiscal year end (prc4), basic current earnings per share excluding extraordinary items (epspx), analysts' forecasts of earnings per share in year t+1 to t+5 (eps_1 - eps_5), analysts' forecasts of dividends per share in year t+1 to t+5 (dps_1 - dps_5), respectively earnings and dividends long-term growth rates (eps_g and dps_g) close the basic set of variables used in the four models. Any additional variables relevant to the study will be presented and described before its mention and use.

3.3.3. Data

This section's purpose is to portrait the performance of the residual income and dividend discount models as well as the price to forward earnings and price to book value multiples. The initial dataset described in the appendix I[1] totalizes 37106 observations from 2004 to 2016 as in table 1, although additional adjustments are performed to standardize the information due to model related choices and missing figures, to be detailed ahead.

The first trimming procedure consists in the exclusion of financial firms due to their complexity in using the models previously chosen. Follows an exclusion

[1] Appendix I – Initial dataset before preparation

of all firms that have less than 10 observations of t+1 and t+2 forecasted earnings per share. Hereafter the initial dataset is composed of 16209 observations as shown in table 2.

Variable	N	Mean	SD	Min	P1	P25	P50	P75	P99	Max
prc4	16209	40.20	52.14	0.19	1.60	15.72	30.44	50.70	187.06	1661.29
bvpsadj	16209	50.34	2647.63	-1186.62	-5.61	5.63	10.77	18.99	78.48	253503.00
epsadj	16209	-14.70	1379.63	-126900.30	-12.90	0.29	1.28	2.48	11.00	39750.10
eps_1_adj	16209	5965.73	2871722.00	-192000000.00	-6.15	0.60	1.51	2.84	20.70	290000000.00
eps_2_adj	16209	21452.16	2796043.00	-62300000.00	-3.85	0.81	1.80	3.26	28.00	328000000.00
dps_0_adj	16209	0.6322814	2.238623	-0.0007406	0	0	0.1909326	0.7999839	4.882491	159.2022
dps_1_adj	16209	28.43	2848.77	0.00	0.00	0.00	0.17	0.80	5.21	359999.30
dps_2_adj	16209	73.52	6333.54	0.00	0.00	0.00	0.18	0.84	5.45	719998.60

Table 1: Descriptive statistics before trimming

Observing the table above we can state that in a preliminary stage the the data has outliers which will bias the analysis and should be accounted for. For such, two adjustments are made, first the variables ceq, epspx, eps_1 and eps_2 are adjusted for stock splits resulting in a new set of variables, bvpsadj, epsadj, eps_1_adj and eps_2_adj to be used from now on. Second, an attempt of eliminating outliers has been processed in every variable eliminating the top and bottom 1% of the data for these variables.

Variable	N	Mean	SD	Min	P1	P25	P50	P75	P99	Max
prc4	14222	34.28	24.46	0.68	2.01	15.79	29.21	47.13	113.00	148.52
bvpsadj	14222	12.48	9.61	-4.44	-1.13	5.56	10.15	17.27	43.79	64.66
epsadj	14222	1.18	1.97	-11.35	-5.74	0.31	1.21	2.22	5.71	7.32
eps_1_adj	14222	1.63	1.55	-10.15	-1.73	0.58	1.40	2.52	5.88	7.74
eps_2_adj	14222	1.97	1.63	-2.02	-1.08	0.78	1.68	2.88	6.60	7.26
dps_0_adj	14222	0.43	0.57	0.00	0.00	0.00	0.16	0.69	2.23	3.12
dps_1_adj	14222	0.42	0.57	0.00	0.00	0.00	0.15	0.69	2.25	2.78
dps_2_adj	14222	0.45	0.60	0.00	0.00	0.00	0.16	0.72	2.37	2.61

Table 2: Descriptive statistics after trimming

Subsequently, overall improvements in the sample are obtained, such as lower standard deviation and values in the minimum, maximum and edge percentiles (bottom -P1 and top-P99, one percent). To be noticed that the minimums, p1 and p25 are not exactly but approximately zero.

3.3.4. Preliminary analysis

3.3.4.1. Absolute valuation models

The purpose of this subsection is to perform a sensitivity analysis on the properties of terminal value, namely the growth rate and respective weight on the models' value estimates. Accordingly, the following assumptions underpin the calculations of intrinsic value for the RIVM and DDM - a three year forecast horizon, a cost of equity calculated using the capital asset pricing model yearly Treasure Bills yields and equity risk premiums (respectively named re, rf and rpm in the dataset) (Hillier, *et al.*, 2016), and a constant growth rates of 0% and 4% to perpetuity (Francis *et al.*, 2000).

3.3.4.2. Relative valuation models

As in the previous subsection, the present's objective is the calculation of the multiples - price to earnings and price to book value; to be used in the intrinsic value calculation. The eight approaches used are detailed in figure 4 below and consist solely in modelling the valuation models with two different averaging methods and two different industry definitions for both multiples.

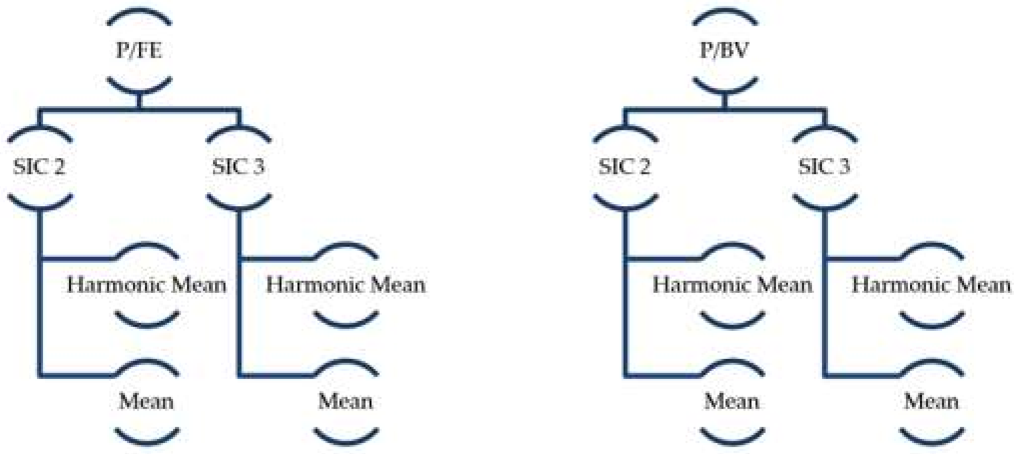

Figure 4: Relative valuation feature analysis

Firstly, according to Liu *et al.* (2000), there is a significant improvement in using the harmonic mean instead of the mean in the calculation of price to value driver. Secondly, according to Alford (1992) and Bhoraj & Lee (2001) the use of both two and three digits from the SIC as the criteria to choose the pool of comparable firms, yield performance improvements. Nonetheless, one could choose the peers pool using industry, risk (measured by firm size), and earnings growth as Alford (1992) published in his paper, although, the author ultimately stated that these criteria actually decreased model accuracy.

3.3.5. Performance measurement

3.3.5.1. Absolute valuation models

After calculating the intrinsic values for the RIVM and the DDM, errors were calculated in order to assess models' performances using 0% and 4% terminal growth rates to assess the impact on value estimates both for the DDM and RIVM.

3.3.5.2. Relative valuation models

Thereafter, the intrinsic value is calculated using combinations of the harmonic mean and mean with two and three digits from the SIC. Afterwards, these are used to calculate a deviation (henceforth error) from the market price four months ahead of fiscal year end, presented in the results section for both price to forward earnings and price to book value.

3.3.5.3 Section summary

In summary, after the performance measure calculation in the last two subsections, decile portfolios are going to be created following a binary code, where number one represents the top two deciles and zero represents the bottom two deciles. Afterwards, a selection of two firms, with opposite pricing performance for both RIVM and P/FE, will follow through onto the small sample.

3.1. Small sample

3.1.1. Data recovery and analysis methods

According to the previous sections and subsections, analysts' reports will be retrieved from Thomson Reuters Eikon for each of the selected firms. This collection of information is intended to yield conclusions on model valuations and provide insights on market pricing.

Additionally, using the initial dataset, although lacking the adequate amount of data, univariate and multivariate regression models based upon prices and earnings data from 2005 to 2015 for AMAT and 2006 to 2015 for JEC will be performed to assess the relationship between market prices returns, earnings levels and changes to assess earnings value relevance for each firm considering stock returns calculated using prc4, and adjusted earnings data from the initial dataset. Ultimately,this last step is intended to understand to what extent do earnings determine market prices and what is the relation to the models' value estimates.

Chapter IV: Results

4.1. Large sample

4.1.1. Absolute valuation models

Variable	N	Mean	SD	Min	P1	P25	P50	P75	P99	Max
v_rim_g0_error_signed	12150	0.44	7.35	-1.00	-0.92	-0.25	0.18	0.65	3.34	699.74
v_rim_g0_error_abs	12150	0.75	7.32	0.00	0.01	0.21	0.46	0.77	3.34	699.74
v_rim_g4_error_signed	10681	8.07	109.95	-1.00	-0.92	0.20	1.36	3.52	86.05	8550.31
v_rim_g4_error_abs	10681	8.25	109.94	0.00	0.02	0.55	1.36	3.52	86.05	8550.31
v_ddm_g0_error_signed	3625	-0.36	0.67	-1.00	-0.99	-0.80	-0.56	-0.11	1.92	12.89
v_ddm_g0_error_abs	3625	0.62	0.44	0.00	0.02	0.36	0.63	0.84	1.92	12.89
v_ddm_g4_error_signed	2280	12.61	339.32	-1.00	-0.99	-0.52	0.29	2.16	85.65	16075.29
v_ddm_g4_error_abs	2280	13.10	339.31	0.00	0.02	0.44	0.82	2.16	85.65	16075.29

Table 3: Absolute valuation performance measures before trimming

After calculating the intrisic values for the RIVM and the DDM, errors were calculated in order to assess models' performances. Following, the first round of results the absolute methods yield the following performance measures descriptive statistics. In a first glance it worth noting that there are some outliers which should be accounted for before proceeding with the analysis. From the errors above were created percentile portfolios which the purpose of trimming 1% top and bottom for signed valuation errors and 1% top for absolute valuation errors. The results after this trimming are presented in the table below.

Variable	N	Mean	SD	Min	P1	P25	P50	P75	P99	Max
v_rim_g0_error_signed	11359	0.23	0.65	-0.92	-0.86	-0.24	0.17	0.60	2.05	2.58
v_rim_g0_error_abs	11359	0.53	0.44	0.00	0.01	0.20	0.43	0.72	2.05	2.58
v_rim_g4_error_signed	10069	2.70	4.62	-0.89	-0.82	0.22	1.31	3.18	24.92	37.89
v_rim_g4_error_abs	10069	2.86	4.52	0.00	0.02	0.52	1.31	3.18	24.92	37.89
v_ddm_g0_error_signed	3303	-0.43	0.48	-0.99	-0.98	-0.80	-0.58	-0.17	1.01	1.35
v_ddm_g0_error_abs	3303	0.58	0.29	0.00	0.02	0.35	0.61	0.82	1.01	1.35
v_ddm_g4_error_signed	2047	1.20	2.77	-0.97	-0.96	-0.51	0.20	1.67	12.62	16.00
v_ddm_g4_error_abs	2047	1.69	2.50	0.00	0.01	0.40	0.76	1.67	12.62	16.00

Table 4: Absolute valuation performance measures after trimming

In a first glance, it is observable that the RIVM yields a lower mean, hence a better performance than the DDM for a terminal growth rate of zero percent. Moreover, the value resulting from the RIVM is by nature more sensitive to terminal growth rates, as opposed to the DDM. This can be observed in the table above by assessing the changes in the mean value of signed errors from assuming 0% or 4% terminal growth rates, concluding that the DDM is less sensitive to terminal growth rate changes than the RIVM.

Additionally, paired t-tests were performed on combinations of growth rate assuming that the dependent variable (v_ddm_erro) is continuous, the independent variable (v_rim_error) uses the same related groups (firms) as the dependent variable, that there are no significant outliers (after the trimming) and that the differences in the dependent variable are approximately normally distributed. The results presented in table 5 for a 95% confidence level are that all the averaging method/industry combinations are statistically significant and different from zero. The null hypothesis is that the mean is zero and the alternative hypothesis number is that the mean is different from zero.

Diff(Mean)	v_ddm_g0_error_signed	v_ddm_g0_error_abs	v_ddm_g4_error_signed	v_ddm_g4_error_abs
v_rim_g0_error_signed	0.84	-	-	-
v_rim_g0_error_abs	-	0.04	-	-
v_rim_g4_error_signed	-	-	2.53	-
v_rim_g4_error_abs	-	-	-	2.15

Table 5: Mean difference between performance measures in absolute valuation

In average terms, the means are statistically significant and different from zero.

Whilst, it is worth noting that the number observations in the DDM is significantly smaller than in the RIVM due to a large number of missing values of forecasted dividends per share and negative values resulting from the model's sensitivity to the terminal growth rates undoubtedly leading the RIVM as the choice to create decile portfolios.

4.1.2. Relative valuation models

Industry	Average	Variable	N	Mean	SD	Min	P1	P25	P50	P75	P99	Max
SIC 2	Harmonic Mean	fpe_error_signed	21385	1.42	85.05	-1089.85	-18.06	-0.73	0.05	0.78	40.95	9331.33
		fpe_error_abs	21385	4.22	84.96	0.00	0.01	0.31	0.75	1.87	56.01	9331.33
	Mean	fpe_error_signed	21383	0.22	8.07	-206.71	-20.20	-0.41	0.34	1.30	13.96	347.84
		fpe_error_abs	21383	2.36	7.72	0.00	0.01	0.37	0.89	2.03	26.16	347.84
SIC 3	Harmonic Mean	fpe_error_signed	21135	1.13	89.92	-2322.92	-14.24	-0.68	0.00	0.59	20.93	9331.33
		fpe_error_abs	21135	3.98	89.84	0.00	0.01	0.25	0.64	1.45	44.51	9331.33
	Mean	fpe_error_signed	21092	1.65	7.98	-1.00	-0.92	-0.04	0.55	1.61	19.67	491.95
		fpe_error_abs	21092	1.87	7.93	0.00	0.01	0.30	0.71	1.61	19.67	491.95

Price to Book Value

Industry	Average	Variable	N	Mean	SD	Min	P1	P25	P50	P75	P99	Max
SIC 2	Harmonic Mean	pbv_error_signed	24643	0.06	1.37	-1.00	-0.94	-0.53	-0.20	0.29	4.47	35.52
		pbv_error_abs	24643	0.63	1.22	0.00	0.01	0.23	0.46	0.71	4.47	35.52
	Mean	pbv_error_signed	24643	7.41	105.27	-1.00	-0.88	-0.03	0.69	1.89	30.85	9594.81
		pbv_error_abs	24643	7.62	105.26	0.00	0.01	0.34	0.79	1.89	30.85	9594.81
SIC 3	Harmonic Mean	pbv_error_signed	24385	0.14	3.21	-148.22	-0.95	-0.52	-0.17	0.34	5.45	361.47
		pbv_error_abs	24385	0.73	3.13	0.00	0.01	0.23	0.47	0.73	5.63	361.47
	Mean	pbv_error_signed	24372	5.97	124.17	-1.00	-0.89	-0.14	0.49	1.57	23.11	8667.37
		pbv_error_abs	24372	6.21	124.15	0.00	0.01	0.30	0.67	1.57	23.11	8667.37

Table 6: Relative valuation performance measures before trimming

After a first round of results the multiples methods yield the following performance measures descriptive statistics. In a first glance it worth noting that there are some outliers which should be accounted for before proceeding with the analysis. From the errors above were created percentile portfolios which the purpose of trimming 1% top and bottom for signed valuation errors and 1% top for absolute valuation errors. The results after this trimming are presented in the table below.

Price to Forward Earnings

Industry	Average	Variable	N	Mean	SD	Min	P1	P25	P50	P75	P99	Max
SIC 2	Harmonic Mean	fpe_error_signed	17244	0.10	2.98	-17.72	-7.20	-0.68	0.05	0.70	9.04	40.95
		fpe_error_abs	17244	1.47	2.59	0.00	0.01	0.28	0.68	1.68	12.18	40.95
	Mean	fpe_error_signed	17243	0.37	2.18	-11.58	-7.61	-0.34	0.35	1.21	6.80	11.57
		fpe_error_abs	17243	1.41	1.71	0.00	0.01	0.35	0.82	1.77	8.79	11.58
SIC 3	Harmonic Mean	fpe_error_signed	17013	-0.01	1.66	-9.42	-5.07	-0.61	0.00	0.50	5.90	9.64
		fpe_error_abs	17013	1.00	1.32	0.00	0.01	0.22	0.55	1.22	6.81	9.64
	Mean	fpe_error_signed	16981	0.90	1.38	-0.92	-0.85	-0.03	0.52	1.45	5.84	7.05
		fpe_error_abs	16981	1.10	1.23	0.00	0.01	0.28	0.66	1.45	5.84	7.05

Price to Book Value

Industry	Average	Variable	N	Mean	SD	Min	P1	P25	P50	P75	P99	Max
SIC 2	Harmonic Mean	pbv_error_signed	19727	-0.10	0.54	-0.93	-0.86	-0.52	-0.22	0.20	1.56	2.12
		pbv_error_abs	19727	0.46	0.32	0.00	0.01	0.21	0.42	0.64	1.56	2.12
	Mean	pbv_error_signed	19727	1.08	1.95	-0.79	-0.69	-0.02	0.61	1.58	8.52	23.96
		pbv_error_abs	19727	1.25	1.85	0.00	0.01	0.30	0.66	1.58	8.52	23.96
SIC 3	Harmonic Mean	pbv_error_signed	19490	-0.08	0.56	-0.88	-0.85	-0.50	-0.20	0.22	1.62	1.91
		pbv_error_abs	19490	0.46	0.33	0.00	0.01	0.21	0.42	0.64	1.62	1.91
	Mean	pbv_error_signed	19490	0.77	1.26	-0.75	-0.70	-0.12	0.42	1.28	5.28	6.90
		pbv_error_abs	19490	0.98	1.11	0.00	0.01	0.26	0.56	1.28	5.28	6.90

Table 7: Relative valuation performance measures after trimming

Beforehand it is worth noting that there were improvements in the mean and standard deviation. In detail, the features analyzed have an impact on the intrinsic value calculation and consequently affect both relative and absolute valuation errors. Accordingly, the information displayed in table 7 above illustrates the differences between using the harmonic mean or median to calculate the industry multiple as well as the differences between the choice of two and three digits SICs as comparable firms. As previously stated and observed in the results above, there are meaningful improvements in model performance when using the harmonic mean instead of the mean as well as in a narrower choice of comparable firms. Nonetheless, there are some points worth mentioning, such as the higher standard deviation as a result from using the harmonic mean when compared to the mean verified in both two and three digits SICs.

Additionally, paired t-tests were performed on combinations of averaging method and industry assuming that the dependent variable (fpe) is continuous, the independent variable (pbv) uses the same related groups (firms) as the dependent variable, that there are no significant outliers (after the trimming) and that the differences in the dependent variable are approximately normally distributed. The null hypothesis is that the mean is zero and the alternative hypothesis number is that the mean is different from zero.

Diff(Mean)	pbv_hm_sic2_error_signed	pbv_hm_sic2_abs	pbv_m_sic2_error_signed	pbv_m_sic2_error_abs	pbv_hm_sic3_error_signed	pbv_hm_sic3__error_abs	pbv_m_sic3_error_signed	pbv_m_sic3_error_abs
fpe_hm_sic2_error_signed	0.14	-	-	-	-	-	-	-
fpe_hm_sic2_error_abs	-	0.99	-	-	-	-	-	-
fpe_m_sic2_error_signed	-	-	0.04	-	-	-	-	-
fpe_m_sic2_error_abs	-	-	-	0.21	-	-	-	-
fpe_hm_sic3_error_signed	-	-	-	-	0.13	-	-	-
fpe_hm_sic3_error_abs	-	-	-	-	-	0.58	-	-
fpe_m_sic3_error_signed	-	-	-	-	-	-	0.22	-
fpe_m_sic3_error_abs	-	-	-	-	-	-	-	0.21

Table 8: Mean difference between performance measures in relative valuation

The results presented in table 8 for a 95% confidence level are that all the averaging method/industry combinations are statistically significant and different from zero.

In average terms, the means are statistically significant and different from zero.

Whilst, it is worth noting that the number observations is significantly different between both multiples resulting from missing values in the earnings forecast. Although the results presented above point price to forward earnings using harmonic mean and three digits SIC as the best performing multiple, undoubtedly leading this multiple as the choice to create decile portfolios along with the RIVM.

4.2. Small sample

Model	(1) AMAT	(2) JEC	(3) AMAT	(4) JEC	(5) AMAT	(6) JEC	(7) AMAT	(8) JEC
Constant	0.27	0.54	-	-	0.28	0.09	-0.07	0.47
Std. Error	0.16	0.62	-	-	0.18	0.86	0.19	1.26
t-ratio	1.67	0.88	-	-	1.56	0.10	-0.38	0.37
p-value	0.14	0.41	-	-	0.17	0.92	0.72	0.74
epspx_adj	-0.31	-0.21	-	-	-0.32	-0.04	-0.43	0.21
Std. Error	0.19	0.22	-	-	0.21	0.31	0.14	0.44
t-ratio	-1.67	-0.95	-	-	-1.55	-0.13	-3.05	0.48
p-value	0.14	0.38	-	-	0.17	0.90	0.04	0.66
delta epsx_adj	-	-	0.06	-0.02	0.01	-0.45	0.01	-0.79
Std. Error	-	-	0.12	0.10	0.06	0.57	0.03	0.86
t-ratio	-	-	0.48	-0.20	0.22	-0.79	0.22	-0.91
p-value	-	-	0.65	0.85	0.83	0.47	0.84	0.43
eps_1_adj	-	-	-	-	-	-	0.50	-0.35
Std. Error	-	-	-	-	-	-	0.16	0.37
t-ratio	-	-	-	-	-	-	3.19	-0.93
p-value	-	-	-	-	-	-	0.03	0.42
delta eps_1_adj	-	-	-	-	-	-	0.01	0.94
Std. Error	-	-	-	-	-	-	0.05	1.11
t-ratio	-	-	-	-	-	-	0.25	0.85
p-value	-	-	-	-	-	-	0.82	0.46
R Squared	0.29	0.13	0.01	0.22	0.29	0.23	0.83	0.42
Adjusted R Squared	0.18	-0.01	-0.13	0.09	0.06	-0.08	0.67	-0.36
Overall F-Test	2.80	0.90	0.06	1.73	1.23	0.73	5.01	0.54

Table 9: Regression results

According to the data described in table 9 above, it can be seen that although only two variables are statistically significant at a 95% confidence level, the models show improvements in their explanatory power as the number of variables increase. Firstly, to show the differences between earnings levels (1 and 2) and earnings changes (3 and 4) it is observable that earnings levels explain 29% of stock returns for (1) AMAT comparing with 13% for (2) JEC. The opposite can be seen for earnings changes, where the model yields 1% for (3) AMAT and 22% for (4) JEC. The combination of the previous models yields similar explanatory power, 29% for (5) AMAT and 23% for (6) JEC. When introducing the same feature, but considering one year forecasted earnings, the difference takes off yielding 83% for (7) AMAT and 42% for (8) JEC. Henceforth, models 7 and 8 will be used as the basis for comparison with qualitative information retrieved from analysts' reports.

According to several analysts' reports retrieved from Thomson Reuters Eikon, this differences in explanatory power can be due to firm specific risks. To be approached in more detail during the following discussion section.

4.3. Chapter summary

After creating the decile portfolios, a ranking of firm-model performance is created. This ranking serves as guidance to the choice of firms to study in more detail in the former section. As such, AMAT – Applied Materials Inc. - appears as a firm in the top two deciles, contrary to JEC – Jacobs Engeneering Group Inc. - which appears in the bottom two deciles. The purpose of chosing these firms is to understand what factors may lead to the difference in value estimates and how that is connected to earnings value relevance.

Chapter V: Discussion and conclusions

5.1. Results discussion

Beforehand, it is worth noting that this study intends to measure the extent to which earnings explain market prices. Hence, the regression performed in the previous section yields some interesting inferences namely that AMAT's market price is largely explained by current earnings levels and changes, as well as future earnings levels and changes while the contrary is observable for JEC.

Decile	Firm	PRC4 - Price 30th April 2015	FPE	Relative Error	Qualitative Valuation	RIVM	Relative Error	Qualitative Valuation
Top	AMAT	25.05	30.80	0.23	Overvalued	30.85	0.23	Overvalued
Bottom	JEC	38.10	62.82	0.65	Overvalued	61.28	0.61	Overvalued

Table 10: AMAT and JEC comparison between price to forward earnings, residual income valuation model value estimates and observed market prices

In line with the ranking from the large sample results section, where in April 2015 – year chosen for the comparison of the firms, AMAT had a market price of 25.05 and the models value estimate was within 25% of the market price. This value is slightly higher than the value admissible for an hold recommendation, although given the regression results it is plausible that the models' value estimates represent true value. Correspondingly, Sidney Ho and Ross Seymore – Deutsche Bank's equity analysts, review in their 15 of May equity report, strong trends regarding revenues and capex spending, only reiterating as risks the cyclicality of the business, global demand for consumer electronics, highly concentrated consumer based with 25% of the revenue coming from two major clients and the reduction in government subsidies in their energy businesses.

For JEC, we observe the opposite, the models' value estimates were outside the limits for an hold recommendation, giving a pricing error of 65%, which can be justifiable by JEC's stock returns being less explainable by earnings, fundamental on which both residual income valuation model and forward price to earnings are based on. To be noticed that at the time of the report JEC was going through a restructuring process with no certain outcomes at the date. Consequently, the reasons for this difference in pricing errors can be due to changes in competitive environment, macro conditions which may affect capex spending and cost overruns impacts on margins, as detailed in Vishal Shah and Chad Dillard- Deustche Bank's equity analysts; 29 of April 2015 equity report.

5.2. Conclusions

Altogether, we can state for conclusion that firms with weaker earnings value relevance, tend have larger valuation errors in models which use earnings as the value driver.

5.2.1. Answer to the research question

Earnings do determine market prices, recurring to econometric analysis and regression techniques, R^2 provides a meaningful measure of earnings value relevance.

$$Stock\ Returns_t = \alpha_t + \beta_1 epspx\_adj_t + e_t$$

$$Stock\ Returns_t = \alpha_t + \beta_1 delta\_epspx\_adj_t + e_t$$

$$Stock\ Returns_t = \alpha_t + \beta_1 epspx\_adj_t + \beta_2 delta\_epspx\_adj_t + e_t$$

$$Stock\ Returns_t = \alpha_t + \beta_1 epspx\_adj_t + \beta_2 delta\_epspx\_adj_t + \beta_3 eps\_1_t + \beta_4 delta\_eps\_1_t + e_t$$

5.3. Limitations

In this section there are presented some caveats to this study. First, the modern literature on valuation theory is currently almost sixty years old. Accordingly, the models chosen to perform the valuation were based on the information available in the databases, and although the theory fundamentals have remained mostly the same over time, there are some firms and industries included in the large sample section which these models do not explain as well causing outliers in the performance measure.

Second, the number of observations in the large sample analysis is not uniform along forecasted items (earnings and dividends) which causes the number of observations to drop drastically on the dividend discount model. Third, the industry proxy used (2 and 3 digits SIC), is deemed insufficient and there are measures available that complement the former proxy. Fourth, when calculating the cost of equity, the equity risk premium is being considered yearly for the U.S. and not for industry or firm biasing the discount rate. Fifth, the regressions performed in the small sample section, lack a minimum number of observations causing a dimensional bias, along with multicollinearity issues and noise in the error term, which given the sample size, were chosen not to be corrected for.

Sixth in the small sample analysis, analysts' opinions are subjective, and are based upon their cognitive matrix and should not be taken as absolute.

5.4. Strenghts of the study

Notwithstanding the former limitations, this study has predictive power to the extent that it has not contradicted the theory. As per, has allowed the researcher to present a measure of earnings value relevance, and to show its connection to pricing errors. Moreover, it allows the reader to assess the differences in value estimates recurring to different terminal growth rates, industry choices and averaging methods.

5.5. Further research recomendations

In a further study would be interesting to assess the performance of other absolute and relative valuation models, as well as to deepen the analysis on terminal value assumptions and industry proxy effects on the absolute and relative valuation models respectively. In detail, adapt the accounting flows models for different growth stages and assess firm size, riskiness and future cash flow generating opportunities as a proxy for comparable firms to be used in relative valuation models.

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*, 94-108.

Baker, M., & Wurgler, J. (2004). Appearing and disappearing dividends: The link to catering incentives. *Journal of Financial Economics, 73*(2), 271-288.

Baker, M., & Wurgler, J. (2004). A catering theory of dividends. *The Journal of Finance, 59*(3), 1125-1165.

Barth, M. E., Beaver, W. H., & Landsman, W. R. (2001). The relevance of the value relevance literature for financial accounting standard setting: another view. *Journal of accounting and economics, 31*(1-3), 77-104.

Beatty, R. P., Riffe, S. M., & Thompson, R. (1999). The method of comparables and tax court valuations of private firms: An empirical investigation. *Accounting Horizons, 13*(3), 177-199.

Bhojraj, S., & Lee, C. M. (2002). Who is my peer? A valuation-based approach to the selection of comparable firms. *Journal of accounting research, 40*(2), 407-439.

Bhojraj, S., Lee, C. M., & Oler, D. K. (2003). What's my line? A comparison of industry classification schemes for capital market research. *Journal of Accounting Research, 41*(5), 745-774.

Boulding, K. E. (1935). The theory of a single investment. *The Quarterly Journal of Economics, 49*(3), 475-494.

Cheng, C. A., & McNamara, R. (2000). The valuation accuracy of the price-earnings and price-book benchmark valuation methods. *Review of Quantitative Finance and Accounting, 15*(4), 349-370.

Damodaran, A. (2007). Valuation approaches and metrics: a survey of the theory and evidence. *Foundations and Trends® in Finance, 1*(8).

Daniels, M. B. (1934). Principles of asset valuation. *Accounting Review*, 114-121.

DeAngelo, H., DeAngelo, L., & Skinner, D. J. (2004). Are dividends disappearing? Dividend concentration and the consolidation of earnings. *Journal of financial economics, 72*(3), 425-456.

Dechow, P. M., Hutton, A. P., & Sloan, R. G. (1999). An empirical assessment of the residual income valuation model. *Journal of accounting and economics, 26*(1-3), 1-34.

Durand, D. (1957). Growth stocks and the Petersburg paradox. *The Journal of Finance, 12*(3), 348-363.

Fama, E. F. (1991). Efficient capital markets: II. *The journal of finance, 46*(5), 1575-1617.

Feltham, G. A., & Ohlson, J. A. (1995). Valuation and clean surplus accounting for operating and financial activities. *Contemporary accounting research, 11*(2), 689-731.

Fisher, I., & Barber, W. J. (1907). *The rate of interest.* Garland Pub.

Fisher, I. (1930). The theory of interest. *New York*, 43.

Francis, J., Olsson, P., & Oswald, D. R. (2000). Comparing the accuracy and explainability of dividend, free cash flow, and abnormal earnings equity value estimates. *Journal of accounting research, 38*(1), 45-70.

Frankel, R., & Lee, C. M. (1998). Accounting valuation, market expectation, and cross-sectional stock returns. *Journal of Accounting and economics, 25*(3), 283-319.

Fuller, R. J., & Hsia, C. C. (1984). A simplified common stock valuation model. *Financial Analysts Journal*, 49-56.

Glesne, C. P., & Peshkin, S. A. (1992). *Becoming quantitative researchers: An introduction.* White Plains, NY: Longman.

Hancock, B., Ockleford, E., & Windridge, K. (1998). *An introduction to qualitative research.* Notthingham: Trent focus group.

Healy, M., & Perry, C. (2000) Comprehensive criteria to judge validity and reliability of quantitative research within the realis paradigm. Qualitative markey research: *An international journal.* **3** (3), 118-126.

Hillier, D., Ross, S., Westerfield, R., Jaffe, J. & Jordan, B. (2016). Corporate Finance (3rd ed), *Risk and Return: The Capital Asset Pricing Model* (pp. 253-293). McGraw-Hill Education.

Keynes, J. M. (1937). The general theory of employment. *The quarterly journal of economics, 51*(2), 209-223.

Lee, C. M., Myers, J., & Swaminathan, B. (1999). What is the Intrinsic Value of the Dow? *The Journal of Finance, 54*(5), 1693-1741.

Lie, E., & Lie, H. J. (2002). Multiples used to estimate corporate value. *Financial Analysts Journal*, 44-54.

Liu, J., Nissim, D., & Thomas, J. (2002). Equity valuation using multiples. *Journal of Accounting Research, 40*(1), 135-172.

Mazzariol, P., & Thomas, M. (2016). Theory and practice in M&A valuations. *Strategic Direction*, **32** (6), 8–11. https://doi.org/10.1108/SD-03-2016-0032

Ohlson, J. A. (1995). Earnings, book values, and dividends in equity valuation. *Contemporary accounting research, 11*(2), 661-687.

Penman, S. H., & Sougiannis, T. (1997). The dividend displacement property and the substitution of anticipated earnings for dividends in equity valuation. *Accounting Review*, 1-21.

Richardson, G., & Tinaikar, S. (2004). Accounting based valuation models: What have we learned? *Accounting and Finance,* **44** (2), 223-255. https://doi.org/10.1111/j.1467-629X. 2004.00109.

Robson, C. (2002). *Real World Research: A Resource for Social Scientists and Practioner-Researchers.*, 2nd ed.

Samuelson, P. A. (1937). Some aspects of the pure theory of capital. *The Quarterly Journal of Economics, 51*(3), 469-496.

https://en.oxforddictionaries.com/definition/determine, consulted in 28/08/2018 at 17h34m.

https://en.oxforddictionaries.com/definition/extension, consulted in 28/08/2018 at 16h46m.

https://www.nasdaq.com/investing/glossary/e/earnings-per-share, consulted in 31/08/2018 at 16h47m

https://www.nasdaq.com/investing/glossary/m/market-prices, consulted in 31/08/2018 at 16h47m

Appendices

Appendix I: Initial dataset before preparation

Dataset Variable Description					
Variable	Obs	Mean	Std. Dev.	Min	Max
permno	37106	70431.81	26419.75	10001	93436
date	37106	18703.48	1165.247	16587	20573
exchcd	37106	2.047378	.9845965	1	4
siccd	37106	4967.093	2090.678	0	9999
cusip	0				
ticker	0				
comnam	0				
tsymbol	0				
naics	0				
trdstat	0				
beta	37106	1.281538	.7864947	-18.22339	9.013506
n	37106	57.06689	8.036589	24	60
prc4	37106	30.22909	43.36951	.0376	1661.29
ib_tic	0				
eps_1	37106	-90.27023	8544.74	-975000	338625
eps_2	35853	-37.77812	6421.559	-702000	400500
eps_3	22542	-47.30604	6055.824	-791666.7	51600
eps_4	11055	-13.22926	4678.644	-427500	243333.3
eps_5	7799	88.67959	6650.282	-882	582500

eps_g	25496	13.84251	13.59837	-246.5	985
dps_1	23680	.7671191	4.663744	0	600
dps_2	22969	.870785	9.310148	0	1200
dps_3	16312	.8736629	2.604806	0	160
dps_4	4924	1.46107	3.759334	0	117
dps_5	3394	2.111709	30.45265	0	1761
dps_g	2510	9.45543	12.62465	-6.54	236
gvkey	0				
datadate	37106	18583.18	1165.164	16467	20453
tic	0				
conm	0				
exchg	37106	12.8653	2.038788	1	19
cik	0				
costat	0				
fyear	37106	2009.903	3.194245	2004	2015
ajex	37106	1.105819	7.808574	2.00e-06	1500
act	29596	2091.016	7232.026	0	152629
aqc	35735	112.6287	833.488	-31804	43123
at	37106	16830.51	120077.7	.118	3065553

bkvlps	37039	20.70226	1010.09	-3311	173499
capx	36970	371.5132	1844.427	-401.609	50233.74
ceq	37083	3071.032	12159.58	-96620	255550
che	37104	1896.793	17608.2	0	728111
csho	37092	203.9687	719.4421	.001	29058.36
cshpri	37092	201.5537	698.8139	.966	28776
dlc	37077	1510.164	17956.41	0	562857
dp	35453	284.9768	1253.035	0	33750.97
dpc	36111	293.9644	1303.376	-21.363	44667
dvc	36955	157.3098	818.4356	-10	67643.8
epspx	37092	1.229385	27.20103	-196.12	5193
ib	37103	355.7106	1923.853	-99289	53394
invt	36578	889.2668	11761.35	0	472266.2
ivch	35183	1881.508	25227.9	-.022	2405971
lct	29640	1559.745	6369.401	0	329795
ni	37103	364.4114	2008.05	-99289	98806.04
sale	37103	5211.474	20605.69	-1964.999	483521
siv	35354	1513.129	21165.86	-362.84	2093057
spi	36718	-57.45707	650.1462	-41250	17675.38
txt	37101	164.1964	995.0519	-34831	36530

xad	15542	120.556	478.2531	0	9729
xint	31563	143.955	1087.255	-.875	57302
xrd	19668	196.2499	821.4385	0	12540
prcc_c	37089	31.91139	516.6544	.07	99200
mkvalt	32749	5325.01	20840.3	.0181	626550.4
sic	0				
rank	37106	1	0	1	1
au	0				
auop	0				
auopic	0				
ap	36772	4727.094	53906.17	0	1692680
fincf	37067	-29.63135	4638.997	-199568	250506
gdwl	36086	922.7309	4219.63	0	129709.6
intan	36476	1423.61	6672.919	0	225278
ivncf	37067	-636.1086	4841.428	-286346	157925
oancf	37067	730.4882	4045.795	-110560	129731
ppegt	31966	4407.8	20046.92	0	501069.8
rect	36866	4944.264	49189.72	0	1291825

Appendix II: Stata large sample code

/// Data preparation and observation standardization

/// 0. Open Valuation Data File

```
use "/Users/diogomarques/Dropbox/Diogo/9.FormacÃßaÃÉo/Dissertation - Double Degree/Dissertation/5. Dissertation Data/V.1/valuation_data.dta"
```

/// 1. Initial dataset tabulation

/// Note:

```
sum permno date exchcd siccd cusip ticker comnam tsymbol naics trdstat beta n prc4 ib_tic eps_1 eps_2 eps_3 eps_4 eps_5 eps_g dps_1 dps_2 dps_3 dps_4 dps_5 dps_g gvkey datadate tic conm exchg cik costat fyear ajex act aqc at bkvlps capx ceq che csho cshpri dlc dp dpc dvc epspx ib invt ivch lct ni sale siv spi txt xad xint xrd prcc_c mkvalt sic rank au auop auopic ap fincf gdwl intan ivncf oancf ppegt rect
```

/// 2. Eliminate Financial Firms

/// Note:

```
gen sic_1=substr(sic,1,1)
drop if sic_1=="6"
drop if fyear==2004
```

/// 3. Generate 2 digit SIC code and eliminate industries with less than 10 observations

/// Note:

```
gen sic_2=substr(sic,1,2)
gen sic_3=substr(sic,1,3)
egen num_0=count(eps_1), by(sic_2 fyear)
drop if num_0<5
egen num_1=count(eps_2), by(sic_2 fyear)
drop if num_1<5
```

/// 4. Compute the dividend payout (use "ni" and "dvc"), adjusted (for stock splits) book value of equity per share (use "ceq", "csho", "ajex") and adjusted earnings per share (for stock splits) use ("epspx", "ajex")

/// Note:

```
gen divpayout=dvc/ni
gen bvpsadj=ceq/(csho*ajex)
gen epsadj=epspx/ajex
gen eps_1_adj=eps_1/ajex
gen eps_2_adj=eps_2/ajex
gen eps_3_adj=eps_3/ajex
gen dps_0_adj=dvc/(csho*ajex)
gen dps_1_adj=dps_1/ajex
gen dps_2_adj=dps_2/ajex
gen dps_3_adj=dps_3/ajex
```

/// 5. Descriptive Statistics Table

/// Note:

```
tabstat prc4 bvpsadj epsadj eps_1_adj eps_2_adj dps_0_adj dps_1_adj dps_2_adj, stat(n me sd min p1 p25 p50 p75 p99 max) col(stat) case
```

/// 6. Generate percentiles for prc4 (price four months ahead the fiscal year end) and drop 1,99 and 100 percentiles.

/// Note:

```
xtile xprc4=prc4, nq(100)
drop if xprc4==1 | xprc4==100
```

/// --

```
xtile xbvpsadj=bvpsadj, nq(100)
drop if xbvpsadj==1 | xbvpsadj==100
```

/// --

```
xtile xepsadj=epsadj, nq(100)
drop if xepsadj==1 | xepsadj==100
xtile xeps_1_adj=eps_1_adj, nq(100)
drop if xeps_1_adj==1 | xeps_1_adj==100
xtile xeps_2_adj=eps_2_adj, nq(100)
drop if xeps_2_adj==1 | xeps_2_adj==100
```

/// ---

```
xtile xdps_0_adj=dps_0_adj, nq(100)
drop if xdps_0_adj==100
xtile xdps_1_adj=dps_1_adj, nq(100)
drop if xdps_1_adj==100
xtile xdps_2_adj=dps_2_adj, nq(100)
drop if xdps_2_adj==100
```

/// 7. Descriptive Statistics Table

/// Note:

```
tabstat prc4 bvpsadj epsadj eps_1_adj eps_2_adj dps_0_adj dps_1_adj dps_2_adj, stat(n me sd min p1 p25 p50 p75 p99 max) col(stat) case
sum prc4 bvpsadj epsadj eps_1_adj eps_2_adj dps_0_adj dps_1_adj dps_2_adj
```

/// 8. Generate Multiples: Trailing and Forward Price/Earnings [(prc4/(epsadj=epspx/ajex) and prc4/(eps_1_adj=esp_1/ajex)]

/// Note:

```
gen pe=prc4/epsadj if epsadj>0 & epsadj!=.
gen fpe=prc4/eps_1_adj if eps_1_adj>0 & eps_1_adj!=.
```

/// 9. Generate Multiples: Price to Book Value (prc4/bvpsadj)

/// Note:

```
gen pb=prc4/bvpsadj if bvpsadj>0 & bvpsadj!=.
```

/// 10. Identify Panel Structure

/// Note:

```
egen i= group(gvkey)
duplicates drop i fyear, force
```

```
xtset i fyear
```

/// 11. Generate Forward P/E - Harmonic Mean and Mean, SIC_2 and SIC_3, and Calculate Intrinsic Value

/// Note:

/// 11.1 Generate Harmonic Mean of Forward PE and Calculate Instrinsic Values based on obtained SIC_2 industry FPE multiples

/// Note:

```
gen fep_sic2=eps_1_adj/prc4
gen fep_duplicate_sic2=fep_sic2

gen fpe_industry_hm_sic2=.
egen i_fyear_8=group(gvkey fyear)
sum i_fyear_8
local min=r(min)
local max=r(max)
forvalues k=`min'(1)`max'{
        qui replace fep_sic2=. if i_fyear_8 ==`k'
        qui egen help=mean(fep_sic2), by(sic_2 fyear)
        qui replace fpe_industry_hm_sic2=help if i_fyear_8==`k'
        qui drop help
        qui replace fep_sic2=fep_duplicate_sic2 if i_fyear_8 ==`k'
        di `k'
        }
replace fpe_industry_hm_sic2=1/fpe_industry_hm_sic2
gen v_fpe_harm_sic2=eps_1_adj*fpe_industry_hm_sic2 if eps_1_adj>0
```

/// 11.2 Generate Mean of Forward PE and Calculate Instrinsic Values based on obtained SIC_2 industry PE multiples

/// Note:

```
gen fpe_duplicate_sic2=fpe
gen fpe_industry_m_sic2=.
egen i_fyear_9=group(gvkey fyear)
sum i_fyear_9
local min=r(min)
local max=r(max)
forvalues k=`min'(1)`max'{
        qui replace fpe=. if i_fyear_9==`k'
        qui egen help=mean(fpe), by(sic_2 fyear)
        qui replace fpe_industry_m_sic2=help if i_fyear_9==`k'
        qui drop help
        qui replace fpe=fpe_duplicate_sic2 if i_fyear_9==`k'
        di `k'
        }
gen v_fpe_mean_sic2=epsadj*fpe_industry_m_sic2 if eps_1_adj>0
```

/// 11.3 Generate Harmonic Mean of Forward PE and Calculate Instrinsic Values based on obtained SIC_3 industry PE multiples

/// Note:

```
gen fep_sic3=eps_1_adj/prc4
gen fep_duplicate_sic3=fep_sic3

gen fpe_industry_hm_sic3=.
egen i_fyear_10=group(gvkey fyear)
sum i_fyear_10
local min=r(min)
local max=r(max)
forvalues k=`min'(1)`max'{
        qui replace fep_sic3=. if i_fyear_10==`k'
        qui egen help=mean(fep_sic3), by(sic_3 fyear)
        qui replace fpe_industry_hm_sic3=help if i_fyear_10==`k'
        qui drop help
        qui replace fep_sic3=fep_duplicate_sic3 if i_fyear_10==`k'
        di `k'
        }
replace fpe_industry_hm_sic3=1/fpe_industry_hm_sic3
gen v_fpe_harm_sic3=eps_1_adj*fpe_industry_hm_sic3 if eps_1_adj>0
```

/// 11.4 Generate Mean of Forward PE and Calculate Instrinsic Values based on obtained SIC_2 industry PE multiples

/// Note:

```
gen fpe_duplicate_sic3=fpe
gen fpe_industry_m_sic3=.
egen i_fyear_11=group(gvkey fyear)
sum i_fyear_11
local min=r(min)
local max=r(max)
forvalues k=`min'(1)`max'{
        qui replace fpe=. if i_fyear_11==`k'
        qui egen help=mean(fpe), by(sic_3 fyear)
        qui replace fpe_industry_m_sic3=help if i_fyear_11==`k'
        qui drop help
        qui replace fpe=fpe_duplicate_sic3 if i_fyear_11==`k'
        di `k'
gen v_fpe_mean_sic3=eps_1_adj*fpe_industry_m_sic3 if eps_1_adj>0
```

/// 12. Generate P/BV - Harmonic Mean and Mean, SIC_1 and SIC_2, and Calculate Intrinsic Value

/// Note:

/// 12.1 Generate Harmonic Mean of PBV and Calculate Intrinsic Values based on obtained SIC2 industry PBV multiples

/// Note:

```
gen bvp_sic2=bvpsadj/prc4
gen bvp_duplicate_sic2=bvp_sic2
gen pbv_industry_hm_sic2=.
egen i_fyear_4=group(gvkey fyear)
sum i_fyear_4
local min=r(min)
local max=r(max)
forvalues k=`min'(1)`max'{
        qui replace bvp_sic2=. if i_fyear_4 ==`k'
        qui egen help=mean(bvp_sic2), by(sic_2 fyear)
        qui replace pbv_industry_hm_sic2=help if i_fyear_4 ==`k'
        qui drop help
        qui replace bvp_sic2=bvp_duplicate_sic2 if i_fyear_4 ==`k'
        di `k'
        }
replace pbv_industry_hm_sic2=1/pbv_industry_hm_sic2
gen v_pbv_harm_sic2=bvpsadj*pbv_industry_hm_sic2 if bvpsadj>0
```

/// 12.2 Generate Mean of PBV and Calculate Intrinsic Values based on obtained SIC2 industry PBV multiples

/// Note:

```
gen pb_duplicate_sic2=pb
gen pbv_industry_m_sic2=.
egen i_fyear_5=group(gvkey fyear)
sum i_fyear_5
local min=r(min)
local max=r(max)
forvalues k=`min'(1)`max'{
        qui replace pb=. if i_fyear_5 ==`k'
        qui egen help=mean(pb), by(sic_2 fyear)
        qui replace pbv_industry_m_sic2=help if i_fyear_5 ==`k'
        qui drop help
        qui replace pb=pb_duplicate_sic2 if i_fyear_5 ==`k'
        di `k'
        }
gen v_pbv_mean_sic2=bvpsadj*pbv_industry_m_sic2 if bvpsadj>0
```

/// 12.3 Generate Harmonic Mean of PBV and Calculate Intrinsic Values based on obtained SIC3 industry PBV multiples

/// Note:

```
gen bvp_sic3=bvpsadj/prc4
gen bvp_duplicate_sic3=bvp_sic3
gen pbv_industry_hm_sic3=.
egen i_fyear_6=group(gvkey fyear)
sum i_fyear_6
```

```
        local min=r(min)
        local max=r(max)
        forvalues k=`min'(1)`max'{
                qui replace bvp_sic3=. if i_fyear_6 ==`k'
                qui egen help=mean(bvp_sic3), by(sic_3 fyear)
                qui replace pbv_industry_hm_sic3=help if i_fyear_6 ==`k'
                qui drop help
                qui replace bvp_sic3=bvp_duplicate_sic3 if i_fyear_6 ==`k'
                di `k'
                }
        replace pbv_industry_hm_sic3=1/pbv_industry_hm_sic3
        gen v_pbv_harm_sic3=bvpsadj*pbv_industry_hm_sic3 if bvpsadj>0
/// 12.4 Generate Mean of PBV and Calculate Intrinsic Values based on obtained SIC3 industry PBV multiples
/// Note:
        gen pb_duplicate_sic3=pb
        gen pbv_industry_m_sic3=.
        egen i_fyear_7=group(gvkey fyear)
        sum i_fyear_7
        local min=r(min)
        local max=r(max)
        forvalues k=`min'(1)`max'{
                qui replace pb=. if i_fyear_7 ==`k'
                qui egen help=mean(pb), by(sic_3 fyear)
                qui replace pbv_industry_m_sic3=help if i_fyear_7 ==`k'
                qui drop help
                qui replace pb=pb_duplicate_sic3 if i_fyear_7 ==`k'
                di `k'
                }
        gen v_pbv_mean_sic3=bvpsadj*pbv_industry_m_sic3 if bvpsadj>0
/// 13. Calculate Absolute and Signed Valuation Errors for the 8 loops
/// Note:
        /// FPE - Harmonic Mean - SIC 2
         gen fpe_hm_sic2_error_signed=(v_fpe_harm_sic2-prc4)/prc4
         gen fpe_hm_sic2_error_abs=abs(v_fpe_harm_sic2-prc4)/prc4
         /// PE - Mean - SIC 2
         gen fpe_m_sic2_error_signed=(v_fpe_mean_sic2-prc4)/prc4
         gen fpe_m_sic2_error_abs=abs(v_fpe_mean_sic2-prc4)/prc4
         /// PE - Harmonic Mean - SIC 3
         gen fpe_hm_sic3_error_signed=(v_fpe_harm_sic3-prc4)/prc4
         gen fpe_hm_sic3_error_abs=abs(v_fpe_harm_sic3-prc4)/prc4
         /// PE - Mean - SIC 3
         gen fpe_m_sic3_error_signed=(v_fpe_mean_sic3-prc4)/prc4
```

```
gen fpe_m_sic3_error_abs=abs(v_fpe_mean_sic3-prc4)/prc4
// --------------------------------------------------------
/// PBV - Harmonic Mean - SIC 2
gen pbv_hm_sic2_error_signed=(v_pbv_harm_sic2-prc4)/prc4
gen pbv_hm_sic2_error_abs=abs(v_pbv_harm_sic2-prc4)/prc4
/// PBV - Mean - SIC 2
gen pbv_m_sic2_error_signed=(v_pbv_mean_sic2-prc4)/prc4
gen pbv_m_sic2_error_abs=abs(v_pbv_mean_sic2-prc4)/prc4
/// PBV - Harmonic Mean - SIC 3
gen pbv_hm_sic3_error_signed=(v_pbv_harm_sic3-prc4)/prc4
gen pbv_hm_sic3_error_abs=abs(v_pbv_harm_sic3-prc4)/prc4
/// PBV - Mean - SIC 3
gen pbv_m_sic3_error_signed=(v_pbv_mean_sic3-prc4)/prc4
gen pbv_m_sic3_error_abs=abs(v_pbv_mean_sic3-prc4)/prc4
/// 14. Tabulate the errors calculated in step 13
/// Note:
tabstat fpe_hm_sic2_error_signed fpe_hm_sic2_error_abs fpe_m_sic2_error_signed fpe_m_sic2_error_abs fpe_hm_sic3_error_signed fpe_hm_sic3_error_abs fpe_m_sic3_error_signed fpe_m_sic3_error_abs pbv_hm_sic2_error_signed pbv_hm_sic2_error_abs pbv_m_sic2_error_signed pbv_m_sic2_error_abs pbv_hm_sic3_error_signed pbv_hm_sic3_error_abs pbv_m_sic3_error_signed pbv_m_sic3_error_abs, stat(n mean sd min p1 p25 p50 p75 p99 max) col(stat)
/// 15. Create and drop 1st and 100th %tile of PE and PB value estimates
/// Note:
/// PE - Harmonic Mean - SIC 2
xtile xfpe_hm_sic2_error_signed=fpe_hm_sic2_error_signed, nq(100)
drop if xfpe_hm_sic2_error_signed==1 | xfpe_hm_sic2_error_signed==100
xtile xfpe_hm_sic2_error_abs=fpe_hm_sic2_error_abs, nq(100)
drop if xpe_hm_sic2_error_abs==100
/// PE - Mean - SIC 2
xtile xfpe_m_sic2_error_signed=fpe_m_sic2_error_signed, nq(100)
drop if xfpe_m_sic2_error_signed==1 | xfpe_m_sic2_error_signed==100
xtile xfpe_m_sic2_error_abs=fpe_m_sic2_error_abs, nq(100)
drop if xfpe_m_sic2_error_abs==100
/// PE - Harmonic Mean - SIC 3
xtile xfpe_hm_sic3_error_signed=fpe_hm_sic3_error_signed, nq(100)
drop if xfpe_hm_sic3_error_signed==1 | xfpe_hm_sic3_error_signed==100
xtile xfpe_hm_sic3_error_abs=fpe_hm_sic3_error_abs, nq(100)
drop if xfpe_hm_sic3_error_abs==100
/// PE - Mean - SIC 3
xtile xfpe_m_sic3_error_signed=fpe_m_sic3_error_signed, nq(100)
drop if xfpe_m_sic3_error_signed==1 | xfpe_m_sic3_error_signed==100
xtile xfpe_m_sic3_error_abs=fpe_m_sic3_error_abs, nq(100)
```

drop if xfpe_m_sic3_error_abs==100

/// ---

/// PBV - Harmonic Mean - SIC 2

xtile xpbv_hm_sic2_error_signed=pbv_hm_sic2_error_signed, nq(100)

drop if xpbv_hm_sic2_error_signed==1 | xpbv_hm_sic2_error_signed==100

xtile xpbv_hm_sic2_error_abs=pbv_hm_sic2_error_abs, nq(100)

drop if xpbv_hm_sic2_error_abs==100

/// PBV - Mean - SIC 2

xtile xpbv_m_sic2_error_signed=pbv_m_sic2_error_signed, nq(100)

drop if xpbv_m_sic2_error_signed==1 | xpbv_m_sic2_error_signed==100

xtile xpbv_m_sic2_error_abs=pbv_m_sic2_error_abs, nq(100)

dorp if xpbv_m_sic2_error_abs==100

/// PBV - Harmonic Mean - SIC 3

xtile xpbv_hm_sic3_error_signed=pbv_hm_sic3_error_signed, nq(100)

drop if xpbv_hm_sic3_error_signed==1 | xpbv_hm_sic3_error_signed==100

xtile xpbv_hm_sic3_error_abs=pbv_hm_sic3_error_abs, nq(100)

drop if xpbv_hm_sic3_error_abs==100

/// PBV - Mean - SIC 3

xtile xpbv_m_sic3_error_signed=pbv_m_sic3_error_signed, nq(100)

drop if xpbv_m_sic3_error_signed==1 | xpbv_m_sic3_error_signed==100

xtile xpbv_m_sic3_error_abs=pbv_m_sic3_error_abs, nq(100)

drop if xpbv_m_sic3_error_abs==100

/// 16. Run again tabstat to checki if mean and median improved after dropping 1st and 100th %tile

/// Note:

tabstat fpe_hm_sic2_error_signed fpe_hm_sic2_error_abs fpe_m_sic2_error_signed fpe_m_sic2_error_abs fpe_hm_sic3_error_signed fpe_hm_sic3_error_abs fpe_m_sic3_error_signed fpe_m_sic3_error_abs pbv_hm_sic2_error_signed pbv_hm_sic2_error_abs pbv_m_sic2_error_signed pbv_m_sic2_error_abs pbv_hm_sic3_error_signed pbv_hm_sic3_error_abs pbv_m_sic3_error_signed pbv_m_sic3_error_abs, stat(n mean sd min p1 p25 p50 p75 p99 max) col(stat)

/// 17. Using t-test to check for equility of two means and the equality of two medians (paired testing method)

/// Note:

/// PE - Harmonic Mean - SIC 2 // PBV - Harmonic Mean - SIC 2

ttest fpe_hm_sic2_error_signed=pbv_hm_sic2_error_signed

ttest fpe_hm_sic2_error_abs=pbv_hm_sic2_error_abs

/// PE - Mean - SIC 2 // PBV - Mean - SIC 2

ttest fpe_m_sic2_error_signed=pbv_m_sic2_error_signed

ttest fpe_m_sic2_error_abs=pbv_m_sic2_error_abs

/// PE - Harmonic Mean - SIC 3 // PBV - Harmonic Mean - SIC 3

ttest fpe_hm_sic3_error_signed=pbv_hm_sic3_error_signed

ttest fpe_hm_sic3_error_abs=pbv_hm_sic3_error_abs

/// PE - Mean - SIC 3 // PBV - Mean - SIC 3

ttest fpe_m_sic3_error_signed=pbv_m_sic3_error_signed ttest fpe_m_sic3_error_abs=pbv_m_sic3_error_abs

/// 18. Using Wilcoxon Test

/// Note:

// PE - Harmonic Mean - SIC 2 // PBV - Harmonic Mean - SIC 2

signrank fpe_hm_sic2_error_signed=pbv_hm_sic2_error_signed

signrank fpe_hm_sic2_error_abs=pbv_hm_sic2_error_abs

// PE - Mean - SIC 2 // PBV - Mean - SIC 2

signrank fpe_m_sic2_error_signed=pbv_m_sic2_error_signed

signrank fpe_m_sic2_error_abs=pbv_m_sic2_error_abs

// PE - Harmonic Mean - SIC 1 // PBV - Harmonic Mean - SIC 3

signrank fpe_hm_sic3_error_signed=pbv_hm_sic3_error_signed

signrank fpe_hm_sic3_error_abs=pbv_hm_sic3_error_abs

// PE - Mean - SIC 1 // PBV - Mean - SIC 3

signrank fpe_m_sic3_error_signed=pbv_m_sic3_error_signed

signrank fpe_m_sic3_error_abs=pbv_m_sic3_error_abs

/// 19. From now on - RIVM & DDM

/// 19.1 Tabulate the variables used both in RIVM and DDM as done with the multiples.

/// Note:

/// 19.2 The cost of equity is 10%, the forecast horizon is 3 years and the terminal value is a growing (future residual income growth rate is 4%)

/// Note:

gen rf=.

gen rpm=.

replace rf=0.0439 if fyear==2005

replace rf=0.0470 if fyear==2006

replace rf=0.0402 if fyear==2007

replace rf=0.0221 if fyear==2008

replace rf=0.0384 if fyear==2009

replace rf=0.0329 if fyear==2010

replace rf=0.0188 if fyear==2011

replace rf=0.0176 if fyear==2012

replace rf=0.0304 if fyear==2013

replace rf=0.0217 if fyear==2014

replace rf=0.0227 if fyear==2015

replace rpm=0.0220 if fyear==2005

replace rpm=0.0197 if fyear==2006

replace rpm=0.0206 if fyear==2007

replace rpm=0.0405 if fyear==2008

replace rpm=0.0260 if fyear==2009

replace rpm=0.0224 if fyear==2010

replace rpm=0.0271 if fyear==2011

replace rpm=0.0247 if fyear==2012

```
replace rpm=0.0203 if fyear==2013
replace rpm=0.0224 if fyear==2014
replace rpm=0.0246 if fyear==2015
gen re=.
replace re=rf+beta*rpm
/// 19.3. Calculate book value of equity for year 2 and 3
/// Note:
gen bvps_2=bvpsadj+(1-divpayout)*eps_1_adj
gen bvps_3=bvps_2+(1-divpayout)*eps_2_adj
gen bvps_4=bvps_3+(1-divpayout)*eps_3_adj
/// 19.4 Calculate median of long term forecast rate eps_g for each industry separately
/// Note:
gen eps_g_duplicate=eps_g/100
egen eps_g_industry=mean(eps_g_duplicate), by (sic_2 fyear)
/// 19.5 Calculate eps_3_new whcih is all eps_3 values including missing values with imputed growth rate calculated as industry mean growth rate
/// Note:
gen eps_3_new=eps_3_adj
replace eps_3_new=eps_2_adj*(1+eps_g_industry) if eps_3_adj==.
gen eps_4_new=eps_3_new*(1+eps_g_industry)
/// 19.6 Terminal Value Growth Rate
/// Note:
gen g0=0
gen g4=0.04
/// 19.7 Calculate intrinsic value with RIVM for each industry separately
/// Note:
/// g=0%
gen v_rivm_g0=bvpsadj+(eps_1_adj-re*bvpsadj)/(1+re)+(eps_2_adj-re*bvps_2)/(1+re)^2+(eps_3_new-re*bvps_3)/(1+re)^3+(eps_4_new-re*bvps_4)/(re-g0)/(1+re)^4
///  g=4%
gen v_rivm_g4=bvpsadj+(eps_1_adj-re*bvpsadj)/(1+re)+(eps_2_adj-re*bvps_2)/(1+re)^2+(eps_3_new-re*bvps_3)/(1+re)^3+(eps_4_new-re*bvps_4)/(re-g4)/(1+re)^4
/// 19.8 Drop negative value estimates
/// Note:
gen v_rim_1_g0=v_rivm_g0 if v_rivm_g0>0
gen v_rim_1_g4=v_rivm_g4 if v_rivm_g4>0
/// 19.9 Calculate absolute and signed valuation errors by industry
/// Note:
/// g=0%
gen v_rim_g0_error_signed=(v_rim_1_g0-prc4)/prc4
gen v_rim_g0_error_abs=abs(v_rim_1_g0-prc4)/prc4
/// g=4%
```

```
gen v_rim_g4_error_signed=(v_rim_1_g4-prc4)/prc4
gen v_rim_g4_error_abs=abs(v_rim_1_g4-prc4)/prc4
/// 20. Dividend Discount Model
/// Note:
/// 20.1 Generate Dividends per Share Long Term Growth Rate by SIC3
/// Note:
gen dps_g_duplicate=dps_g/100
egen dps_g_industry=mean(dps_g_duplicate), by (sic_3 fyear)
/// 20.2 Generate dps_3 missing values using the LT growth rate
/// Note:
gen dps_3_new=dps_3_adj
replace dps_3_new=dps_2_adj*(1+dps_g_industry) if dps_3_adj==.
gen dps_4_new=dps_3_new*(1+dps_g_industry)
/// 20.3 Calculate Intrinsic Value Using DDM
/// Note:
///  g=0%
gen v_ddm_0_g0=dps_1_adj/(1+re)+dps_2_adj/(1+re)^2+dps_3_new/(1+re)^3+(dps_4_new)/(re-g0)/(1+re)^4
gen v_ddm_1_g0=v_ddm_0_g0 if v_ddm_0_g0>0
///  g=4%
gen
v_ddm_0_g4=dps_1_adj/(1+re)+dps_2_adj/(1+re)^2+dps_3_adj/(1+re)^3+(dps_3_new*(1+dps_g_industry))/(re-
g4)/(1+re)^4
gen v_ddm_1_g4=v_ddm_0_g4 if v_ddm_0_g4>0
/// 20.4 Calculate absolute and signed valuation errors by industry
/// Note:
gen v_ddm_g0_error_signed=(v_ddm_1_g0-prc4)/prc4
gen v_ddm_g0_error_abs=abs(v_ddm_1_g0-prc4)/prc4
/// ---------------------------------------------------
gen v_ddm_g4_error_signed=(v_ddm_1_g4-prc4)/prc4
gen v_ddm_g4_error_abs=abs(v_ddm_1_g4-prc4)/prc4
/// 20.5 Check for outliers
/// Note:
tabstat v_rim_g0_error_signed v_rim_g0_error_abs v_rim_g4_error_signed v_rim_g4_error_abs
v_ddm_g0_error_signed v_ddm_g0_error_abs v_ddm_g4_error_signed v_ddm_g4_error_abs, stat(n mean sd min p1
p25 p50 p75 p99 max) col(stat)
/// 21. Create and drop 1st and 100th %tile of RIVM and DDM value estimates
/// Note:
// RIVM - g=0%
xtile xv_rim_g0_error_signed=v_rim_g0_error_signed, nq(100)
drop if xv_rim_g0_error_signed==1 | xv_rim_g0_error_signed==100
xtile xv_rim_g0_error_abs=v_rim_g0_error_abs, nq(100)
drop if xv_rim_g0_error_abs==100
```

```
// RIVM - g=4%
xtile xv_rim_g4_error_signed=v_rim_g4_error_signed, nq(100)
drop if xv_rim_g4_error_signed==1 | xv_rim_g4_error_signed==100
xtile xv_rim_g4_error_abs=v_rim_g4_error_abs, nq(100)
drop if xv_rim_g4_error_abs==100
// DDM - g=0%
xtile xv_ddm_g0_error_signed=v_ddm_g0_error_signed, nq(100)
drop if xv_ddm_g0_error_signed==1 | xv_ddm_g0_error_signed==100
xtile xv_ddm_g0_error_abs=v_ddm_g0_error_abs, nq(100)
drop if xv_ddm_g0_error_abs==100
// DDM - g=4%
xtile xv_ddm_g4_error_signed=v_ddm_g4_error_signed, nq(100)
drop if xv_ddm_g4_error_signed==1 | xv_ddm_g4_error_signed==100
xtile xv_ddm_g4_error_abs=v_ddm_g4_error_abs, nq(100)
drop if xv_ddm_g4_error_abs==100
/// 22. Summary statistic of errors after trimming for outliers
/// Note:
tabstat v_rim_g0_error_signed v_rim_g0_error_abs v_rim_g4_error_signed v_rim_g4_error_abs v_ddm_g0_error_signed v_ddm_g0_error_abs v_ddm_g4_error_signed v_ddm_g4_error_abs, stat(n mean sd min p1 p25 p50 p75 p99 max) col(stat)
/// 23. Using t-test to check for equality of two means and the equality of two means (paired testing method)
/// Note:
/// G=0%
ttest v_rim_g0_error_signed=v_ddm_g0_error_signed
ttest v_rim_g0_error_abs=v_ddm_g0_error_abs
/// G=4%
ttest v_rim_g4_error_signed=v_ddm_g4_error_signed
ttest v_rim_g4_error_abs=v_ddm_g4_error_abs
/// Constructing valuation error deciles for RIVM - g=0%///
xtile xv_rim_g0_error_abs_dec=v_rim_g0_error_abs, nq(10)
gen high_error_rim=1 if xv_rim_g0_error_abs_dec==9 | xv_rim_g0_error_abs_dec==10
replace high_error_rim=0 if xv_rim_g0_error_abs_dec==1 | xv_rim_g0_error_abs_dec==2
/// Constructing valuation error deciles for P/FE - SIC_3
xtile xfpe_hm_sic3_error_abs_dec=fpe_hm_sic3_error_abs, nq(10)
gen high_error_fpe=1 if xfpe_hm_sic3_error_abs_dec==9 | xfpe_hm_sic3_error_abs_dec==10
replace high_error_fpe=0 if xfpe_hm_sic3_error_abs_dec==1 | xfpe_hm_sic3_error_abs_dec==2
```