HIT OR MISS:
A STUDY OF ANALYSTS' TARGET PRICE ACCURACY
DANISH TITLE: ANALYSE AF NØJAGTIGHEDEN AF ANALYTIKERES AKTIEKURSMÅL

A THESIS PRESENTED FOR THE DEGREE:
MASTER OF SCIENCE IN FINANCE AND ACCOUNTING (CAND. MERC. FIR)

COPENHAGEN BUSINESS SCHOOL
SUPERVISOR: **PROFESSOR THOMAS PLENBORG**
MAY 2021

AUTHORS:
CHRISTOFFER GJESSING TERKILDSEN (102830)
OLIVER LØNSTRUP THORSEN (133937)

NUMBER OF STANDARD PAGES: 105
NUMBER OF CHARACTERS: 232.800

Abstract

I dette speciale undersøger vi sammenhængen mellem nøjagtigheden af aktiekursmål og forskellige forklarende variable. Gennem en litteraturgennemgang når vi frem til 20 forskellige variable som vi mener potentielt kan have indflydelse på nøjagtigheden af aktiekursmål, alle delt op på enten analytiker specifikke, finansielle eller metodiske variable. Fremfor at undersøge faktiske aktieanalyser fra professionelle aktieanalytikere, undersøger vi værdiansættelser begået af studerende som led i deres afsluttende afgangsprojekt på Copenhagen Business School.

Litteraturen viser, at professionelle analytikere deltager i et principal-agentproblem mellem arbejdsgiver og investorer, der leder analytikere til at udgive kursmål med positiv bias. Da vores speciale fokuserer på studerende, er vores hypotese at de ikke er udsat for de samme incitamenter som professionelle analytikere, og dermed kan vores undersøgelser danne baggrund for en unik indsigt i hvordan f.eks. valg af værdiansættelsesmodel, budgetadfærd, uddannelsesbaggrund, køn og antagelser gjort i terminalværdileddet påvirker nøjagtigheden af aktiekursmål uden påvirkning fra ovenstående faktor.

Vores dataset består af 321 værdiansættelser af danske og udenlandske børsnoterede selskaber. Vi benytter os af gængse statistiske metoder, hvor vi kontrollerer for potentielle variable der kan påvirke vores konklusioner, til at undersøge vores hypoteser. Vi finder bl.a. at uddannelsesbaggrund, køn, historisk omsætningsvolatilitet, valg af værdiansættelsesmodel(ler), den historiske periode medtaget i projektet og sammenhængen mellem ROIC og WACC i terminalperioden spiller en rolle i nøjagtigheden af studerendes aktiekursmål. Vi finder også at studerende generelt er unøjagtige og at de udviser konsistent positiv bias i deres kursmål.

Imens konklusionen på vores indledende analyse, som viser at studerende generelt er unøjagtige og udviser positivt bias, antages at være robust, lider styrken af vores konklusioner omhandlende forklarende variable generelt under, at vi ikke har kunne kontrollere for alle variable, der potentielt kan påvirke vores resultater. Dette skyldes til dels generelle begrænsninger i størrelsen af vores datasæt, når vi kontrollerer for disse variable. Vi foreslår derfor en mere gennemgribende gennemgang af disse forhold i fremtidige undersøgelser, for at sikre stærkere og mere robuste konklusioner.

Table of Contents

List of Abbreviations

AEF	M.Sc. Applied Economics & Finance (Study programme)
ANOVA	Analysis of Variance
ASC	M.Sc. Accounting, Strategy & Control (Study programme)
ATO	Asset Turnover
AUD	M.Sc. Auditing (Study programme)
CAGR	Compound Average Growth Rate
CAPM	Capital Asset Pricing Model
DCF	Discounted cash flow (model)
EBIT	Earnings Before Interest and Tax
EIR	Excess Implied Return
EPS	Earnings per Share
EV	Enterprise Value
EVA	Economic Value Added
FC	Forecast
FCFF	Free cash flow to the firm
FIN	M.Sc. Finance and Investments (Study programme)
FIR	M.Sc. Finance and Accounting (Study programme)
GDP	Gross Domestic Product
HD_AFM	Graduate Diploma in Accounting and Financial Management (Study progr.)
HD_F	Graduate Diploma in Finance (Study programme)
IC	Invested capital
ICR	Intercoder Reliability
IQR	Interquartile range
LBO	Leveraged Buyout
NIBD	Net Interest-Bearing Debt
NOPAT	Net Operating Profit after Tax
NOPLAT	Net Operating Profit Less Adjusted Tax
OLS	Ordinary least squares
RI	Residual Income
ROIC	Return on Invested Capital
RSD	Relative Standard Deviation
STPFE	Standardized Target Price Forecast Error
TPA	Target Price Accuracy
TPFE	Target Price Forecast Error
TV	Terminal Value
TY	Terminal Year
WACC	Weighted Average Cost of Capital

List of Figures

List of Tables

1.0 Introduction

Valuation might be based on sound theoretical principles and numerical inputs, but it is not an exact science. Academics, financial analysts, and other professionals might have access to the same information and, to a large degree, utilize the same set of models, but the outputs will often differ. Valuation is heavily reliant on the assumptions made and the methodology applied. Further, analysts are notoriously known for making fundamental errors in their models (Petersen & Plenborg, 2008; Green et al., 2016), as well as exhibiting positive bias (Abarbanell, 1991; Brown, Foster, & Noreen, 1985; Stickel, 1990), all affecting the valuation output. The purpose of valuation is fundamentally to estimate fair prices, and any wrongdoing can have non-trivial consequences. Pensions and personal savings are to a wide extent tied up in equity investments and other financial instruments whose value depends on fair prices. Prices must reflect risk and the assets true value, if not, unsustainable financial markets can consequently lead to financial crises and other loses.

A 2017 report by McKinsey highlighted that equity research spending is on a decline, as asset owners are turning increasingly toward passive investment strategies as a consequence of investment managers and analysts struggling to deliver excess returns (McKinsey & Company, 2017). Further, several scandals involving equity analysts have emerged in the past couple of years. In Denmark, one of the most notable was arguably the Pandora IPO in 2011, where 12 of 14 analysts associated with the facilitating bank issued buy recommendations prior to the IPO (Nielsen, 2011), just to find the stock lose 80% of its value less than a year after. More recently, a scan of eight danish IPOs on the Nasdaq First North exchange in 2021 reveals that only three stocks have stayed above IPO prices, with the remaining five being down -28,8% on average. Whether these are results of poor analyst performance or downright wrongdoing is up for debate. However, it begs the question whether analyst target price forecasts are essentially accurate and usable as investment guidance.

The cause of this debate though, is profoundly clear; the true value of an asset is an intangible concept. Followers of the efficient market hypothesis perceive the freely formed prices in the financial markets to be a good approximation, while some analysts beg to differ. Through thorough analysis and due diligence, they believe they can derive a better approximation, and that financial markets eventually will come to the same conclusion (Damodaran, 2002). The existing literature is generally not too positive regarding the performance of analysts. The strand of literature focusing on the accuracy of valuations often revolve around equity analysts and their target prices (Bradshaw, Brown, & Huang, 2013; Imam, Chan, & Ali Shah, 2013; etc.) where analysts are generally shown to be overly optimistic and inaccurate. All is not lost though. On the positive side, Gleason, Johnson and Li (2011) state that analysts can significantly improve their target price forecasts using a rigorous valuation technique.

Another strand of literature focuses on the incentive structure regarding analysts (Dechow, Hutton, & Sloan, 2000; Hong & Kubik, 2003; Lin & McNichols, 1998; McNichols & O'Brien, 1997). Here, analysts are shown to partake in a principal-agent problem, where analyst incentives lead them to publish optimistic forecasts, in order to secure and maintain lucrative relationships with company management, ultimately to the disadvantage of investors. Our subject of analysis will however not be professional analyst, but rather students, who as a part of their final project have performed valuations of various companies. Our hope is that this will minimize the effect of the incentive structures usually applicable to analysts, and thus give us less biased results. Without these incentives present, we can solely focus on the model input and its implications on model output.

Petersen & Plenborg (2008) distinguishes between two types of erros; estimation error and methodological errors. The latter is due to misinterpretation of the theoretical foundation of the models, or simple errors in the implementation of the model. The former is of less concrete matter, as it relies on the analysts' ability to succesfully interpret the current information at hand, and extrapolate this into an estimation of the company's future performance. In our study we seek to uncover both of these errors, and their effect on the attained accuracy in the models. We will look at the assumptions applied in the forecasts, as well as examine how common errors and theoretical misconceptions affect accuarcy. We will also be looking at different analyst characteristics in this context. This lead us to the two following research questions:

i. "Are target prices derived from student valuations accurate and unbiased?"
ii. "Which factors influence the accuracy or bias derived from student valuations?"

The first question is a general assessment of the accuracy and bias present in student valuations, while the second seeks to uncover the specific factors influencing accuracy and bias. Potential factors that might influence accuracy and bias will be uncovered through a thorough literature review in section 2. In section 3, the specific hypotheses will be formed. Section 4 provides us with the methodological foundation of this thesis, while section 5 is a presentation of our results. Finally, section 6 is a discussion of the results in relation to existing literature, and section 6 is the thesis conclusion.

2.0 Literature review

The objective of this thesis is to uncover the degree of accuracy or bias of analysts' valuations, and if possible, measure the degree of impact of certain theorized explanatory variables on the accuracy or bias uncovered. Therefore, a thorough study of the relevant literature is essential.

A multitude of research has been published, both on earnings forecasts and target price (in)accuracies of analysts. Previous research on the field suggests that analysts are overly optimistic, and thus exhibit positive bias – both in earnings (Abarbanell, 1991; Brown, Foster, & Noreen, 1985; Stickel, 1990) and accuracy measures (Bradshaw, Brown, & Huang, 2013; Imam, Chan, & Ali Shah, 2013). For instance, Imam et al. (2013) found overall target price accuracy to be 49,09% in their study, i.e., that more than half of the analysts in the study did not release accurate target price estimates on a one-year horizon. Bradshaw et al. (2013) found that only 38% of target prices are met during a one-year horizon, with target price forecast errors averaging 45% (the deviation from the realized price for non-met target prices). Thus, it seems well established that analysts are neither accurate nor unbiased when providing earnings and target price estimates. Intuitively, inaccuracies and biases are bound to happen – after all, target prices are, using most conventional methods, an act of forecasting future earnings potential, and as such must be prone to error. However, several variables of interest exist when trying to explain these inaccuracies.

To structure our literature review, and ultimately also our analysis, we have sought to consider explanatory variables that have been of impact in previous literature on firm value and the ability to perform target price forecasts – from the analysis and understanding of financial metrics to methodological and analyst specific variables. Therefore, throughout this section, four general themes will be discussed:

I. **Measures of accuracy and bias:** How has it been done, and how we will do it.
II. **Analyst characteristics:** The role of analysts, and the analyst specific characteristics that have previously been studied.
III. **Financial factors:** An overview of the financial metrics under scrutiny in this study.
IV. **Methodological factors:** The assumptions and best practices proposed in literature when performing firm valuations.

2.1 Measures of accuracy and bias

When reviewing the literature on analyst accuracy and bias, two strands seem to emerge; the literature on earnings forecast accuracy (or bias), and the literature of target price accuracy (or bias).

Much research has been conducted on both strands. For earnings forecast accuracy, commonplace research suggests that analysts exhibit optimistic or positive bias, meaning that the earnings forecasts exceed actual earnings in hindsight. For target price accuracy, similar findings of positive biases exist. As mentioned, Imam et al. (2013) found overall target price accuracy to be 49,09% when looking at a one-year horizon, and Bonini, Zanetti, Bianchini, & Salvi (2010) found that target price prediction errors are consistent and large, with errors up to 36,6%. Similar results were found by Asquith, Mikhail, & Au (2005), concluding that target price forecasts are only accurate in 45,7% of cases.

One overarching problem within the literature, however, is the question of deriving an appropriate and reliable measure of accuracy and bias. For both strands of literature – those interested in target price forecasts, and those interested in earnings forecasts, these measurements are of concern. Most researchers within the earnings forecast category have opted for the solution of measuring relative distance between earnings forecasts (typically by use of EPS) and realized earnings, and thus analyses the accuracy in hindsight at the end of a given period – typically 12 months (Brown & Mohd, 2003; Hilary & Hsu, 2013). This method requires accurate reconciliation of earnings, with one-off earnings or costs eliminated, and comfort that the financials figures from which the forecast is based, is readily available in similar form (e.g., non-adjusted, no changes in accounting methods), at both the time of forecast publication and at the time of the determined period post forecast release.

For target price accuracy researchers, the above concerns have been less of a hindrance. Asquith et al. (2005) proposed an accuracy metric measuring whether the target price has been met (or exceeded) at the end of a given time horizon. One problem associated with this method is however, that the underlying share price could hit the target price sometime in-between the date of target price publication and the end of the time horizon, and still deem the analysts as being inaccurate. This issue was circumvented by Bonini et al. (2010) who, by similar measures, checked whether the price was met at any time within the time horizon, by including high and low prices for the horizon as the points of measurement. However, another issue emerged regarding the measurement of target prices – that of picking an appropriate time horizon. Bonini et al. (2010) manually inspected more than 9.800 analyst reports and found that analysts, in general, do not add explicit time-periods for when the market is expected to adjust to fundamental (or target) values. However, when analysts do explicitly state a horizon, it is often 12 months post the issue date of the report. Intuitively, the time horizon is not without interest, as corporate events (scandals, stakeholder defaults etc.) and market externalities (financial crises, pandemics etc.) can trigger large and unpredicted fluctuations in share prices. Richardson, Teoh, & Wysocki (1999) found that, while optimistic earnings forecast biases exists, they seem to be amplified in long-term forecasts (e.g., yearly forecasts), and smaller in shorter forecasts (e.g., quarterly forecasts). In our thesis, we will seek to utilize a similar metric to Bonini et al. (2010), namely that of yearly high and low prices as determinants of a binary target price accuracy measure (target price hit or not hit), and thus we will include intra-year share price movements in determining accuracy. We will however expand on previous literature by including 12, 24 and 36-month horizons post target price publication date, if available, to see whether the findings of Richardson et al. (1999) hold true on target price accuracies as well.

While the question of whether the underlying stock hits the target price or not on an intra-year perspective is interesting, there is still a need for measuring the forecast errors of target prices. From our review of the literature, this measurement exists in various forms, and with great debates to follow on the utility of each measure. For instance, Patel (2018) proposes a simple metric for Target Price Accuracy (TPAcr), where TPAcr is defined as the percentage of the target price that is achieved by the underlying stock. The downfall here is that only buy (long) recommendations are included, and all short or hold recommendations are stripped from the dataset. Sayed (2015) circumvents this issue and proposes a measure of Target Price Forecast Error (TPFE), where TPFE is equal to zero in cases where the target price has been hit, judged by intra-year high/low. He then provides a percentage error by taking the difference between the 52-week high (or low, depending on buy/sell recommendation at T_0) and the target price, divided by the target price, in cases where the target price has not been hit. Thereby, Sayed does acknowledge that the target price can be hit sometime within the time horizon, and not necessarily at the end of the horizon – however, this would also introduce a possible bias in TPFE by a lot of zero-error measurements, as well as the incentive for analysts to publish conservative price targets, as those are – all else equal – more likely to be met, and thus give zero-error measurements.

Another measure is seen in Bonini et al. (2010), who also includes both short and long recommendations, but measures the relative deviation from the target price rather than absolute deviation. Finally, Kerl (2011) uses a similar measure to that of Patel (2018), but also includes short recommendations. An overview of the measurements is presented in the table below.

Table 1: Target price forecast accuracy and forecast error measures from literature

Issue #1					**Issue #2**		
Assumptions	Scenario 1	Scenario 2	Scenario 3	Scenario 4	**Assumptions**	Scenario 1	Scenario 2
Price at t_0	100	100	140	140	Price at t_0	100	100
Target price	120	120	120	120	Target price	110	110
Recommendation	LONG	LONG	SHORT	SHORT	Recommendation	LONG	LONG
P_{th}	115	125	115	125	P_{th}	109	115
Target Price Error	Scenario 1	Scenario 2	Scenario 3	Scenario 4	**Target Price Error**	Scenario 1	Scenario 2
Patel (2018)	95,83%	104,17%	na	na	Patel (2018)	99,09%	104,55%
Sayed (2015)	4,17%	0%	0%	4,17%	Sayed (2015)	0,91%	0,00%
Bonini et al. (2010)	4,35%	-4,00%	-4,35%	4,00%	Bonini et al. (2010)	0,92%	-4,35%
Kerl (2011)	-95,83%	-104,17%	95,8%	104,17%	Kerl (2011)	-99,09%	-104,55%

t_0 = Time of valuation publication

P_{th} = Price at end of time horizon

As evident from issue #1 in Table 1, when considering all four scenarios, great variety is present among the measures. Patel, Sayed and Kerl all consider the absolute measure, and thus has perfect error symmetry regardless of over/under prediction. Bonini et al. measures relative values, creating asymmetric forecast errors. Essentially, this means that a nominal 5 USD target price error is a larger forecast error when overestimating the target price, than when underestimating it. In our view, a measurement model intended to show analyst forecast error should seek to highlight the deviation from the actual share price, regardless of the direction of the deviation. Another notable difference between the measurements is seen within scenario 2 and 3 using Sayed's TPFE. Here, both analysts have a target price error of 0%, as both target prices have been hit sometime during the time horizon, although a deviation is present at the end of the horizon. The downfall of this measurement is illustrated to a greater degree in issue #2, where a 1 USD nominal deviation, in Sayed's model, generates a 0.91% error, while a 5 USD nominal deviation has zero-error, ultimately leading to inaccurate results when seeking to measure analysts' absolute target price forecast errors.

In our thesis, we will, as proposed by Sayed, utilize intra-year highs and lows on a binary Target Price Accuracy measure, as we acknowledge that hitting the target price intra-year is a relevant measure. However, we will, as proposed by Kerl, Patel and Bonini et al., utilize closing prices as measurement points at end-of-horizon when measuring forecast error, as to decrease the interference of zero-error target prices and thus eliminate discrimination between positive or negative error deviations.

2.2 Analyst characteristics

The role of the analyst has long been studied by academics. Doukas, Kim, & Pantzalis (2005) found analysts to bear an important role in markets, as they provide a key source of information for investors, and thus play an important role in keeping stocks close to their fundamental values. Amiram, Owens, & Rozenbaum (2016) documented that analysts' earnings forecasts decrease the information asymmetry at the time of announcement. Thus, analysts arguably possess an important role in the functioning of markets, as providers of information, which cements the need for accurate analyst estimates to improve investors' basis for decision-making. Also, for the analysts' themselves', accuracy is of utmost importance. Stickel (1992) showed that more accurate analysts were receiving greater professional recognition, and Hong & Kubik (2003) showed that they were subject to better career outcomes.

Analyst biases

Considering the extent of literature on systematic analyst bias, the reason for said bias is naturally of interest. Lim (2001) found that the magnitude of systematic bias could be extrapolated and predicted from public information – namely by uncovering that the size of the firm, and thus, he argued, richness of the information environment, was negatively correlated with forecast bias. On the contrary, Doukas, Kim, & Pantzalis (2005) found that stocks with strong analyst coverage was associated with overvaluation and low future returns. A large strand of research on the bias of analysts tend to focus on the flow of information, and the relationships necessary to sustain it. Lim (2001) argues that the company in question is a key source of information for analysts, and thus, they fear a hinderance in future flow of information if unfavorable forecasts are issued. This notion is backed by earlier research (Dechow, Hutton, & Sloan, 2000; Hong & Kubik, 2003; Lin & McNichols, 1998; McNichols & O'Brien, 1997), arguing that the incentives analysts face leads to optimistic forecasts, in order to secure and maintain lucrative relationships with company management. Further, Huang et al. (2017) found analysts to show herding behavior, and that 60% of analyst herd toward the current consensus estimate. This is found to be associated with analyst following and the size of their employer. Nautrally, as our sample comprises of students, as proxies for analysts, neither of these relationships or incentive structures are assumed to be of influence in the potential biases uncovered – therefore, this thesis will seek to uncover analyst characteristics besides their affiliation and relationships with management.

Effect of teamwork and gender

Besides the strand of research focusing on analyst affiliation and the flow of information, recent research has investigated the effect of teamwork and gender on the analysts' ability to generate accurate forecasts. For instance, Hope & Fang (2020), showed that team size (more than one analyst) and diversity (educational background, experience and gender) are significantly correlated with both target price forecast and earnings forecast accuracy. Earlier research suggests that the gender of analysts does indeed matter. Bosquet, de Goeij, & Smedts (2014) found that, while controlling for differences in risk characteristics of recommended stocks, analyst experience and task complexity, female analysts were 40% less likely to issue optimistic investment advice (buy or strong-buy recommendations) than their male counterparts, thus concluding that female analysts are more conservative than male analysts. On the contrary, Green, Jagadeesh, & Tang (2009) found that absolute forecast errors were slightly larger for female analysts than for male analysts. Through our thesis, we will seek to illuminate whether the gender, size of team and educational background (study programme) can help explain target price accuracy.

2.3 Financial variables

When investigating analyst accuracy and bias, it is helpful to decompose the inputs of a typical firm valuation, as to gain a more nuanced view of the process and possible explanatory variables. Here, it helps to look at the theoretical framework underlying firm valuations, as only then the financial inputs can be specified. Both in research and practice, there exists a large number of methods – ranging from traditional cash flow-based models such as the Discounted Cash Flow analysis, to more obscure and sometimes highly industry-specific relative valuation models, such as the "Enterprise Value to Kilos of Salmon Harvested"-multiple used in one of the valuations in our sample data.

Overall, it is possible to group the plethora of valuation techniques into four primary categories, namely: Present value models[1], relative models (multiples), net asset-based models and real-options models (Petersen, Plenborg, & Kinserdal, 2017). Valuation researchers have long advocated for the use of present value models (Damodaran, 2002; Penman, 1997), and in practice, these are commonly used. In a survey presented by Petersen, Plenborg, & Kinserdal (2017), the frequency of valuation models amongst practitioners were uncovered, with nearly all respondents in the sample using present value models, approximately 90% using relative valuation models (multiples), and less than one-third using asset-based valuation models or contingent claim valuation. This share of popularity is reflected not only in academia and practical settings, but also among the analysts in our sample. A Copenhagen Business School library search[2] reveals almost twice as many DCF valuations than any other method combined. Due to the popularity in use, and the sound academic research governing them, we will focus

[1] Including, but not limited to Discounted Cash Flow (DCF), Dividend Discount Model (DDM) & Economic Value Added Model (EVA)
[2] Search performed on 3/4/2021.

our project and analysis on the present value models, more specifically the DCF model and the inputs yielding firm value.

Levers of firm value in Discounted Cash Flow models

The components in the DCF are visualized below in Figure 1: Levers of firm value in DCF. Own contribution inspired by Banghøj et al. (2009)., which will serve as a guideline in this section:

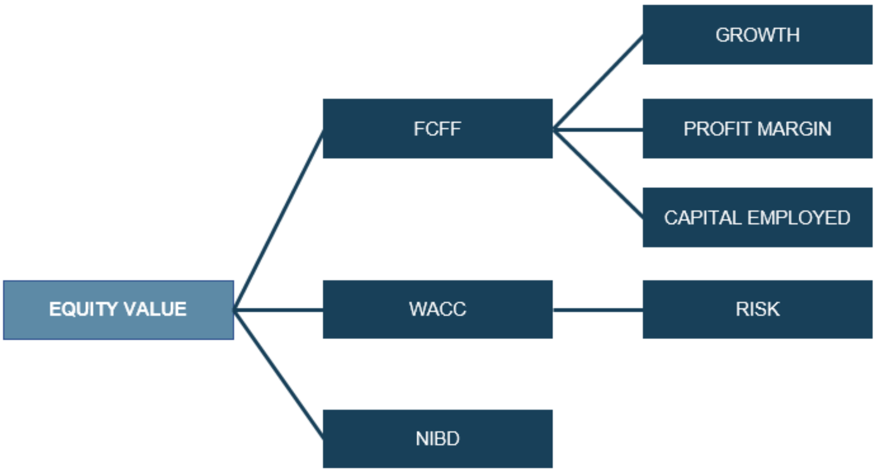


Figure 1: Levers of firm value in DCF. Own contribution inspired by Banghøj et al. (2009).

As described in the previous section, the DCF model is in some studies regareded as the most popular valuation model, which is also the case in our dataset, as 297 out of 321 students apply the DCF model in their valuation. We will therefore take a closer look at the DCF in the following section. Petersen & Plenborg (2008) describe the DCF as a two-stage present value model with the following formula:

Equation 1

$$EV = \sum_{\tau=1}^{T} \frac{FCFF_{t+\tau}}{\prod_{j=1}^{\tau}(1 + WACC_{t+j})} + \frac{FCFF_{T+1}}{WACC_{T+1} - g} * \frac{1}{\prod_{j=1}^{T}(1 + WACC_{t+j})}$$

Where:

$EV = Enterprise\ Value$
$FCFF = Free\ Cash\ Flow\ to\ Firm$
$WACC = Weighted\ Average\ Cost\ of\ Capital$
$g = Terminal\ Growth\ Rate$

The firm value, or enterprise value, in the DCF model is determined by the free cash flow to the firm discounted by the WACC. The model consists of two periods; an explicit forecast period and a terminal (continuing) period, where the growth in FCFF is assumed to be constant. In order to arrive at the equity value, it is necessary to subtract the market value of the net interest-bearing debt. It is assumed that any cash surplus is reinvested in the company or paid out as dividends and invested in projects with a net present value of zero, which would be equivalent to a return equal to the WACC.

g is the constant growth rate of the FCFF in the terminal period, and is therefore the growth rate which the analyst believes the company will grow at in perpetuity. Damodaran (2006) notes that companies can maintain high growth rates in extended periods, but are less likely to sustain this in perpetuity, as new entrants will enter industries of high growth, eventually leveling the playing field for existing actors. Companies will therefore, eventually, reach a stable growth rate which cannot be higher than the industry in which the company operates. This is the intuittion behind the two-state model; the first period is characterized by growth which is different from the market. As it reaches the terminal period, the growth will normalize into a level equal to the rest of the company's industry, and thus the company will enter steady state. For this reason, the assumptions behind the terminal value growth rate (g) is highly important – as this is the level of growth that will, theoretically, be sustained forever.

Another key component of the DCF is the WACC, which is the weighted average cost of the debt and equity employed by the company. Petersen, Plenborg, & Kinserdal (2017) define the WACC as:

Equation 2

$$WACC = \frac{Market\ value\ of\ equity}{Enterprise\ value} * r_d * (1 - t) + \frac{Market\ value\ of\ NIBL}{Enterprise\ value} * r_e$$

Where

$NIBL = Net\ interest\ bearing\ liabilities$
$Enterprise\ value = NIBL + Market\ value\ of\ equity$
$r_d = Required\ rate\ of\ return\ on\ NIBL$
$r_e = Required\ rate\ of\ return\ on\ equity$
$t = Corporate\ tax\ rate$

WACC is therefore the weighted average required return to lenders and shareholders. The tax rate is included in order to include the tax benefit associated with interest rate payments. Further assumptions regarding the WACC will be examined in section 2.4.

Decomposition of FCFF

The FCFF is the free cash flow to the firm, e.g. the cash flow available to shareholders and lenders after tax. The formula for FCFF from Petersen, Plenborg, & Kinserdal (2017) is as seen in Table 2 to the left. From our exploratory research of student valuations, the most frequent definition used in these valuations, is the one seen to the right:

Table 2: Different FCFF formulas (Petersen, Plenborg, & Kinserdal, 2017)

Operating income (EBIT)	**Operating income (EBIT)**
+/- Adjustment for items in EBIT with no cash flow effect	+/- Adjustment for non recurring items in EBIT
+/- Change in net working capital	***= EBIT, adjusted***
+/- Corporate tax	- EBIT, adjusted * (effective tax rate)
= Cash flow from operations	***= Operating income after tax (NOPAT)***
+/- Investments in non-current assets, net	+/- Change in operating assets
= free cash flow to the firm, FCFF	+/- Change in operating liabilities
	= Free cash flow to the firm, FCFF

The definition to the right includes changes in operating assets and change in operating liabilities, which is algebraically equal to the change in invested capital, as evident from equation 3.

Equation 3

$$Invested\ Capital = Operating\ assets - Operating\ liaibilities = Equity + NIBL$$

As evident, FCFF can be drived from two measures, namely FCFF = NOPAT + ΔInvested Capital. In this fashion, Banghøj et al. (2009) divides the drivers behind FCFF into three categories: top-line growth, profit margin and asset turnover, as evident from Figure 1. This approach bears some resemblance to the DuPont model, which decomposes ROIC into two multiplicative parts; profit margin and net operating asset (invested capital) turnover, as evident from equation 4.

Equation 4

$$ROIC = \frac{NOPAT}{Revenue} * \frac{Revenue}{Invested\ capital}$$

Soliman (2008) shows that the DuPont decomposition offers positive incremental information regarding the operating characteristics of a company, than solely looking at ROIC. These results are backed by Fairfield & Lombardi (2001), who show that disintegrating the change in ROIC into the change in asset turnover and the change in profit margin has incremental predictive power in relation to one-year-ahead ROIC, compared to change in ROIC alone. It is shown that an increase in asset turnover represents an increase in asset utilization, which has predictive power in relation to increases in furture earnings. Further, Cheng, Chu, & Ohlson, (2020) found that analysts' genereally are better at accurately forecasting revenue in comparison to margin forecasts. Earnings forecast errors (e.g. EBIT, NOPAT, net income) are therefore argued to be more a product of inaccurate margin forecasts than revenue forecast errors. This finding implies that decomposing FCFF into a revenue growth part, and a profit margin part would add incremental information into our analysis of forecast accuracy. Further, the decomposition of FCFF would allow us to, in a more nuanced fashion, test the forecasting behavior laid out by the analysts.

It should be noted that ROIC and FCFF are two different measures. The first being a ratio that seeks to measure how well a company uses its capital to generate profit, while the latter is a cash flow measure. They do, however, share similiarites. From Table 2 it is seen that:

Equation 5

$$FCFF = NOPAT + \Delta \text{Invested Capital}$$

While DuPont analysis consists of relative figures, NOPAT-margin and asset turnover, FCFF consists of absolute figures, but the same figures are included; NOPAT and invested capital, which are both dependent on revenue. Having the above findings in mind – we find it likely that decomposing FCFF into three parts can give us a better understanding of analyst forecast accuracy than analyzing FCFF on a stand-alone basis.

To summarize, we will, in our disintegration of FCFF, focus on the growth in revenue, NOPAT-margin and the turnover rate of invested capital, as this, according to our investigation of the valuations, is the most common measures used to arrive at FCFF. In Petersen, Plenborg, & Kinserdal (2017) WACC is calculated as an after-tax measure. Therefore, using NOPAT, which is also an after-tax measure, is chosen to be the most reasonable approach. Our calculation of ROIC will therefore also be calculated as NOPAT / invested capital, and thus be after-tax ROIC. This has the advantage that ROIC and WACC will become comparable. Based on the findings in Bonini et al. (2010), that share price forecast accuracy decreases with the predicted growth in the share price, we believe that our decomposition of FCFF has the potential to uncover the underlying drivers behind aggressive or conservative target prices.

The information environment and risk

The relationship between target price forecast accuracy and stock specific risk has also been investigated by previous literature. For instance, to study this relationship, Kerl (2011) uses the volatility of the stock and its price-to-book[3] ratio as a proxy for the riskiness of the stock and finds a negative correlation between the magnitude of these factors and the target price

[3] Market value of equity divided with book value of equity

forecast error. That is, higher risk (i.e., higher price-to-book ratio or higher historic volatility), leads to lower forecast accuracy. Ikromov & Yavas (2012) conduct an experimental study where they examine the effect of volatile dividend payments on the share price. With the same expected value of dividends in all trials, they find that volatility decreases share prices. Intuitively, these findings are expected, as an increase in volatility will increase the beta value in the CAPM model, and thus increase the required return on equity. This will drive down the share price, in order to compensate investors for the increase in risk. If we use FCFF as a proxy for dividend payments, and apply the findings in Kerl (2011), this could imply that volatility in FCFF would affect target price forecast error aswell. As we have already covered the drivers behind FCFF, we will examine the effect of volatility in historical revenue, NOPAT-margin and turnover of invested capital and its effect on the target price forecast error. The notion of volatility affecting forecast errors is consistent with the findings of Das et al. (1998), where they find that analysts publish more optimistic forecasts for companies that have historically been difficult to forecast. Further, they find that variability in historical earnings is positively correlated with how difficult it is to predict furture earnings.

The relationship between the market cap and forecast accuracy is also described in the litterature. As previously mentioned in Lim (2001), market cap is used as a proxy for the richness of the information environment, thus assuming that the information available is posetively correlated with market cap. More information should therefore lead to more accurate share price forecasts, which is also consistent with the findings in Falkenstein (1996). He finds that prediction erros regarding shareprice is inversely related with market factors, such as market cap. Cheng, Chu, & Ohlson (2020) derives to the same finding, concluding that sales and margin forecast accuracy improves with increased market capitalization. However, Bonini et al. (2010) proposes the opposite finding, concluding that share price forecast errors are postively correlated with the size of the company. The litterature is therefore torn on this subject. Nonetheless, we will seek to test whether the findings of Falkenstein (1996) hold true on both forecast accuracy and bias variables by analyzing the relationship between market capitalization and target price forecast accuracy and bias.

Also analyst coverage has been found to influence share prices. Doukas et al. (2005) found that the relationship between analyst coverage is positively correlated with stock overvaluation and low future returns. They argue that analysts are a part in a two-principal-agent problem between corporate managers and investors. Analyst face pressure from bankers to withhold negative information, which leads analysts to publish earnings forecasts with positive bias. This raises investor optimism, and causes stock prices to soar, thus causing them to trade above their fundamental value. Conversely, weak coverage causes investors to believe that the stock is more prone to being influenced by information asymmetry and agency problems, which causes shares to trade below the fundamental value. Chan et al. (2003) also find that analyst earnings forecast are affected by their desire to win investment banking business. We therefore hypothesize that high P/E ratios, can be a proxy for analyst coverage, as shares trading above fundamental values are theoretically more likely to trade on high P/E ratios.

2.4 Methodological variables:

Petersen & Plenborg (2008) list several assumptions that need to be satisfied in order to properly implement and use present value models. Present value models like the DCF, dividend discount models, economic value added etc. are theoretical equivalent, and should yield the same results, given the correct input and identical assumptions (Petersen & Plenborg, 2008). Not satisfying these assumptions can lead to faulty and incorrect valuations (Green et al., 2016). In the following section, some of the common pitfalls regarding the input of these models, as suggested by literature, will be presented,

WACC and coherent financial statements

When calculating the cost of capital (WACC), the capital structure must be based on market values and the cost of capital must reflect changes in the capital structure (Petersen & Plenborg, 2008). This leads to a circulating argument. To find the true capital structure, the market value of equity is needed, and in order to derive the market value of equity, the proper WACC is needed (and thus the capital structure). Larkin (2011) proposes a solution to this by proposing an iterative process. To apply the iterative method, you start out by using the weights from the current market value of equity. The value of the debt is assumed to be constant. Next it is checked whether the value of equity is equal to the equity value that was used to calculate the original weights of debt and equity. If this is not the case, you repeat the valuation using the new value of equity, in order to calculate new weights to use in the second attempt. This process is continued until the value of equity does not change anymore. A similar iterative process is described in Petersen, Plenborg, & Kinserdal (2017), where it is recommended that analysts rely on the capital structure of peers as well as the iterative process when determining the WACC. It is not our perception that either of these methods have been used to wider extends in our sample, nor that dynamic WACC levels are applied during forecast periods in our sample, as only 20 out of 297 DCF valuations contain changes in WACC during forecast. With this in mind, we will try to examine the relationship between the share price forecast accuracy, and the use of different values for WACC in the forecast.

The second assumption is that the cash flows in the explicit forecast period must be based on coherent pro-forma financial statements (Petersen & Plenborg, 2008). For instance, it is not allowed to create a “plug” in order to make the assets and liabilities balance, and Lundholm & O’Keefe (2001) argues that forecasts should follow the clean surplus criteria, namely that forecasted equity values should be explained when reconciling the net income and net dividends. Palepu et al. (1996) find that excess cash must also be reinvested in projects with a net present value of zero (return equal to the WACC), which has the consequence that any excess cash generated in the explicit forecast period cannot destroy nor create value. As it is inherently difficult to uncover whether the analysts of our sample have utilized balancing "plugs" or satisfied clean surplus criteria without complete access to their spreadsheets, we will seek to use the deviations of historical financials to forecasted financials to measure the coherence and realism of pro-forma forecasts.

The terminal period

As mentioned previously, the assumptions and inputs relating to the terminal value are of utmost importance in valuation, as these cement the expectations for the company in perpetuity. One assumption regarding terminal value inputs is that the cash flows must growth at a constant rate in the terminal period together with a constant capital structure. If Gordon's growth model is used, Levin & Olsson (2000) show that it is necessary to forecast the income statement, balance sheet and cash flow statement two years into the terminal period, to ensure that all variables grow at the same rate for two consecutive years, and thus satisfies the steady state assumption. This is quite an important point as 291 out of 297 valuations in our dataset have utilized Gordon's growth in their DCF.

Another assumption regarding the terminal period, is the relationship between ROIC and WACC, and the terminal growth rate and the risk-free rate. Damodaran (2002) has noted that if ROIC ≠ WACC in the continuing period, you will either be creating excess value forever if ROIC > WACC or destroying value forever if ROIC < WACC. Something similar goes for the growth rate in the terminal period. In the long run a company cannot grow faster than the economy in which the company operates. A good proxy for this is the risk-free rate, often in the shape of a government bond (Damodaran, 2002). If the terminal growth rate is higher than the risk-free rate, you assume that the company will be able to grow faster than the economy in all eternity. Ultimately, this has the unrealistic implication that the company eventually will become bigger than the total economy – something Damodaran coins "obeying the growth cap".

The relationship between the value in the terminal period and the explicit forecast period is further investigated in Cassia et al. (2007), where it is shown that the value of the terminal period is associated with the length of the explicit forecast period. The longer the explicit forecast period is, the smaller proportion of the total value will be attributed to the terminal period. The DCF model is often critiqued for the high proportion of value created in the terminal period (Platt, Demirkan, & Platt, 2009; Green et al., 2016). Further, the general decline of interest rates, and therefore discount rates, may magnify the relevance of this critique, as, all else equal, a decrease in discount rates will inflate the value of the terminal period relative to the total enterprise value. This could hypothetically imply that having a short explicit forecast period is negatively correlated with accuracy, as the proportion of value created in the terminal period could ultimately become too high, which cements the need for additional consideration when deriving terminal value assumptions. On the other hand, a longer forecast period also places new demands to analysts and their forecasting abilities, as it is inherently more difficult to forecast for longer periods than shorter periods. This might also pose a challenge to the achieved share price forecast accuracy. As noted above, the explicit forecast period also needs to be long enough to satisfy the steady state assumption, which might not be the case if the period is too short. In our analysis, we will try to uncover the relationship between the explicit forecast period in order to examine its effect on the share price forecast accuracy.

Empirical findings

The importance of the abovementioned assumptions is highlighted in Green et al. (2016), where 120 DCF-models from U.S. brokers are analyzed, showing that analysts make a median of three theory related and/or execution errors per valuation. After correcting for these errors, the mean valuations and target prices change between -2% to 14% per error corrected. Out of the 120 DCF valuations analyzed, 7% set the terminal growth rate above 5%, and 84% of the analysts set the terminal growth rate ±30 bps away from the relevant 10-year U.S. Treasury note. 24% of the analyst have a revenue growth rate in the terminal year that is either higher than 10% or at least two times the growth rate in the explicit forecast period. These findings do not align well with the arguments from Damodaran (2002), who proposes to use the risk-free rate (e.g., a government bond) as a proxy for the growth rate in the terminal period. In our analysis we will examine whether deviating from using the risk-free rate, as a proxy for the growth rate in the terminal period, will affect the share price forecast accuracy.

When calculating WACC it is shown that 30% of the analyst use an equity value weight which is ±10% away from the weight implied by the ratio between the enterprise value and equity value implied from the valuation. 14% of the analyst also fail to adjust the WACC over time when the weight of the equity value is ±20% away from the equity weight initially found by the analyst. These findings deviate from the previously discussed assumptions regarding the WACC calculation, namely that it should be based on market values and change if the capital structure changes (Petersen & Plenborg, 2008; Larkin, 2011). As this might lead to inaccurate or biased estimates, we will seek to uncover whether the analysts in our sample are prone to make the same methodological errors, and whether these errors can help explain the differences in target price accuracies and forecast errors.

Green et al. (2016) also found that 14% of analysts set the terminal year four years or less from the valuation date. In addition to this, in 22% of the analyzed DCF models the continuing period accounts for at least 85% of the total value. These findings tie in well with the criticism raised by Platt, Demirkan, & Platt (2009) regarding the high levels of value created in the terminal period. We will, on the background of the above findings, analyze the relationship between the length of the forecast period and its relationship with the target price forecast accuracy.

Valuation model choice

The accuracy of the DCF compared to other valuation models is examined in Sayed (2015). Sayed conducts a study of sell side analysts in emerging markets and finds that the accuracy of the DCF model is generally more accurate than multiples and return on capital models. He finds that the DCF valuation have an average TPA of around 70%, while book asset-based models perform the worst with a TPA of around 50%. Multiples are the second most accurate with a TPA of around 60%. Most importantly he finds that the use of earnings multiples together with a cash flow-based model can increase share price forecast accuracy. Imam et al. (2013) also finds that a cash flow based valuation model together with multiples can improve forecast accuracy, but on the other hand, find that accrual based multiples outperfom cash flow based models. Nevertheless, Bonini & Kerl (2012) find that valuations that deviate from simple multiple based target prices are more accurate, thus indicating that more sophisticated models like DCF's and residual income models, that demand more sophisticad inputs, have better accuracy. This is in line with Gleason, Johnson & Li (2011) who find that analysts that formulate earnings forecast, instead of merely using a heuristic multiple approach, produce more accurate target prices on a 12 month horizon. Asquith et al. (2002) on the other hand finds no relationship between the valuation model choice and share price forecast accuracy. It is certain that the litterature regarding forecast accuracy of different valuation models, is at best unclear. The different results in the studies can probably be attributed to different research methodologies (e.g. different measures of forecast accuracy), the sample used, the behavior of analysts across regions and the companies used in the studies might also be causing the discrepancy. Nonetheless, we will try uncover the forecast accuracy of different valuation models in our dataset. We will mainly focus on the target price forecast accuracy of multiples together with present value models versus only using a present value models. We will not be looking at the sole use of multiples as our sample does not contain enough such observations to substansiate an analysis of this kind. We will, however, be looking at the forecast accuracy of different multiples, but these are in most instances used together with a present value model of some sort, and therefore cannot be reviewed on a stand-alone basis. 198 have used multiples together with af present value model, while the remaining 123 valutaion reports have only used a present value model.

3.0 Research questions

Based on the motivation outlined in the introduction, this thesis seeks to uncover whether analysts target price estimates are in fact accurate, or if they exhibit bias in their recommendations. The primary research question is therefore as follows:

RQ1: "Are target prices derived from student valuations accurate and unbiased?"

To answer this question, two sets of hypotheses were formulated; the first seeks to test analysts' overall accuracy, and the second aims to uncover whether accuracy and bias measures differ between buy and sell recommendations. The hypotheses are as follows:

$H1_0$: Analyst target prices are accurate and unbiased

$H1_a$: Analyst target prices are inaccurate and biased

$H2_0$: Accuracy on buy recommendations = Accuracy on sell recommendations

$H2_a$: Accuracy on buy recommendations ≠ Accuracy on sell recommendations

As previous literature has, almost unanimously, ascertained that bias exists, we also expect to find bias in our sample. Therefore, we seek to test whether there exist distinct factors that can explain the accuracy or bias within these target price estimates. Our secondary research question is therefore as follows:

RQ2: "Which factors influence the accuracy or bias derived from student valuations?"

Naturally, a multitude of factors could potentially explain bias or inaccuracies, as also uncovered in the literature review. Therefore, based on the literature review findings, a list of hypotheses has been formulated, grouped by three distinct themes, namely: analyst characteristics, financial factors, and methodological factors. What follows are the hypotheses that will be tested to answer our secondary research question, based on these theorized explanatory variables.

3.1 Analyst characteristics

Previous research has suggested that analyst accuracy differs between men and women. Hope & Fang (2020) found the gender variable to be significantly explanatory on differences in forecast accuracy. Bosquet et al. (2014) found women to be more conservative than men, and Green et al. (2009) found forecast errors to be slightly larger for female analysts than for their male counterparts. Therefore, we will seek to test whether gender plays a role in the target price accuracy of our sample. The hypothesis is therefore:

$H3_0$: Male analysts' accuracy = Female analysts' accuracy

$H3_a$: Male analysts' accuracy ≠ Female analysts' accuracy

Hope & Fang (2020) also found explanatory power in the experience and educational backgrounds of analysts. As our sample includes the study programme of analysts, we will seek to test whether there is a difference in analysts' accuracy between study programmes. The hypothesis is therefore stated as follows:

$H4_0$: The study programme of the analyst does not influence the analysts' accuracy

$H4_a$: The study programme of the analyst does influence the analysts' accuracy

The last variable of interest in our analyst characteristics group is that of analyst teams. Hope & Fang (2020) also showed that analyst teams, on average, proposed more accurate target prices. As our sample contains valuations by either one or two analysts, we will seek to test the following hypothesis:

$H5_0$: Analyst groups' accuracy = individual analysts' accuracy

$H5_a$: Analyst groups' accuracy ≠ individual analysts' accuracy

3.2 Financial variables

Previous research has found varying associations between the forecasts of financial variables and the target prices issued by analysts. In this section, hypotheses will be formulated and presented based on the levers of FCFF presented in section 2.3.

Revenue is one of the underlying drivers in our decomposition of FCFF, as revenue bears impact on all variables further down the income statement. We will therefore test the deviation between the growth in the historical period versus the forecast period. In this way, we seek to find whether aggressive or conservative forecasts lead to better or worse target price forecast accuracy.

We will also be looking at the historical volatility in revenue in order to uncover if historical volatility leads to a decrease in target price forecast error. Intuitively, historical volatility should make it harder to forecast, as trends are harder to identify. In our review of the literature, we found evidence that volatility makes it more difficult for analysts to perform accurate forecasts – but also that increased volatility made forecasts more optimistic. We there hypothesize that:

$H6_0$: Historic revenue volatility does not influence analysts' accuracy

$H6_a$: Historic revenue volatility does influence analysts' accuracy

$H7_0$: The deviation of forecasted revenue growth to historical growth does not influence analysts' accuracy

$H7_a$: The deviation of forecasted revenue growth to historical growth does influence analysts' accuracy

The next driver in our decomposition of FCFF is the NOPAT-margin. Again, we seek to find the relationship between historical volatility in NOPAT and target price forecast accuracy. We will also examine the relationship between conservative and aggressive forecasts and target price forecast accuracy. Our hypotheses are therefore as follows:

$H8_0$: Historic NOPAT-margin volatility does not influence analysts' accuracy

$H8_a$: Historic NOPAT-margin volatility does influence analysts' accuracy

$H9_0$: The deviation of forecasted NOPAT-margin to historical NOPAT-margin does not influence analysts' accuracy

$H9_a$: The deviation of forecasted NOPAT-margin to historical NOPAT-margin does influence analysts' accuracy

The final driver in our decomposition of FCFF is the asset turnover, where we will examine the revenue to invested capital ratio and its relationship with the target price forecast accuracy. As with revenue and NOPAT-margin, we will examine the relationship between conservative / aggressive forecasts and its relationship with our accuracy variables. We will also examine the relationship between historical asset turnover volatility and target price forecast accuracy. To do this, the following hypotheses were formulated:

$H10_0$: Historic ATO volatility does not influence analysts' accuracy

$H10_a$: Historic ATO volatility does influence analysts' accuracy

$H11_0$: The deviation of forecasted ATO to historical ATO does not influence analysts' accuracy

$H11_a$: The deviation of forecasted ATO to historical ATO does influence analysts' accuracy

From our literature review, we found evidence that firm size is associated with forecast errors, as market capitalization was used as a proxy for the richness of the information environment. The literature is however divided on whether market capitalization positively influences forecast accuracy. The discrepancy in the literature might be caused by difficulty isolating a single variable reflecting the availability of information. The analysts' access to information is not limited to the official channels of information (e.g., company reports and announcements), but also private information (Sayed, 2015). As we believe students' access to private information is limited (e.g., more uniformly distributed), we believe our sample can lead to more unbiased results on this variable. To test the impact of market capitalization, we have therefore formulated the following hypotheses:

$H12_0$: The company's market capitalization at T_0 does not influence analysts' accuracy

$H12_a$: The company's market capitalization at T_0 does influence analysts' accuracy

3.3 Methodological variables

As evident from section 2.4, several methodological choices have influence on the valuation of equities. In this section, we will state the hypotheses derived from our literature review.

The literature regarding the relationship between valuation model choice and target price forecast accuracy was found to be unclear. This was attributed to differences in sample and research methodology across the studies. In our sample, 198 have used multiples together with af present value model, while the remaining 123 valuation reports have only used a present value model. We can therefore not analyse the accuracy of only using multiples (as opposed to present value models), due to limitations in our dataset. In this specific regard, the litterature was quite consistent, as two studies in our literature review found present value models to increase forecast accuracy (Imam et al., 2013; Sayed, 2015). However, we will test if the number of mulitples used affects target price forecast accuracy, and whether additional present value models apart from the DCF model influences the target price forecast accuracy.

> *$H13_0$: The amount of multiples used in valuation does not influence accuracy*
>
> *$H13_a$: The amount of multiples used in valuation does influence accuracy*
>
> *$H14_0$: Using additional valuation models does not influence accuracy*
>
> *$H14_a$: Using additional valuation models does influence accuracy*

As shown in section 2.4, the influence of the length of forecast horizon is often debated. We will examine how the length of the forecast period affects the target price forecast accuracy. A longer forecast horizon is expected to make it increasingly difficult to forecast, thus potentially increasing the target price forecast error. However, a longer forecast horizon will also decrease the value of the terminal period relative to the enterprise value. Letting this ratio become too large (i.e., terminal value makes up a larger part of total enterprise value) is unfavorable according to Platt, Demirkan, & Platt (2009) and Green et al. (2016). Thus, literature is rather undecided on the optimal forecast period length. We will therefore seek to test the following:

> *$H15_0$: The length of forecast period does not influence accuracy*
>
> *$H15_a$: The length of forecast period does influence accuracy*

Apart from the forecast period length, we will also examine the effect of the historical period length in each valuation to see whether longer periods provide more accurate forecasting abilities. Intuitively, a longer historical period would allow the analyst to perform the valuation on a more informed basis and should therefore lead to a more accurate target price forecast. The hypothesis is therefore stated as follows:

H16$_0$: The length of historical period does not influence accuracy

H16$_a$: The length of historical period does influence accuracy

As previously argued from equation 2 in section 2.3, the DCF model is a two-stage model, with an explicit forecast period and a final, terminal value period. The model is divided into two stages, as to accompany excess growth for a period before reaching steady state. The steady state assumption requires all parameters to grow at the same rate in the terminal period, with constant capital structure. From literature outlined in section 2.3, it is necessary to forecast the income statement, balance sheet and cash flow statement two years into the terminal period in order to ensure that the parameters grow with the same constant rate if Gordon's growth model is used in the terminal period. Therefore, we will examine the whether the fulfillment of this assumption has any influence on the target price forecast accuracy. The hypothesis is therefore as follows:

H17$_0$: Satisfying steady state assumptions does not influences accuracy

H17$_a$: Satisfying steady state assumptions influences accuracy

According to Damodaran (2002), growth in the terminal period cannot exceed the growth in the economy in which the company operates. If this was the case, the company would eventually become bigger than the economy itself. It is suggested that a likely approximation for maximum level of perpetuity growth is the risk-free rate. We will therefore compare the risk-free rate used in the calculation of the WACC, to the growth rate in the terminal period, as to see whether this relationship has any impact on analysts' ability to derive accurate target price estimates. Our hypothesis is as follows:

H18$_0$: Satisfying the growth sanity check does not influences accuracy

H18$_a$: Satisfying the growth sanity check does influences accuracy

Another assumption listed in Petersen & Plenborg (2008) is that changes in capital structure should be reflected in the WACC. We will therefore test how the fulfillment of this assumption affects the target price forecast accuracy. As evident from Appendix A, the leverage level in the forecast period of our sample is very rarely constant, and thus, our approach is simple: if WACC changes in the forecast period, we assume this assumption to be satisfied, and vice-versa. Our hypothesis is therefore:

H19$_0$: Changing WACC during forecast does not influences accuracy

H19$_a$: Changing WACC during forecast does influences accuracy

The last explanatory variable of our analysis is based on another assumption mentioned in Damodaran (2002), namely the relationship bewteen WACC and ROIC in the terminal period. If ROIC is larger than WACC in the terminal period, it is implied that the company will create excess value forever, thus keeping its competitive advantage in all eternity. As noted by Damodaran this is an unrealistic scenario, and we therefore seek to examine the relationship between WACC and ROIC in the terminal year and its influence on forecast accuracy. To do this, the following hypothesis will be tested:

> *$H20_0$: The relationship between WACC and ROIC in terminal year does not influence accuracy*
>
> *$H20_a$: The relationship between WACC and ROIC in terminal year does influence accuracy*

4.0 Methodology

Having introduced the literature on analysts forecast accuracy, the methodology from which we will base our analysis will be explained. In section 4.1, the general research design will be elaborated upon. From here, we will explain the data collection process in section 4.2, split between primary and secondary data sources. The reliability of our primary data source will be covered in section 4.3. The calculation of dependent and independent variables will be covered in section 4.4, after which the process of outlier removal will be elaborated in section 4.5. The statistical methodology will briefly be explained in section 4.6, and finally, the chapter will be rounded off with a brief discussion of the methodological limitations of the thesis in section 4.7.

4.1 Research design and considerations

In order to study the target price accuracy of analysts, including possible explanatory factors, a quantitative research design was chosen. We will be utilizing both a correlational research design to derive impacts of certain variables on the accuracy of analysts, but also an explanatory design, as we rely on prior research on the field. Thus, we work deductively in our thesis, meaning that our variables of interest are largely derived from existing theories that we seek to test.

As the data collection process is rather extensive, it was decided to work together with two other groups on data collection and general methodology, as this would increase the group resources and thus allow for a larger dataset. Further, the inclusion of more people, and thus more coders, would, all else equal, bring down the impact of potential biases from individual perceptions – but likewise a larger potential for inconsistencies. Therefore, as recommended by O'Connor & Joffe (2020), the formulation of the coding scheme and codebook consisted of an iterative process, including mutual discussions of discrepancies and several revisions and additions to the coding scheme. The process is illustrated in Figure 2 below and explained in further detail in the next section.

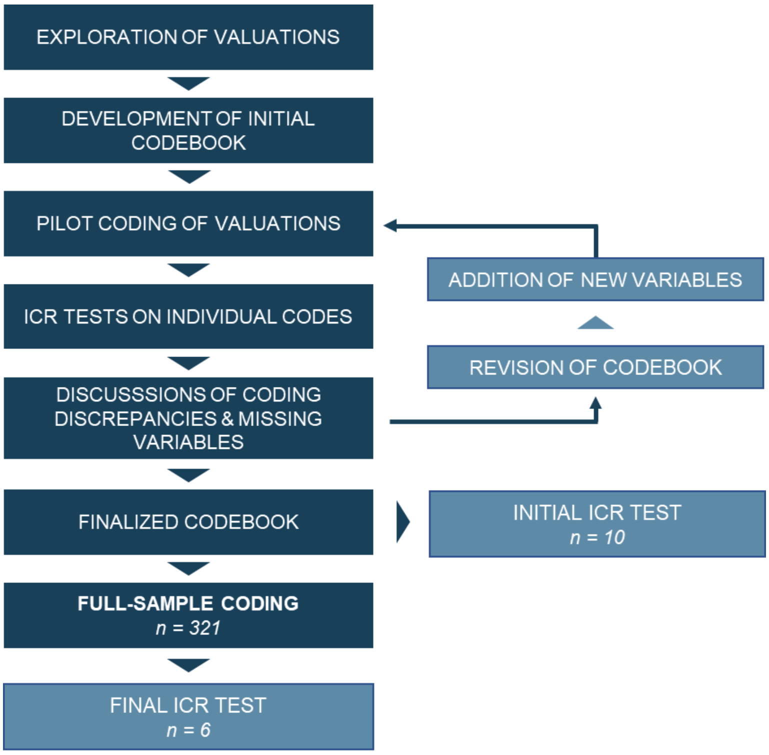


Figure 2: Research design process. Own contribution adapted from O'Connor & Joffe (2020)

As all six members of the data collection group are trained in valuation techniques and finance-related topics, we decided to work deductively with our research design, thus formulating a preliminary codebook prior to coding the valuations, with a predefined coding scheme. This initial coding scheme was developed after exploring several randomly selected valuations from the database, in order to, first and foremost, get an understanding of the format and contents of student's valuations, and secondly, to determine which variables were reasonable to assume could be populated at high frequency during coding.

Working deductively, we initiated our research design by a hypothesis that analyst biases could be explained by the input and assumptions of their valuations – and those are the correlations that we seek to uncover. Using archival data, we therefore seek to systematically draw conclusions from underlying assumptions behind valuations - thus, our research findings are limited by the variables included in our coding scheme. It is therefore of utmost importance that our coding scheme is both well thought-out and extensive enough to allow for potential insights to be uncovered, therefore speaking to the importance of an iterative process of codebook design.

4.2 Data collection

To analyze analysts' target price forecast accuracy in listed company valuations, archival valuation data needed to be collected. Initially, we sought to gather data on analyst valuations from danish investment banks, including insights into their use of methods, forecasting assumptions and underlying calculations. However, this proved to be out of reach, as there seemed to be no systematic archives nor accessible storage of these valuations, and thus we had to consider alternative solutions. One of which was to use students' theses as proxies to analysts, as one common topic for theses at Copenhagen Business School (CBS) is valuations of listed companies.

Through an initial search at the CBS database for students' theses containing the abstract keywords; "valuation of" and/or "værdiansættelse af", gave 856 search results at the time of writing, thus indicating a sizable sample. However, to manually search through the thesis database, archive and index all relevant valuations accordingly would prove to be a rather time intensive task. Therefore, a web-scraper was built to fetch all theses from the database, accompanied by the information provided about the thesis at the CBS website, namely, thesis title, page number, publication date, author names and the link to the PDF file. The scrape-flow is illustrated below:

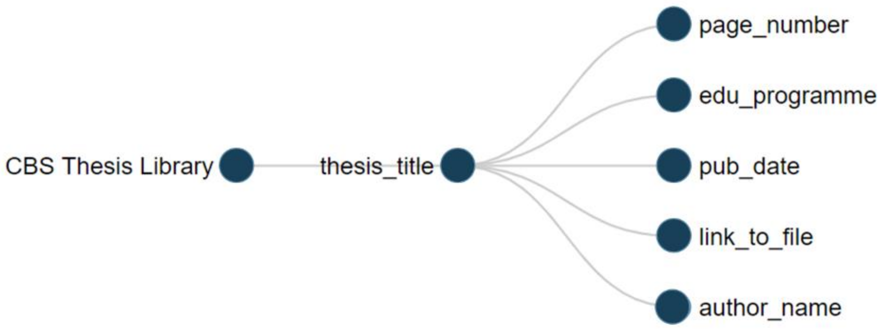


Figure 3: Flow of web-scraper at CBS Thesis Database. Own contribution

As evident, the scraper allowed us to index each thesis and gather valuable data on the aforementioned variables. From the edu_programme and author_name variables, we were able to distinguish between the number of authors and study programmes, as both were delimited by special characters in the database. Thus, we can later test whether analysts perform better given their gender and/or study programme, and thus test the findings of Hope & Fang (2020).

From the scrape file, the sample was narrowed down further. First, through an exclusion formula in excel, we removed all valuations including "LBO" in the title, as these are naturally difficult for us to analyze, as the companies in question are rarely listed at a stock exchange at the time of writing, and they are rarely valued using DCF-models in the student theses. All valuations including “options” in the title were then stripped from the sample, as these are not within the scope of our project. This initial cleaning of our dataset narrowed down our sample from n = 856 to n = 363.

As this exclusion process potentially excludes both feasible and relevant valuations, but also fails to exclude irrelevant projects, one final and manual cross-check was initiated, which yielded 85 irrelevant projects - projects about e.g., “Valuation of Danish Callable Mortgage Bonds”, which do include “valuation” in the title, but do not revolve valuation of publicly listed companies, or project where no appendices were available, and thus no data to code. Through a PowerShell script, all valuations were then downloaded in bulk, allowing for the formulation of the codebook to begin.

As the CBS thesis library only includes theses that are marked as non-confidential by the authors, some valuations were still submitted, but not published. Therefore, our supervisor, by help of the CBS Library, contacted the authors, asking for permission to use these in our project. This increased our sample by 43 new valuations, thus arriving at a total sample size of n = 321, consisting of more than 40.000 coded observations combined.

4.2.1 Primary data source

In this section, the variables of the codebook and the thought process behind them will be elaborated, as the codebook is our primary source of data. The raw codebook variables are attached in Appendix B.

After developing the initial codebook, we coded two pairs of five valuations each (same five valuations for each group member), with discussions and comparisons in-between. A couple of issues were raised and corrected, namely: how to normalize the calculation of total number of pages, as well as pages on strategic versus financial analysis. The period for historical data was also altered, as we initially used five years of historical data, but decided to extend it to eight years, as to gather as much data as possible on financial figures. Also, the forecasting period was extended to 15 years, as that seemed to be the upper limit from our pilot coding. Further, we discovered that our list of multiples was not sufficient, as we had predefined the list to include only the most used valuation multiples from the pilot coding such as EV/EBIT, P/E, EV/SALES and so forth – split between income statement and balance sheet multiples. However, different industries call for very specific industry multiples. Oil and gas exploration companies often used EV/MBOEPD[4], salmon harvesting companies used EV/KG, and shipping companies EV/Vessels. These distinctive valuation multiples were discussed, and we agreed to

[4] Enterprise Value / Thousand Barrels of Oil Equivalents Per Day

note them as "*Other*" in the main codebook, and keep a secondary list of unique multiples, as to not flood our main codebook with numerous multiple options that are rarely occurring.

The financial figures coded from the income statement consists of revenue, EBIT, NOPAT and net income, and assets, equity, net debt and invested capital (IC) from the balance sheet. Further, the free cash flow to the firm (FCFF) was coded, along with DCF valuation input such as WACC, terminal growth rate, risk-free rate and implied enterprise and equity values.

The revenue was set to include minorities insofar these were accounted for in the enterprise value calculations. If no nominal revenue numbers were forecasted, but only growth rates were noted, we calculated the revenue figures in nominal terms from the growth rates. In terms of EBIT, we sought to primarily code the EBIT from all activities, as this is what is usually reported in official financial statements. If EBIT from all activities was not forecasted, EBIT from sales was noted, and the EBIT numbers were marked as adjusted. The net income variable follows the same logic, where regular net income was of priority – however, in some cases, comprehensive income was forecasted instead, which was also coded, but noted as being adjusted.

NOPAT was also coded, which we found was more often used than EBIT, as the students in our sample preferred to report analytical income statements and balance sheets. A few valuations included NOPLAT rather than NOPAT, which is considering adjusted tax. In our analysis we have not distinguished between these two.

Moving to the balance sheet figures, more variety came into play. First and foremost, the total assets were rarely forecasted, as most students reformulate their balance sheets to analytical balance sheets focusing on net operating assets on the asset side and invested capital on the liabilities side (equivalents). Thus, the assets figure was, if not forecasted, calculated as operating assets + financial assets, or total non-current assets + total current assets + total interest-bearing assets, as a rough proxy.

The equity was in most instances forecasted, as it is of large importance for any forecast to balance, in terms of satisfying the clean surplus principle. Net debt was also often forecasted – however, in the instances that it was not, it was calculated from equity and invested capital – and vice-versa in cases when invested capital was not forecasted, but equity and net debt was.

Below we have visualized the variables and assumptions we will focus on in this thesis. Based on the findings in our literature review, we have divided our variables into financial variables, methodological factors and analyst specific characteristics as possible theorized explanatory variables for target price accuracy or bias. The first focuses on the budgeting behavior of the analyst, the second on the methodological choices and assumptions applied by the analyst, and the third on the characteristics of the analyst, namely gender, study programme and whether the valuation was made by a team of analysts or an individual analyst.

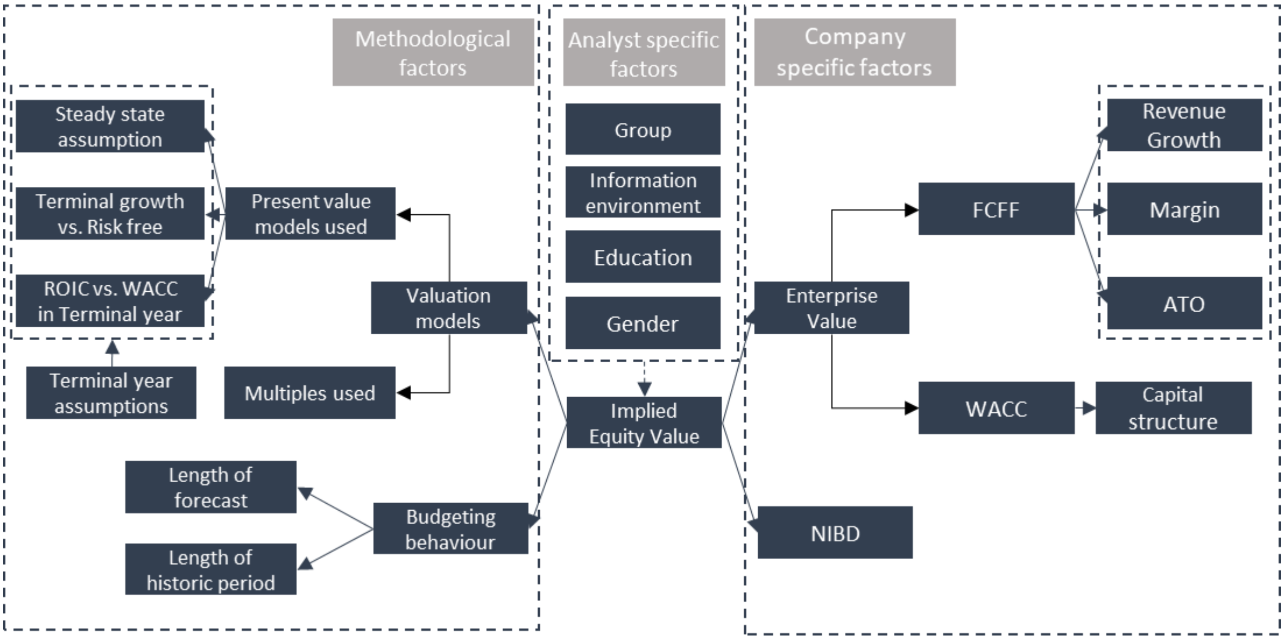


Figure 4: Overview of variables researched. Own Contribution.

The most common valuation model applied in the projects is by far the discounted cash flow to the firm model (DCFF) as it is applied by 297 out of 321 analysts. We have therefore chosen to focus on the inputs in this model, in order to examine how it affects the accuracy of the final share price found.

In our literature review it was found that by dividing FCFF into three different drivers, namely revenue growth, margin growth and asset turnover, increased predictive ability was found. As our earnings measure, we have chosen to focus on NOPAT, which is the operating profit less tax. As net asset turnover we will use revenue divided by invested capital (net operating assets). NOPAT divided by invested capital further gives us return on invested capital after tax (ROIC after tax), and such our decomposition of FCFF into NOPAT + Δ invested capital bears some resemblance to the DuPont decomposition of ROIC. Using ROIC after tax results in greater comparison with WACC rates, as both are after-tax measures. This will enable us to uncover the underlying value drivers in the DCFF and their effect on the forecast error and bias. As a measure of the capital employed in the company, we have chosen invested capital as our metric. This figure consists of equity plus net interest-bearing liabilities. In the analyzed projects, the book-value of equity and debt is usually used to calculate WACC in the valuations, and we will therefore use book values of net interest-bearing liabilities and equity in order to calculate invested capital and ROIC, in order to ensure better comparability with WACC.

The list of raw variables coded is attached in Appendix B, and the complete list of variables after computations is attached in Appendix C.

4.2.2 Secondary data source

In order to measure analysts' target price accuracy, naturally, we need data on the share price development of the underlying companies. In this section, a brief description of the data collection and variables fetched from Capital IQ will be explained.

CapitalIQ is a financial database hosted by S&P Global Market Intelligence, from which various financial data can be extracted. In order to extract share price data from Capital IQ, we need to get the corresponding ticker for each company. To do this we use a partially automatic search function built into the Excel plugin. As this is an automatic method, it is naturally prone to errors, and therefore, in order to mitigate this error, we also extract the corresponding company name from the Capital IQ ticker. If there is any discrepancy between the company name noted by the coder and the company name given by the Capital IQ ticker, we find the correct ticker manually.

Even though the company name extracted by Capital IQ corresponds with the company name noted by the coder, errors can still occur. Apart from needing the correct company ticker, the currency of the extracted data also needs to coincide with the currency of the codebook. Capital IQ also extracts data in millions, and it is therefore crucial that the coder has noted down the financials in this format as well. The same goes for the financial years, (e.g., first year in forecast and terminal year). Coder errors can lead to Capital IQ extracting financials in the wrong currency or the wrong financial years, thus a deviations sheet was computed to carefully investigation these errors, after which they could be corrected manually.

As a check of our data quality, we calculate the actual share price noted in the valuation divided with the target share price. We also calculate the market value of equity from Capital IQ at the implied share price date divided with the implied equity value. To put it simple: the actual equity value divided with the implied equity value should equal the actual share price divided with the target share price. These two values should be (approximately) equal given our data source is correct (under the assumption that everything noted down by the coder is also correct).

Equation 6

$$\frac{Actual\ equity\ value_{t=0}}{Implied\ equity\ value} = \frac{Actual\ share\ price_{t=0}}{Target\ share\ price}$$

If these two measures are not equal, we will need to find the source of the deviation (see above). If these on the other hand are equal, we check whether the actual share price$_{t=0}$ from Capital IQ is equal to the one noted in the valuation. If not, the source of the deviation usually stems from splits (or reverse splits). Because Capital IQ adjust for splits back in time, we need to adjust the target price and actual share price given in the valuation. Careful manual examination of the data has been done to minimize this error.

4.3 Intercoder reliability

As our dataset was created by multiple people, we risk inconsistencies in the interpretations of the data at hand. Therefore, a test for intercoder reliability (ICR) was made. In this section, a brief description of the ICR measure will be presented, along with the calculations and results of our reliability test.

ICR is a measure of evaluating the reliability of a multi-coder dataset, as it seeks to uncover whether each individual coder tends to reach the same conclusions when presented with identical data, and to what degree agreeability is met. In our case, ICR is highly important to secure reliability of our dataset, as we – although being trained at the same university, by the same lecturers, still may interpret certain variables in a different manner - or, may simply miss certain variables of importance in the given valuations.

A 2021 search for Intercoder Reliability on Google Scholar yields over 68.000 search results[5], indicating that ICR is widely used – yet the way in which ICR is computed and evaluated is up for large debate. The appearance of ICR is most frequent in qualitative studies (O'Connor & Joffe, 2020), where personal perception and worldview often are at risk of creating higher interference. Arguably, quantitative studies suffer less from the influence of personal perceptions – yet the correctness of our understanding still is of vast importance, and thus ICR was chosen as an appropriate measure of reliability in this study.

As mentioned, ICR can be computed and measured in a variety of ways. Popping (1988) found 39 different methods of measuring relative agreement, where the methods can be roughly divided into two groups:

- A simple agreement percentage (Total agreements / Total observations), indicated as the most common method by previous reviews of the literature (Feng, 2014; Kolbe & Burnett, 1991)
- Several statistical measures, seeking to quantify the risk of deriving the same answer by chance. These include, but not limited to; Cohen's Kappa, Krippendorff's alpha and Scott's pi (O'Connor & Joffe, 2020).

[5] Search performed 01/04-2021

As we are dealing with a codebook consisting of mostly numeric data (229 numeric variables and 35 binary variables), the risk of achieving agreement by chance is arguably limited. Therefore, the choice went on the simple percentage agreement approach, allowing us to rather easily determine whether specific parts of the valuation; certain variables, categories etc., were the sources of disagreements, without interference from statistical adjustments.

Some studies suggest that ICR should be computed on anywhere between 10-25% of all datapoints in a sample (Campbell, Quincy, Osserman, & Pedersen, 2013), however, this rule of thumb is pointed mostly at qualitative studies, specifically interviews, where variability in understanding may be larger. Due to the nature and the scope of this project, a 10% ICR sample would be a rather cumbersome process. The time needed to code 10-25% of the sample would be high, and the potential outcome would be rather low, thus we opted for a fixed number of 10 initial codes, with satisfying ICR, before embarking on the individual coding process of the full sample. To ensure that our quality of coding would still be acceptable after the individual coding process was started, we added in six duplicate valuations in everyone's codebook, allowing us to measure ICR on the on-going coding of the full sample. This was done to allow for both quality-assurance, but also to see whether we became better (or worse) at judging valuations based on the criteria set in our codebook.

Lombard, Snyder-Duch, & Bracken (2002) argue that, to improve ICR, the research design should follow an iterative process, with coder training, formal pilot test and on-going tests during full-sample coding. As previously mentioned, and as illustrated in Figure 2, our research design has been an iterative process, where the codebook has been changed and adjusted according to discussion and debates. Therefore, the initial ICR test (n = 10 valuations) was not performed until we had a final version of the codebook, in which we had specified each variable that had previously been the source of doubt or confusion. We chose to report the distribution of our ICR scores on a per-case basis, rather than as a pooled average or median measure, as that would not allow for monitoring potential outliers. Further, it allowed us to see whether certain types of valuations were more prone to disagreement, thus enabling us to discuss and correct any issues.

Intercoder reliability method

As mentioned, our approach to calculating ICR is by using the 'simple agreement percentage' method. However, few steps are applied before calculating the ICR score.

First and foremost, each valuation is given to each coder in the group, accompanied by the codebook with indexation and scraper-information (as described in section 4.2). After the coding is finalized, the codes are compiled in a single sheet, and a new ICR sheet is made. In the ICR sheet, a random baseline coding is picked, which is the code that the remaining five codes are tested up against. The observations in the baseline case are not counted in the total amount of variables, as the baseline case would, per definition, be 100% 'correct', thus falsely inflating the ICR score.

In the ICR sheet, each cell from code 2 to code 5 is tested to see whether they are equal to code 1 (baseline), returning a 0 if equal, and a 1 if unequal. For numeric values, e.g., total assets, a minor adjustment of 0,5 in nominal currency is allowed, as to circumvent disagreements caused by copy-pasting or decimal errors when importing the codebooks to the ICR sheet. The formula for numerical variables is thus as follows:

Equation 7

$$Agreement_{ij} = Baseline_{ij} - Code_{X_{ij}} \leq 0{,}5$$

Where i = row, j = column, x = given code from 1-5.

Further, as the codebook is made to accompany seven historical years and 15 forecasted years of financials, some years may not be reported in the valuation, thus causing blank cells. These blank cells are not counted as observations, nor included as agreements or disagreements, as that would likewise inflate the ICR score. However, if e.g., $Baseline_{ij}$ has numbers, but $Code_{X_{ij}}$ is blank (or vice-versa), both instances are counted as observations, and one disagreement is counted. An example of the ICR table is shown in Table 3, with red cells displaying disagreements, and green cells displaying agreements.

Table 3: Summary example of ICR test. Own creation

Variable	Baseline	C1	C2	C3	C4	C5	BL_vs_C1	BL_vs_C2	BL_vs_C3	BL_vs_C4	BL_vs_C5
EV - income statement (1/0)	1	1	1	1	1	1	0	0	0	0	0
Used with peer benchmark? (1/0)	1	1	1	0	1	0	0	0	1	0	1
EV/EBITDA	1	1	1	1	1	1	0	0	0	0	0
EV/EBIT	1	1	1	1	1	1	0	0	0	0	0
EV/Sales	0	0	0	0	0	0	0	0	0	0	0
EV/NOPAT	0	0	0	0	0	0	0	0	0	0	0
EV/FCFF	0	0	0	0	0	0	0	0	0	0	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Balance sheet	BLANK	BLANK	BLANK	BLANK	BLANK	BLANK					
Assets (adjusted? 1/0)	0	0	0	0	0	0	0	0	0	0	0
Assets Year -7	BLANK	BLANK	BLANK	BLANK	BLANK	BLANK					
Assets Year -6	BLANK	BLANK	BLANK	BLANK	BLANK	BLANK					
Assets Year -5	BLANK	BLANK	BLANK	BLANK	BLANK	BLANK					
Assets Year -4	2703,3	744,8	2703,3	744,8	2703,3	745,0	1	0	1	0	1
Assets Year -3	525,6	96,1	525,6	96,1	525,6	96,0	1	0	1	0	1
Assets Year -2	584,0	88,8	584,0	88,8	584,0	89,0	1	0	1	0	1
Assets Year -1	627,7	131,7	627,7	131,7	627,7	132,0	1	0	1	0	1
Assets Year 0	924,4	252,2	924,4	252,2	924,4	252,0	1	0	1	0	1

Intercoder reliability results

Having explained the coding process, the results of the tests will be presented. As previously mentioned, the ICR tests were conducted in two iterations; firstly, a test consisting of n = 10 valuations, coded by all six members of the group prior to coding the full sample, and secondly, a test of n = 6 valuations, coded by all six members of the group during the coding of the full sample, to ensure continued reliability.

The ICR results has thus been split into each iteration. Further, to investigate potential differences in reliability between observations in our dataset, we split up our ICR test scores into two groups:

- Binary variables (0/1) – namely those investigating whether an analyst has satisfied certain criteria, e.g., the use of EV/EBIT multiples on peers, satisfaction of steady state growth or similar.
- Financial variables (numerical values) – namely reported revenues, forecasted revenues and so forth.

The complete ICR test results for each code are enclosed in Appendix D. What follows is the aggregated results.

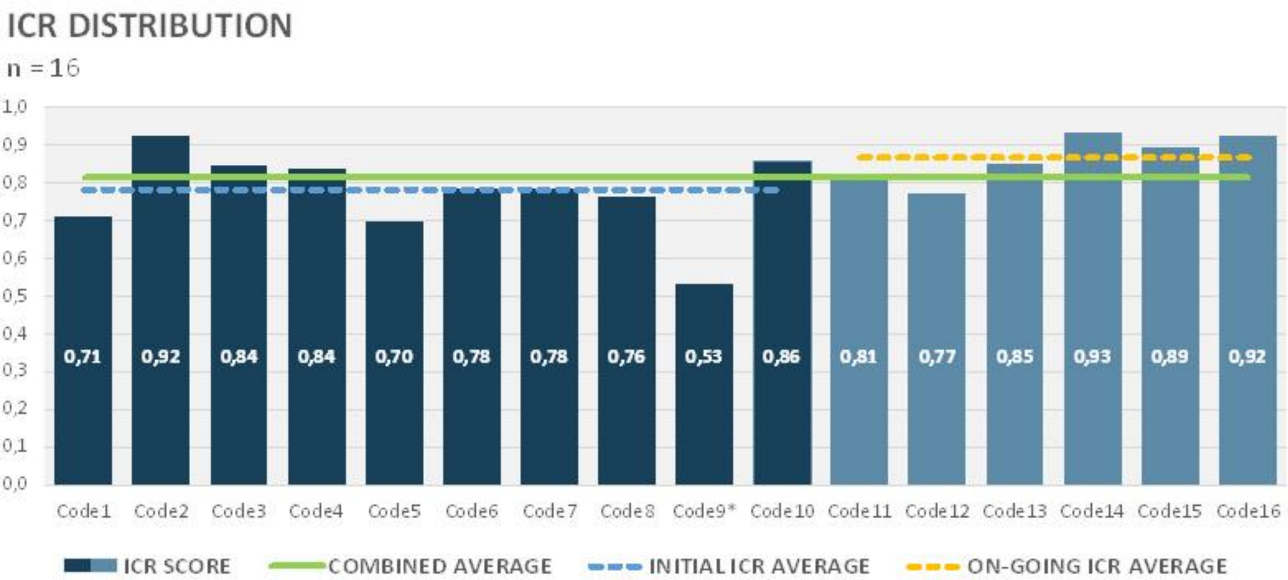


Figure 5: Distribution of ICR scores. Own contribution.

As evident from the above graph, the ICR score is seen to fluctuate between 0,70 and 0,93, with Code 9 being an outlier[6]. In general, the ICR score for the on-going ICR is higher than that of the initial codes, with averages of 0,87 and 0,78 respectively, indicating that we have progressively

[6] Code 9 has a considerably lower ICR score than the rest. This is mostly due to a great variety in the amount of observations coded between each coder. The observations that have been coded by all coders is to a larger degree identical.

become better in aligning our judgments of the sample data. Further, breaking down the ICR scores by the binary and financial groups, we see a clear difference in performance:

Table 4: ICR Scores by group and cluster. Own contribution

INITIAL ICR TEST (N=10)	ICR SCORE	**ON-GOING ICR TEST (N=6)**	ICR SCORE
BINARY VARIABLES	0,94	BINARY VARIABLES	0,93
FINANCE VARIABLES	0,74	FINANCE VARIABLES	0,85
ALL VARIABLES	**0,78**	**ALL VARIABLES**	**0,87**

COMBINED (N = 16)	ICR SCORE
BINARY VARIABLES	0,94
FINANCE VARIABLES	0,78
ALL VARIABLES	**0,82**

Not surprisingly, the financial variables are prone to more inconsistency than the binary variables. This makes sense, as the complexity increases and the possibility for errors is larger – and, intuitively, the possibility of agreeing by chance is much lower. However, with combined ICR scores of 0,94 for binary variables and 0,78 for financial variables, we feel satisfied in using the dataset for further analysis, although the financial variables should be used with attention to the robustness of the derived results, as some discrepancy in our sample data is present.

When evaluating our coding performance, it helps to compare the results to relevant literature. According to Neuendorf (2002), there exist no golden standard in terms of what ICR is acceptable, however, agreeability coefficients of >0.8 are generally acceptable in most situations. Landis & Koch (1977) recommend an interpretation frame with coefficients ranging >0.61<0.8 as substantially agreeable, and >0.81 as nearly perfect agreeability. Ultimately, all guidelines and standards are somehow arbitrary without context. In our case, considering the complexity and depth of our dataset, we consider an ICR score of 0,82 based on 11.414 coded observations in 16 different valuations, by six different coders, as an acceptable basis for further analysis. In comparison, the full sample of our dataset consists of approximately 41.000 observations from 321 valuations, thus the relative basis and size of our ICR sample is considered as satisfactory to indicate a reliable dataset.

4.4 Variables and definitions

In the following section, a brief overview of the variables included in our analysis will be presented. The complete list of variables including explanations is enclosed in Appendix C.

4.4.1 Independent variables

Based on the hypotheses proposed in the earlier section, a list of independent variables was collected. In this section, a brief overview of the computed variables will be given.

Market capitalization

As previously mentioned, market capitalization has often been used in literature as a proxy for information availability (Falkenstein, 1996; Cheng et al., 2020; Bonini et al., 2010), with larger companies having more accessible information. Therefore, we sought to approximate the market capitalization at T_0. From our codebook, we already had the actual share price at the date of valuation, as well as the implied share price and the total equity value. Therefore, we could derive the approximate market capitalization by the following:

Equation 8

$$Market\ Capitalization_{\ T_0} = \frac{Analysts'\ equity\ value}{Analysts'\ implied\ shareprice} * Actual\ shareprice$$

However, in order to compare market capitalizations, we needed to convert it to a common currency. A total of 12 different currencies were used at valuation dates spanning from 2008 to 2021. Through Google Finance, we were able to gather exchange rates from all currencies to USD at the given valuation dates. As some dates were weekends or bank holidays, we fetched the next available trading-day on these currencies, and therefore minor discrepancies may be present on these market capitalization measures. The complete list of currency conversions used is enclosed in Appendix F.

Having converted all figures to USD, we found that the market capitalization ranges from approximately 30 million USD in Brøndbyernes IF to 1.034 billion USD in Apple Inc.. As we do not expect an analyst to be a thousand times more accurate on a 30m USD company compared to a 30bn USD company, the market capitalization was log-transformed at a base of 10. With a log-transformed market cap variable, we seek to flatten out the exponential nature of the market cap variable in our dataset.

In addition to the market capitalization variable as a proxy for information availability, we will also test the price-earnings (P/E) ratio at T_0, as previous literature has shown that larger analyst coverage increases share prices and P/E ratios (Doukas, Kim, & Pantzalis, 2005).

Volatility variables

As argued previously, we have a hypothesis that increased historical volatility would increase the difficulty of target price forecasting. Therefore, we computed the historical volatility on revenue, NOPAT, asset turnover (ATO) and ROIC.

As absolute standard deviation is difficult to compare between companies, we decided to compute the relative standard deviation on the aforementioned financials. The formula is illustrated for revenue below:

Equation 9

$$RSD\ Revenue = \frac{Absolute\ standard\ deviation\ revenue_{T-7;0}}{Average\ absolute\ revenue_{T-7;0}}$$

Forecasting variables

In order to compare analysts' forecasting behavior, and to see whether deviations between historical performance and forecasting had any impact on target price accuracy, we computed average revenue growth rates, average NOPAT margins and average ATO levels for both historical and forecasting periods.

For revenue, we decided to compute the geometric average growth rates (CAGR), as to capture the real performance in terms of sales. Consider the following: A company reports historical revenues of USD 100m, 115m and 100m in years one to three respectively. This company would have growth rates of 15% and -13% in year two and three, giving an arithmetic average growth rate of 1% historically, while the geometric average would be 0%. If the analyst were to forecast future revenues on the basis that average growth has been 1%, he would, all else equal, overestimate the company's ability to grow. Therefore, we decided to compute growth rates as CAGR rates on both historic revenue and forecasted revenue, using the following formula:

Equation 10

$$Revenue\ CAGR = \left(\frac{Revenue_{T_0}}{Revenue_{T=FY}}\right)^{\frac{1}{FY-T_0}} - 1$$

Where FY = First year of historic period and T_0 = Last year of historic period.

For average NOPAT-margin, ATO and ROIC, we have computed the arithmetic mean, as no compounding effect is present on these measures.

Lastly, we are interested in seeing the way in which the analysts deviate from historical measures in their forecasts. The hypothesis is that analysts deviating a lot from historical performance will have larger forecasting errors, as they are exhibiting more bias in either positive or negative direction. To measure this deviation, a simple variable was computed by dividing the forecasted CAGR rates with the historical CAGR rates, minus one. A negative figure would then represent a drop in forecasted CAGR compared to historical CAGR and vice-versa. One issue with this variable is that negative values can occur in two occasions; if forecasted CAGR is lower than historical, or if either the historical or the forecasted revenue CAGR is negative. To circumvent this issue and avoid interference with the interpretation of our analysis, all valuations with negative CAGR values were set to have blank deviation measures. Rather, we computed the deviation in percentage points for these valuations. The intuition behind not choosing percentage points as the main point of measurement is that a growth deviation of one percentage-point would, all else equal, not matter as much if CAGR was 50% historically as if it was only 2% historically.

The same deviation variables – both absolute (percentage points) and relative - were computed on NOPAT, ATO and ROIC, with deviations between historical/forecast and forecast/terminal year.

4.4.2 Dependent variables

The efficient market hypothesis states that the current market price of an equity is the true price (Damodaran A., 2002). The justification of equity analyst relies on the opposite belief, that the market is not totally efficient, and that stocks trade at levels different from fundamental values. Equity analysts believe that they, through thorough due diligence and analysis of a company and its industry, can look through market deficiencies and come up with a fairer price. Therefore, we cannot judge the accuracy of an equity report by comparing it to the current market price of the equity, instead we must look at the price development over a certain period. Our dataset does not include data on the specific price horizon for each project, but the literature seems to agree on a horizon of 12 months (Imam et al., 2013; Sayed, 2015; Bonini et al., 2019; etc.). As we cannot confidently confirm the horizon of the target price from the valuations, we will not limit ourselves to the 12-month horizon, but also be looking at horizons of 24 months and 36 months. Intuitively the forecast accuracy should decline, as intuitively uncertainty increases with the horizon.

As the exact publish date of the projects in our dataset is unknown, we set the "publish date" equal to the mentioned valuation date in each project. This valuation date is in most instances equal to the last day in the most recent fiscal year but is in some instances discounted in order to be in better agreement with the actual publication date of the project, and not the last financial year end. 131 out of 293 analyzed DCF valuations have used some kind of discounting to change the valuation date, either by discounting the equity value / share price with a factor after the valuation, or used a skewed unit of time inside the valuation model itself (e.g., 0.5, 1.5, 2.5 etc.)

Using Capital IQ, we fetch the highest and lowest price for each share in the year 1, 2 and 3 following the valuation date. For each company we also construct a sell (short) / buy (long) recommendation given by the actual share price noted in the thesis, and the target share price given by the valuation. If the target price is higher than the actual share price given in the report, we assign this project a buy recommendation and vice-versa. In the buy scenario we say that the target price is achieved if:

Equation 11

$$Highest\ price\ in\ Year_{t=1,2,3}\ after\ valuation \geq Target\ price$$

And in the sell scenario:

Equation 12

$$Lowest\ price\ in\ Year_{t=1,2,3}\ after\ valuation \leq Target\ price$$

If this condition is met, we will say that the Target price accuracy ($TPA_{t=1,2,3}$) = 1. If it is not met, we will set $TPA_{t=1,2,3}$= 0.

Another common recommendation strategi in equity research is the "hold" strategy. We have decided not to include this strategy, as it would require us to make some serious assumptions regarding the individual valuation reports, as "hold" strategies are rarely explicitly mentioned in our sample. Instead, we have decided to go with only two strategies, "buy" and "sell", which demand much simpler assumptions regarding the valuations. Including a "hold" strategy would also minimize our sample sizes for "buy" and "sell". Our hope is that this will yield more definitive conclusions and better preserve the integrity of our analysis.

In order to more thoroughly understand the accuracy of the forecasts, we also introduce a variable to measure the forecast error, which describes how precise the target price is to the actual realized price. We calculate the target price forecast error as (TPFE):

Equation 13

$$TPFE_{t=1,2,3} = \frac{Target\ price - Actual\ share\ price_{t=1,2,3}}{Target\ price}$$

We will use these variables to determine how the budgeting behavior and the selected valuation approach influence the accuracy of the final target price given by the projects. This approach will enable us to judge both sell and buy scenarios on equal terms. Independently of either being a sell or buy recommendation, each valuation will receive a score determined by how accurate the target price is relative to the realized price in each horizon. A positive TPFE indicates a too high target price, and a negative TPFE indicates a too low target price.

The problem with the above variable arises when calculating the mean. Positive and negative TPFE's will "cancel" each other out. In Table 5 below, different scenarios are illustrated with the corresponding TPFE. The mean of TPFE in this example is equal to zero, and therefore does not tell us much about the forecast error independent of the analyst overshooting or undershooting. We therefore introduced another variable, STPFE, which is the absolute value (or standardized), version of TPFE. This variable tells us about the target price forecast error independent of being a “buy” or “sell” recommendation, or whether the analyst overshot or undershot the target price in relation to the actual price at each horizon t=1,2,3.

Table 5: TPFE and STPFE example

Target price (TP)	100	100	100
Actual price$_{t=1,2,3}$	110	100	90
TPFE$_{t=1,2,3}$	-10%	0%	10%
STPFE$_{t=1,2,3}$	10%	0%	10%

Apart from uncovering the forecast error of each valuation we also seek to uncover the optimism of each valuation. Inspired by Bjerregaard-Nielsen (2015) we need to calculate the implicit return and the realized return. We calculate the implied return by taking the target price given by each valuation and dividing it with the actual share price noted in each valuation.

Equation 14

$$Implied\ return_{t=0} = \frac{Target\ price}{Actual\ share\ price_{t=0}} - 1$$

We calculate the realized return by dividing the actual share price on each horizon (e.g., 1, 2 and 3 years), with the actual share price noted in each project.

Equation 15

$$Realized\ return_{t=1,2,3} = \frac{Actual\ share\ price_{t=1,2,3}}{Actual\ share\ price_{t=0}} - 1$$

The final step is to calculate the excessive implicit return on each horizon:

Equation 16

$$Excessive\ implied\ return_{t=1,2,3} = Implicit\ return - Realized\ return_{t=1,2,3}$$

This is mathematically equivalent to the following expression:

Equation 17

$$Excessive\ implied\ return_{t=1,2,3} = \frac{Target\ price - Actual\ share\ price_{t=1,2,3}}{Actual\ share\ price_{t=0}}$$

The difference between TPFE and Excess Implied Return (EIR) is therefore based on whether you base the discrepancy between the realized price and the target price on the actual share price at t=0, or the target price.

This approach will enable us to uncover the bias in each valuation. If a valuation has a high implicit return but a low excessive implicit return this valuation will have less positive bias than a valuation with a low implicit return but a high excessive implicit return. Different scenarios are illustrated in Table 6: Illustrative examples of accuracy and bias variables . The leftmost example has a target price of 110, but the actual share price on the given horizon is 85, and the analyst therefore has a positive bias. The rightmost example has a target price of 90 but the actual share price on the given horizon is 115, and the analyst therefore has a negative bias.

Table 6: Illustrative examples of accuracy and bias variables

Recommendation	Buy	Buy	Buy	Sell	Sell	Sell
Target price (TP)	110	110	110	90	90	90
Actual price$_{t=0}$	100	100	100	100	100	100
Actual price$_{t=1,23}$	85	100	115	85	100	115
Implicit return	10%	10%	10%	10%	10%	10%
Realized return	-15%	0%	15%	-15%	0%	15%
EIR	25%	10%	-5%	5%	-10%	-25%

When calculating TPA on each horizon we will not be using the closing price at time t=1,2,3 but be using a high / low figure for each horizon. This has the advantage that we take the intra-period volatility into consideration, and thus provide us with a more nuanced look on the share price development. If we only looked at the closing price at the end of each period, we would merely get a "snapshot" of the share price development. We believe that approach to be unfair to the analysts, as we can hardly expect them to forecast the share price to some exact time in the future. When calculating the excessive implicit return and TPFE, we have decided to not take the high / low approach but rather using the closing price at the end of each period. Our approach here differs from the TPA, in order to get a more nuanced look at the share price development and forecast accuracy. When realizing returns you need to actual sell your assets (or buy if short recommendation). Because you never know when an asset has reached its lowest or highest price in real life, we have decided to use the share price at the end of each period in this instance.

4.5 Outliers

As in any empirical study, the dataset needs to be put under scrutiny for the occurrence of outliers. Especially so, since our method of data collection consists of a lot of manual data-coding.

A common way of detecting outliers is by use of the interquartile range (IQR), namely by threating observations outside of ±1.5x the IQR as outliers. However, for datasets with absolute values, IQR tests can prove difficult to utilize. Consider a revenue series with extreme drops in revenue in the historical period, and a somewhat modest growth in the forecast period:

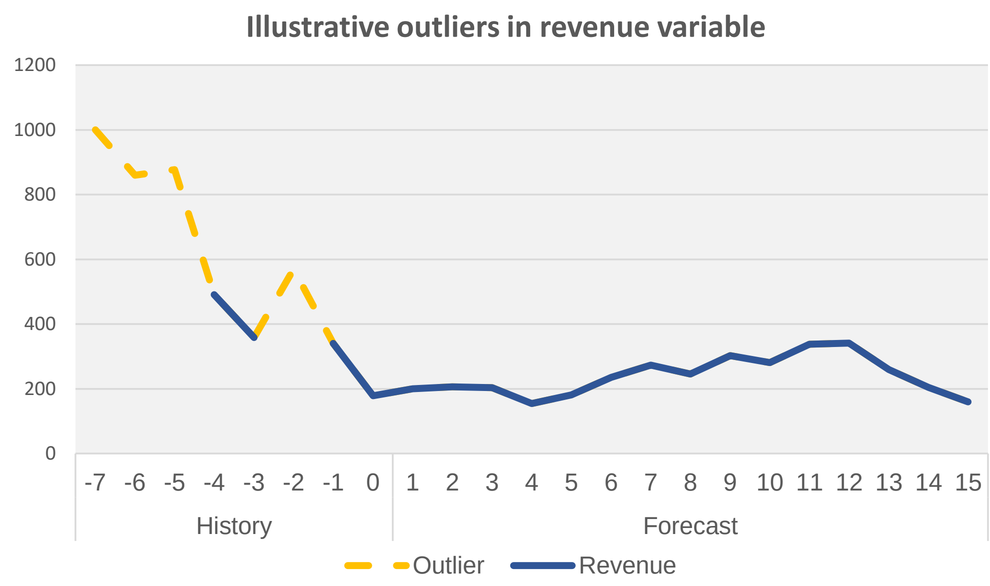


Figure 6: Illustrative example of revenue outliers

In Figure 6 an example is illustrated where an IQR test would exclude revenues within the -28 and 588 range, meaning that the first three years are excluded, along with the -2 year, creating a gap in revenue. This would mean, in practice, that we would be unable to use this entire revenue series for analysis, as gaps in revenue would make certain variables impossible to calculate. While the above is an illustrative example of a revenue data series, the concept applies to all financial variables in our dataset, and as such, intra-case outliers are very difficult to remove on absolute data. To circumvent this issue, we converted all absolute financials to relative measures – either through DuPont decomposition as previously mentioned, or through volatility or growth variables. This way, we were able to filter outliers on between-case basis.

When detecting outliers on relative variables, we judged each variable independently. For instance, if we were to apply the IQR method on our WACC variable, which has a pre-cleaned range of 2,3% - 89,1%, we would remove a total of 31 observations, including viable observations such as WACC = 12,6%. Therefore, before using any automated outlier detection methods, each variable had to be cleaned from obvious outliers – such as a WACC of 89,1%, which was done through visually inspecting each variable through histograms (Appendix E). From here, we proceeded to compute 1st and 99th percentiles and excluded observations outside this range, if the histograms showed outliers. Obviously, this was only necessary on financial and computed variables – as such, binary (e.g., gender) and date variables (e.g. establishment year) were not considered for outliers.

From this process, we removed between zero and 85 outliers from each variable. The variable with most outliers was the price-earnings ratio at T_0, where a lot of outliers occurred due to either currency conversion errors and/or mistyping during coding. We decided to remove outliers on a per-variable basis, and not remove entire cases if one outlier was present, as it intuitively does not matter whether a valuation has currency conversion error in P/E ratios when comparing a non-related variable, such as gender or study programme. However, for analyses containing multiple variables with differences in sample size, we only included valuations where both variables were present for all observations, as to ensure robustness of our statistical tests. For instance, consider a timeline of target price errors, where T_1 errors are 25% with n = 290, and T_2 errors are 15% with n = 250. If we were to compare these means, our ability to infer statistically significant results would suffer, as we do not know whether the forecast errors decrease because the most accurate valuations are removed in T2 forecast errors. Based on the above outlier removal process as well as the data collection process described earlier, our final fallout process is illustrated as follows:

Table 7: Fallout table of excluded cases and outliers

Fallout steps	Δ	n	% of search	% of codebook
1. Initial search		856	100,0%	na
2. Automatic exclusion of options and LBO valuations	(493)	363	42,4%	na
3. Manual exclusion of irrelevant projects	(85)	278	32,5%	na
3. Addition of confidential projects	43	321	37,5%	na
Codebook		321	37,5%	100,0%
4. Removal of non-DCF valuations	(24)	297	34,7%	92,5%
5. Removal of outliers	0-(85)	212-297	24-34%	66-92%
Final N		**212-297**	**24-34%**	**66-92%**

4.6 Statistics methodology

In this section we will describe the different statistical methods applied in our analysis. We will compare means, analyze correlations and perform regression analysis. We will hypothesize based on our literature review and determine whether we can reject the null hypotheses on a either a.90,.95 or.99 significance level. If nothing else is mentioned, the following section will be based on Agresti & Franklin (2013).

Comparing means

In order to compare means we will employ T-tests and perform analysis of variance (one-way ANOVA) and in some instances further apply Turkey's post-hoc test.

The T-test is a statistic used to determine if there are significant differences in the mean of two groups, which could be the mean forecast error of males versus females. The T-test assumes normal distributed data, which we will try to ensure by removing outliers (see previous section). The null hypothesis in the T-test is that the means are equal, $\mu_1 = \mu_2$, which rejection depends on the significance of the test.

While the T-test is a way to compare the means of two groups, ANOVA is a way to compare the means of three groups or more. However, the ANOVA only tests if all group means are equal – as such, the interpretation value is quite low in determining differences between groups. Therefore, if the means are found to be statistically significant, we can perform Turkey's post-hoc test in order to determine which of the comparisons that are significant. Turkey's post-hoc analysis will enable us to compare all possible combinations of the mean. If we for example want to compare the mean forecast error of different study programmes, we would conduct an ANOVA in order to test whether the means differs on a certain level of significance. In order to analyze which combinations led to the result, Turkey's post-hoc analysis will show us all possible combinations of study programmes, and whether the differences in the specific mean are significant.

Correlations

In order to get a quick overview of the relationship between different variables we will use Pearson's correlation. Pearson's correlation examines the linear relationship between two variables. Pearson's correlation takes on a value between 0 and 1, where 1 is a perfect linear relationship between the two variables. Pearson's correlation assumes the variables to be normally distributed, which can be checked by plotting the variables on histograms or boxplots. The variable plots are attached in Appendix E.

Regression

In some instances, we would like to examine how independent variable(s) influence a dependent variable, and in order to achieve this, we will perform both univariate and multivariate linear and logistic regression.

Ordinary least squares (OLS) regression seeks to minimize the absolute distance (error) between actual observations and a linear model, with respect to the coefficients in the model. It therefore seeks to find a linear relationship between the dependent variable and some independent variable(s). Logistic regression applies the same methodology[7], but is used when the dependent variable is binary, as is the case with TPA (e.g., either 0 or 1). A general linear model has the formula:

---

[7] This is a simplification, as our statistical software utilizes a maximum likelihood estimator to determine the coefficients (but OLS could be used), as this has better statistical properties.

Equation 18

$$f(X) = \beta_0 + \beta_1 X$$

In order to map the dependent variable between 0 and 1, it utilizes the logit function:

Equation 19

$$f(x) = \begin{cases} 1, if \ \dfrac{1}{1 + e^{-(\beta_0 + \beta_1 X)}} > 0{,}5 \\ 0, Else \end{cases}$$

Where $p(X) = \frac{1}{1+e^{-(\beta_0+\beta_1 X)}}$ is the probability that a certain observation X is equal to one. Generally, the dependent variable is assigned the value one if the probability is above 0,5. in our analysis we will generally look at the log-odds, which can be directly derived from the probability. Assume p(x) to be the probability of belonging to category 1, then the log-odds can be calculated with the following function:

Equation 20

$$\log -odds = Log\left(\frac{p(X)}{1 - p(X)}\right) = \beta_0 + \beta_1 X$$

If, for example, 60% of the analysts will satisfy TPA, this will imply log-odds of log(1,5)=0,41. This also has the implication, that say, if the coefficient β_1 is equal to 0,2, the log-odds will then increase with 0,2 every time X increases with one. If you further exponentiate the beta coefficient, exp(β_1)=exp(0,2)=1,22, then you get the odds. If X was equal to 1, e.g. using a specific valuation model, then your probability of belonging to category 1, would increase with (1,22-1)=22%, over not using this particular model.

Different assumptions regarding linear least squares must be satisfied in order to obtain an unbiased result. The first one is the assumption of linearity – there must exist a linear relationship between the dependent and independent variables in the model, which can be checked by using scatterplots. The second is the assumption of non-multicollinearity between the independent variables. If the variables in multivariate OLS are too highly correlated, the effect of each variable cannot be separated, thus leading to an unreliable model. The last assumption is that the residual errors in the OLS regression should be normal distributed. With a sufficiently large sample this should not be an issue, as the central limit theorem states that with a reasonable large sample, the errors should be normal distributed. This is generally the case when the sample exceeds 30 observations, which should easily be satisfied in our situation.

4.7 Limitations

As in any study, we have encountered limitations in our methodological and data collection choices which might affect our ability to infer statistically significant results. In this section, the methodological limitations will be elaborated.

In our calculation of TPFE and EIR we have downloaded adjusted closing prices for the first, second and third year following the valuation date. This closing price is adjusted for dividends and will thus give us a precise image of the true return of the equity. In calculating our TPA measure, we have downloaded a 52-week trailing high / low figure of the share price in the first, second and third year after the valuation date. In year two, we use the highest and lowest price in the two-year span. The same procedure is used in year three, taking the highest / lowest price in the three years after the valuation date. The intuition here is that if target prices are hit sometime during year one, it would per default also have been hit sometime from T_0 to T_2, and as such, the range in which the target price is deemed accurate increases with added time. This figure is not adjusted for dividends however and is therefore not equal to the true return of the equity. This error can potentially grow as we get further away from the valuation date. This might pose an issue with companies exhibiting a high pay-out ratio, as analysts covering these companies will be penalized disproportionately compared to companies with lower pay-out ratios in the buy recommendation scenario, as the dividends are not included in the return. The opposite is true in the sell recommendation scenario, as the investor would then need to pay the dividend to whomever the share was borrowed from. Thus, in terms of returns, the buy scenario is negatively affected by the exclusion of dividends, and the sell scenario is positively affected.

When we compare different groups, it is necessary that the two has the same prerequisites. When comparing female and male forecast accuracy for example, we would have to assume that the distribution of companies/industries is the same in both samples, as some industries and companies might be harder to forecast (e.g., volatile industries are generally more difficult to forecast). If the independent variable put under scrutiny is simultaneously skewed toward a specific industry, this could lead to a biased result. Say, if men are generally affectionate toward "easy to forecast" industries, this could potentially be the explanatory factor behind the observed differences in target price forecast error between the groups, rather than a general difference between genders. This is always a balancing act between adjusting for these prerequisites and attaining a sufficient sample size, as smaller sample sizes lead to, all else equal, weaker conclusions. On the other hand, adjusting for these factors can lead to other issues. If, for example, that female analysts are truly more accurate than male analysts, we would not be able to include female and male analyst in the same group when examining the forecast error using different valuation models, unless the proportions using each model is identical. If we then decided to exclude female analyst altogether, because they are underrepresented in our sample, we would not be able to draw any conclusion on analysts as a general group. In this specific example, we would have to rely on the assumption that the genders are equally likely to choose certain valuation models.

The above mostly relates to our secondary research question: "Which factors influence the accuracy or bias derived from student valuations?", where it is necessary to divide the dataset into different comparable groups. When answering our first research question: "Are target prices derived from student valuations accurate and unbiased?", we will be looking at students as a group, and different prerequisites are not as important, as we believe our sample of students to be a good general representation of analysts, as students who chose to do valuations in their final thesis are assumed to share similarities with "real world" analysts.

After coding our dataset, we found that the proportions of companies and industries is rather skewed toward certain companies and industries. As evident from Appendix G and H, 52% of the companies in our sample belong to the same 10 industries, and one-third of the valuations are made on the same 10 danish companies. This speaks to the general applicability of our results across different industries and companies, and thus the inferential abilities may be limited by this distribution of valuations.

This specific example leads to issues regarding the validity and application of our results. As previously mentioned, we use students as proxies for "real world" analysts. This assumption is affected by both how frugal and experienced our subjects are. It is fair to assume that students are less experienced than real world analyst. On the other the hand, we would argue that students are generally more thorough in their analysis, as the fewest professional analyst spend up to half a year on their reports, and they are less likely to, single-handedly, publish a report counting 70-120 pages. Further, our sample can, in some respects, propose less biased results. Students are not to the same degree affected by the principal-agent problem preciously described in section 3.1. If these variables are taken out of the equation, we can get a rawer look upon what drives the target price accuracy of analysts, and their methodology, free from potential biases. If we compare the performance of students in our study with accuracy results reported in previous literature, we find that their performance is not differing by large margins, as evident from Table 8 below.

Table 8: Comparison of TPA and TPFE measures between analysts and students

Source	TPA (52 week)	TPFE (equivalent)
Bonini et. al (2010)	33%	NA
Bradshaw & Brown (2006)	45%	NA
Sayed (2015)	(only DCF) 70%	NA
Asquith, Mikhail and Au (2005)	54%	37%
Kerl (2011)	57%	42%
Students	53%	27%
Average (excl. students)	52%	40%
Median (excl. students)	54%	40%

As evident, on both TPA and TPFE, students have higher accuracy and less target price forecast errors than the average analyst from previous literature, and just one percentage-point worse than the median for actual analysts on the TPA measure. These results are not presented as a definitive answer, but merely as a guide to the differences between real-world analysts and students. While we acknowledge that differences between students and analysts exist, we are comfortable utilizing this group as proxies, as the methodology employed are highly identical.

In our data collection we have adjusted for (reverse) splits, which have taken place after the valuation date. In Bonini et al. (2010), it is noted that these (reverse) splits can affect the volatility of the share and therefore also the share price. Analysts can hardly be responsible for these changes, and it would therefore be unfair to attribute these to their abilities as analyst. When calculating TPFE and TPA measures, we use post-split prices by adjusting the target prices (see section 4.2.2) – however, if the arguments put forth by Bonini et al. (2010) are valid, we are unable to correct for potential increases in volatility caused by stock splits, which might be a source of error in our analysis.

Further, as the economy moves in cycles, this will also naturally affect share prices. Some periods are characterized by booms and other busts. Our sample spans more than a decade, from the end of the 00s to the beginning of the 20s. Events such as the financial crises in 2007-2008 and the most recent corona pandemic can hardly be foreseen by analysts. However, these periods will not be excluded, but rather acknowledged and discussed by descriptively showing the accuracy and bias measures on a per-year basis.

One fundamental limitation lies within the nature of our research design. As we have taken a deductive approach, we risk being biased in the very factors we accredit as being explanatory for our analysis. We cannot test for factors which are not included in our dataset, and as such, we have, to the best of our ability, used exploratory readings of previous literature to cope with this issue. This entails that many factors that may influence analyst's accuracy are left outside the scope of this thesis. An example could be the beta used to estimate the cost of capital in the calculation of WACC. This could tell us something about the specific risk associated with each company, which might have been an interesting addition, to estimate the risk of the company's equity and its effect on target price forecast error and target price accuracy. This is also a balancing act between the extent of the codebook, and the time spent gathering data.

Finally, the nature of our data collection method of collaborative coding further increases risks of data reliability. We sought to circumvent this issue and highlight the influence by performing an intercoder reliability test. As reported in section 4.3, we achieved an ICR score of 82% agreeability based on 11.414 coded observations in 16 different valuations. According to previous literature, this is generally deemed acceptable, but it still leaves room for human error, potentially affecting our results.

5.0 Results

This section will present the results of this thesis by testing the hypotheses stated in section 3.0. The first subsection, section 5.1, will seek to answer the first research question, namely whether analyst valuations are accurate or biased. The second subsection, section 5.2, will seek to answer our secondary research question, namely which factors are of importance in terms of analyst accuracy or bias.

5.1 Primary research question

In this section, we will seek to answer the following research question using the statistical methods outlined in section 4.6:

> *"Are target prices derived from student valuations accurate and unbiased?"*

To answer this research question, a list of accuracy and bias measurements have been introduced, as listed in section 4.4.2, namely Target Price Accuracy (TPA), which is a binary variable indicating whether the target price has been met on intra-year highs and lows, on 12-, 24- and 36-month timeframes. Further, a Target Price Forecast Error (TPFE) variable was introduced, which measures the aggressiveness or conservativeness of target prices. As TPFE is zero for zero-error forecasts, negative for too conservative forecasts and positive for too aggressive forecasts, we computed a standardized version of the TPFE variable (STPFE), with all-positive values, seeking to measure the absolute levels of forecast error. Finally, we introduced an Excess Implied Return (EIR) variable, which measures positive or negative bias. Our first hypothesis is as follows:

> *$H1_0$: Analyst target prices are accurate and unbiased.*
>
> *$H1_a$: Analyst target prices are inaccurate and biased.*

In the Table 9 below, the descriptive statistics are presented on our sample data on the abovementioned variables.

Table 9: Accuracy and bias results on full sample

		Target Price met?		TPFE		STPFE		EIR	
	n	Yes	No	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Year 1	274	0,54	0,46	0,27	0,44	0,43	0,27	0,38	0,84
Year 2	270	0,65	0,35	0,15	0,63	0,51	0,40	0,21	1,34
Year 3	267	0,71	0,29	-0,01	0,84	0,62	0,58	0,03	1,50

From the above table it is evident that analysts in general are fairly inaccurate. 53,6% of the target prices proposed by the analysts of our sample were hit during the first year, with increased TPA to 70,7% in year three. This makes intuitive sense, as the TPA exploits the volatility in the share, and thus increases the chances of hitting the target share price, the longer the horizon is. TPA can only increase, never decrease, with the length of the horizon. TPFE was, on average, 27,1% in the first year, falling to -0,9% in year three, indicating that the analysts in our sample are, on average, overly aggressive in their target price estimates, if they were issued with a one-year time horizon. As evident from the standard deviation of TPFE, there exist great variance in forecast errors, and the rather low TPFE in year three is not to be interpreted as the analysts being accurate, but rather as a product of a wider span of errors, ultimately cancelling each other out, and thus deriving at a small TPFE measure. Looking at STPFE, analysts have on average performed with quite large errors of 43,3% in year one, increasing to 62,4% in year three. On EIR, analysts are on average systematically biased, proposing mean implied returns that are 38,1% larger than actual one-year realized returns. This figure drops severely to only 2,9% excess implied return in year three, but with a larger standard deviation to follow. If investigating the relationship between our independent variables, we derive the following:

Pearson correlations

* Significant at the p = 0.05 level. ** Significant at the p < 0.01 level

	TPA_1	TPA_2	TPA_3	TPFE_1	TPFE_2	TPFE_3	STPFE_1	STPFE_2	STPFE_3	EIR_1	EIR_2	EIR_3
TPA_1	1	0,78**	0,69**	-0,27**	-0,21**	-0,16**	-0,41**	-0,08	0,04	-0,23**	-0,17**	-0,16**
TPA_2		1	0,89**	-0,27**	-0,25**	-0,22**	-0,45**	-0,20**	-0,05	-0,21**	-0,17**	-0,17**
TPA_3			1	-0,20**	-0,23**	-0,22**	-0,41**	-0,27**	-0,13*	-0,16**	-0,13*	-0,15**
TPFE_1				1	0,80**	0,63**	0,55**	-0,08	-0,20**	0,88**	0,64**	0,61**
TPFE_2					1	0,88**	0,44**	-0,30**	-0,46**	0,67**	0,91**	0,88**
TPFE_3						1	0,36**	-0,23**	-0,61**	0,47**	0,76**	0,94**
STPFE_1							1	0,51**	0,19**	0,48**	0,25**	0,21**
STPFE_2								1	0,74**	-0,10	-0,25**	-0,29**
STPFE_3									1	-0,15**	-0,37**	-0,55**
EIR_1										1	0,78**	0,67**
EIR_2											1	0,95**
EIR_3												1

Figure 7: Pearson correlation matrix for all combinations of independent variables

In Figure 7, the correlation coefficients are shown between our independent variables. As evident, and not surprisingly, the TPA is negatively correlated with TPFE on all years, and negatively correlated with STPFE in year one and two. What is interesting, and perhaps a bit surprising, is that STPFE in year two and three are significantly negatively correlated with EIR in all years. This indicates that as absolute forecast errors on longer horizons increase, the level of positive bias decreases significantly.

While the results from Table 9 show that analysts are inaccurate when pooled together, it is interesting to look at the deviation between different years of publication.

Table 10: Accuracy measures grouped by year of publication

Year	N	TPA	TPFE	STPFE	EIR
2008	5	60%	73%	73%	101%
2009	17	35%	46%	47%	67%
2010	18	50%	43%	48%	74%
2011	44	52%	42%	47%	66%
2012	42	45%	42%	47%	67%
2013	36	50%	22%	45%	16%
2014	26	81%	9%	28%	18%
2015	23	70%	24%	32%	27%
2016	37	57%	12%	34%	-8%
2017	30	57%	7%	40%	19%
2018	18	39%	36%	45%	54%
2019	11	55%	-10%	57%	-16%
2020	12	42%	23%	60%	60%
Total	**319**	**53,6%**	**27,1%**	**43,3%**	**38,1%**

In Table 10 we find that there is great variety in both the number of target prices hit and the forecast errors when grouping by year of publication. As evident, analysts are consistently positively biased and publish consistently aggressive target price estimates, with only TPFE in year 2019 and EIR in year 2016 and 2019 being negative, indicating a higher number of conservative estimates and negative bias in those years. As seen from the STPFE measure, analysts consistently make errors in the range of 28% in 2014 to 73% in 2008.

From the results in Table 10 it is clear that, while variety exists, analysts are consistently and significantly biased and inaccurate, even though some years have historically provided better estimates than others. From the results in Table 9, we showed that both TPFE, STPFE and EIR are significantly different from zero, and thus we reject our null hypothesis, as the analysts in our sample are, as shown, severely biased, and only a bit more than half of the target prices proposed had been met on a one-year horizon.

However, as analysts bias and accuracy may differ depending on the recommendation of the analyst, we seek to test the following hypothesis:

$H2_0$: Accuracy on buy recommendations = Accuracy on sell recommendations

$H2_a$: Accuracy on buy recommendations ≠ Accuracy on sell recommendations

To answer this hypothesis, the following descriptive statistics have been computed, along with an independent samples T-test.

Table 11: Analyst accuracy and bias, grouped by long or short recommendation

		TPA	TPFE		STPFE		EIR	
	N	Mean	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Buy								
Year 1	199	0,48	0,35	0,38	0,44	0,27	0,61	0,54
Year 2	198	0,60	0,22	0,58	0,50	0,37	0,48	0,66
Year 3	194	0,67	0,08	0,79	0,57	0,55	0,31	0,92
Sell								
Year 1	69	0,64	0,05	0,52	0,37	0,26	-0,23	1,15
Year 2	69	0,75	-0,05	0,73	0,54	0,48	-0,57	2,22
Year 3	69	0,77	-0,19	0,94	0,70	0,65	-0,71	2,30
Difference								
Year 1		-0,15**	0,30 **		0,06		0,84 **	
Year 2		-0,15**	0,27 **		-0,05		1,05 **	
Year 3		-0,10**	0,27 **		-0,13		1,02 **	

* Significant at the p = 0.05 level. ** Significant at the p < 0.01 level

In Table 11 we find that analysts are much more prone to proposing buy recommendations rather than sell recommendations, with more than twice as many buy recommendations than sell recommendations in our dataset. Perhaps this is simply a consequence of using students as proxies for analysts, as students are free to choose the company they cover, and thus pick companies they find interesting – often being companies of high recognition and growth.

Despite the overweight of buy recommendations, the sell recommendations seem to be a lot more accurate. Looking at mean TPA, the target prices from sell recommendation have been hit 15% more, with one-year TPA of 64% on sell side, compared to 48% on buy side. In terms of aggressiveness and conservatism of recommendations, the sell recommendations are, not surprisingly, much more conservative, with mean one-year TPFE differences of 30%, dropping to 27% in year two and three. However, the variance is larger on sell recommendations. Looking at STPFE, we find that there are no statistically significant differences in actual target price forecast errors, showing that both groups are somewhat equally inaccurate. Finally, we find that analysts proposing buy recommendations are, not surprisingly, positively biased, and analysts proposing sell recommendations are negatively biased. We do however find that analysts with sell recommendations have a lower mean bias in absolute terms, but with much larger variance.

Based on the above results, we reject our null hypothesis, as analyst accuracy is not similar for buy recommendations and sell recommendations.

5.2 Secondary research question

In the previous section, we found that analysts are not accurate, with only 54% hitting their issued target prices on one-year horizons and having absolute forecast errors of 43% on average. We will now seek to answer the following research question using the methods outlined in section 4.6:

"Which factors influence the accuracy or bias derived from student valuations?"

We will seek to answer the problem formulation using various tests on previously identified factors of interest. We have sought to group the theorized explanatory variables on a list of categories, namely:

- Analyst specific characteristics
- Financial factors
- Methodological-related factors

From the above categories, a list of hypotheses was formulated. Naturally, in order to test most of these hypotheses, we had to create groups based on the variable at hand – for instance, grouping men and women to test gender differences. For such tests to satisfy the underlying assumptions, we need the groups to be similar on different variables – however, as discussed in section 4.7, this has proved to be a difficult task, as an endless count of possible control-groupings can be made; study programme, year of publication, gender, buy/short recommendation, just to name a few. Were we to control for all these subsections, our per-group sample size would vastly suffer, and thus we would hardly be in a position to infer statistically significant differences. What follows is a structured presentation of the results of hypothesis testing, based on the categories presented above.

5.2.1 Analyst specific characteristics

Gender

The first hypothesis of interest is that of male and female analysts, and whether significant differences in accuracy or bias exists. To test this, we have formulated the following hypothesis:

$H3_0$: Male analysts' accuracy = Female analysts' accuracy

$H3_a$: Male analysts' accuracy ≠ Female analysts' accuracy

To answer the above hypothesis, an independent samples T-test was computed. The table below contains the descriptive statistics as well as the results of the T-test.

Table 12: Gender differences on accuracy measures

		TPA	TPFE		STPFE		EIR	
	N	Mean	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Male								
Year 1	226	0,54	0,29	0,39	0,41	0,26	0,45	0,85
Year 2	225	0,67	0,15	0,60	0,50	0,40	0,31	1,40
Year 3	222	0,73	0,00	0,84	0,61	0,59	0,15	1,55
Female								
Year 1	42	0,50	0,33	0,38	0,42	0,28	0,50	0,79
Year 2	42	0,60	0,34	0,37	0,41	0,39	0,51	1,01
Year 3	42	0,68	0,26	0,51	0,48	0,31	0,44	1,21
Difference								
Year 1		0,04	-0,03		-0,01		-0,05	
Year 2		0,07	-0,19 *		0,08		-0,21 *	
Year 3		0,06	-0,27 **		0,13		-0,29 **	

* Significant at the p = 0.05 level. ** Significant at the p < 0.01 level

As evident from the above, male analysts are vastly overrepresented in our sample. As evident from the TPA test, there are no significant differences in target price accuracy between men and women. Both groups' target prices have been met in roughly half of the cases on a one-year horizon. Surprisingly, female analysts are more prone to publishing too aggressive target prices, especially on two and three-year TPFE, where the mean female analyst has target price forecast errors of 34% and 26% respectively, whereas the TPFE of male analysts was 15% and 0% for two and three-year TPFE. As evident from the standard deviations, there is a wider span of TPFE measures in the male category, which is further cemented in the STPFE measure. Here, male analysts, on average, have larger absolute forecast errors in year two and three, although not statistically significant, implying that the lower TPFE measures of male analysts is due to conservative and aggressive target prices cancelling each other out. As such, there is no significant difference in the absolute forecast errors produced by male or female analysts.

In terms of bias, female analysts exhibit more positive bias in all three years, with statistically significant differences on two and three-year EIR. This is quite interesting, as it is contradictory to previous literature on the matter – namely that female analysts are more conservative. Based on the above results, we do however fail to reject the null hypothesis that female analyst's accuracy is equal to male analysts accuracy. As evident, there is no significant differences in neither TPA nor STPFE. Only the aggressiveness and bias measures of TPFE and EIR had significant results, indicating that the female analysts are more aggressive and exhibit more positive bias, on average, in our sample.

Study programme

As discovered in the literature review section, Hope & Fang (2020) found explanation power in the educational background of analysts in terms of their forecasting capabilities. As we have access to the students' study programmes, we will seek to test the following hypothesis:

H4$_0$: The study programme of the analyst does not influence the analysts' accuracy

H4$_a$: The study programme of the analyst does influence the analysts' accuracy

In order to investigate whether there exist differences in mean analyst accuracy and bias based on the analysts' education, we compiled the analysts in groups of study programmes. A total of 12 study programmes were present in our dataset. However, after cross-checking, we found some study programmes to be duplicates with different labeling[8]. Further, three study programmes had between one and two participating analysts, which was considered too low of a sample for meaningful analysis, and thus these programmes were stripped from the dataset[9], leaving eight remaining study programmes, with a total of n = 286 analysts. The below chart summarizes the conducted one-way ANOVA test, with grouped accuracy and bias means based on study programme.

Table 13: Differences in accuracy grouped by study programme

		Target Price Hit?			TPFE			STPFE			EIR		
Analyst study programme	N	Year 1	Year 2	Year 3	Year 1*	Year 2	Year 3	Year 1*	Year 2	Year 3	Year 1	Year 2	Year 3
AEF	37	0,49	0,65	0,70	0,14	-0,05	-0,27	0,46	0,52	0,69	0,08	-0,01	-0,26
ASC	23	0,39	0,61	0,65	0,25	0,14	-0,06	0,41	0,39	0,60	0,48	0,36	0,14
AUD	51	0,49	0,71	0,78	0,31	0,24	0,09	0,43	0,45	0,57	0,46	0,33	0,16
FIN	9	0,44	0,56	0,56	0,36	0,25	0,07	0,42	0,53	0,81	0,46	0,35	0,17
FIR	76	0,66	0,72	0,79	0,24	0,10	-0,07	0,34	0,50	0,63	0,30	0,12	-0,02
FSM	36	0,53	0,61	0,61	0,21	0,08	0,00	0,44	0,57	0,62	0,32	-0,24	-0,40
HD_AFM	31	0,42	0,55	0,65	0,40	0,23	0,10	0,44	0,58	0,57	0,55	0,37	0,20
HD_F	23	0,52	0,57	0,61	0,32	0,22	-0,03	0,54	0,56	0,68	0,63	0,48	0,28
Total average		**0,54**	**0,65**	**0,71**	**0,27**	**0,15**	**-0,01**	**0,43**	**0,51**	**0,62**	**0,38**	**0,21**	**0,03**

* Significant at the p = 0.1 level. ** Significant at the p = 0.05 level. *** Significant at the p < 0.01 level

As shown in the table above, the mean differences per study programme is seen to vary. On TPA, the lowest mean accuracy is for ASC in year 1, HD_AFM in year 2 and year 3. In general, FIR analysts were most frequently hitting their target prices on all three years, closely followed by AUD and AEF analysts.

On TPFE, we found statistically significant results between groups in year one, indicating that not all group means are equal in our ANOVA test. For TPFE, the AEF analysts were the most conservative, with TPFE measures of just 14% in year one, dropping to -27% in year three.

---

[8] E.g. "HD Finansiering" and "Graduate Diploma Finance"

[9] The stripped programmes were: HD Regnskab og Økonomistyring (n=1), HD International Business (n=2) and HD Finansiel Rådgivning (n=1)

From the ANOVA it is not evident which groups are statistically significant from each other – we therefore compute a Turkey's post-hoc test, which reveals the between-group differences, where AEF is significantly different from HD_F, with a mean TPFE difference of 0,18 *(p = 0.02).* The most aggressive group of analysts was HD_AFM, with TPFE measures reaching 40% in year one. Both HD-programmes were above the total means on one and two-year TPFE.

As evident, FIR analysts have the lowest absolute forecast errors measured on STPFE, followed by ASC and FIN analysts. One-year STPFE also showed statistical significance between groups. From the post-hoc test, we find statistically significant difference between FIR and HD_F analysts, with a mean difference of 0,20 (p = 0,03)

On EIR, all groups are seen to be positively biased on average, with only the FIR, FSM and AEF analysts being negatively biased on three-year EIR. When comparing on one-year EIR measures, AEF is the least biased group of analysts, with excess implied returns of only 8%. This is perhaps due to a lot of conservative estimates, as we saw that STPFE for this group was above average, indicating that they are still fairly inaccurate in their estimates, but in a more negative direction than the average.

In general, for all three measures, there seem to be a reversion to lower biases and more accuracy as time passes. The TPA is seen to increase from an average of 0,54 in year one to 0,71 in year three. This pattern is even more amplified in the EIR measure, with a rather considerable positive bias of 0,38 in year one, dropping to 0,03 in year three. This intuitively makes sense, as market values of equities historically have increased, and therefore it is expected that the degree of positive bias, when measured on three-year ahead prices, should, on average, be lower.

Based on the above results, we reject our null hypothesis. Analyst study programme does influence accuracy, as we found statistically significant differences between groups on the STPFE measure.

Groups or individuals

As previously mentioned, Hope & Fang (2020) found that analyst teams proposed more accurate and unbiased estimates than individual analysts. We therefore seek to answer the following hypothesis:

> *$H5_0$: Analyst groups' accuracy = individual analysts' accuracy*
>
> *$H5_a$: Analyst groups' accuracy ≠ individual analysts' accuracy*

To test this, we grouped our sample valuations on whether they were made by individual analysts, or by two analysts. The vast majority (n = 152) of valuations were made by individuals, but a sizable group (n = 91) also collaborated on their analysis. The results of our T-test are shown below.

Table 14: Differences in accuracy grouped by individual or groups of analysts

		TPA			TPFE			STPFE			EIR		
Number of analysts	n	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3
Individual	152	0,55	0,67	0,74	0,30	0,20	0,06	0,41	0,49	0,60	0,49	0,39	0,24
Group	91	0,52	0,64	0,70	0,29	0,14	0,01	0,40	0,47	0,57	0,41	0,26	0,13
Difference		0,03	0,03	0,03	0,01	0,06	0,05	0,01	0,02	0,04	0,08	0,13	0,12

* Significant at the p = 0.05 level. ** Significant at the p < 0.01 level

As evident from the above T-test results, there is a small difference on TPA, with groups of analysts being slightly worse in hitting their target prices in our sample. Further, the groups of analysts were shown to be slightly more conservative on their TPFE measure and had slightly fewer forecast errors on the STPFE measure. On EIR we found the largest differences – namely that groups of analysts generally were less positively biased, with mean EIR differences of 8-12% in one-to-three year EIR. None of our tests yielded statistically significant results.

As evident from the above results, we fail to reject the null hypothesis, and thus conclude that there is no difference in accuracy measures between individual analysts and groups of analysts, as none of our tests were statistically significant.

5.2.2 Financial variables

As previously argued, we have sought to decompose the drivers of FCFF into revenue growth, NOPAT-margin development and ATO development, as to uncover whether there is significant correlation between how these levers of firm value have been forecasted, and the impact they have on the accuracy of target prices. In this section, we will test these levers by seeking to answer the hypotheses outlined in section 3.2. Further, this section will seek to uncover whether the firm size has an impact on analysts' ability to accurately predict target prices.

We would expect that the financial factors would have an impact; algebraically, both revenue growth, NOPAT-margin and ATO are directly associated with FCFF. We therefore seek to uncover whether any specific forecasting behaviors lead to increases in accuracy or bias. To give an idea about the average forecasting behavior of our sample, we have illustrated the average development of revenue growth, NOPAT-margin, ATO and ROIC in the years T_{-7} through to T_{15} in Figure 8 below.

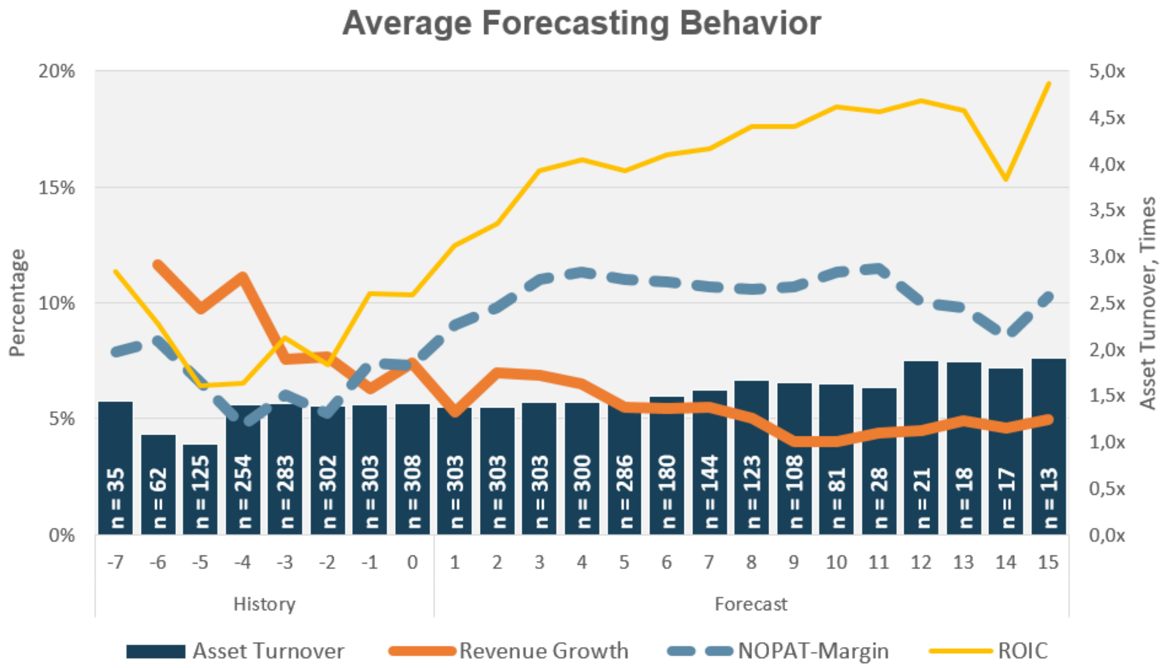


Figure 8: Average forecasting behavior of full sample, based on DuPont decomposition.

As evident, the average ATO level seems fairly stable around 1,5-1,8x throughout the period, with a rather stable transition from historic data to forecast, although with slight increases in the later forecast years, most likely caused by a rather quick fallout of forecasts after the 9th year. The average NOPAT-margin is seen to increase quite rapidly from T_0 to T_4, with average revenue growth rates falling during the same period. Intuitively this assumption seems fair, as the companies are reaching a steady-level, growth is set to decrease but profitability increases slightly. Interestingly however is the ROIC development, increasing very rapidly from the last historical year onwards, indicating that the analysts in our sample are somewhat optimistic about the company's ability to continuously create value.

Revenue growth

The first financial factor of interest is the revenue growth. Not surprisingly, revenue growth is an important metric to forecast, as it has an immediate relation to all other figures in the forecast. From our sample, we have computed three revenue-specific variables, as explained in section 4.4.1, namely the relative standard deviation of historic revenue, the historical and forecasted CAGR, and finally the deviation between historic and forecasted CAGR figures. In the table below, initial correlation tests are presented on the intercorrelation between our four independent variables:

Table 15: Correlation table of revenue variables

Pearson's Correlation	RSD Historic Revenue	Growth Deviation	Historic CAGR	Forecast CAGR
RSD Historic Revenue	1,0	-,405**	0,1	0,0
Growth Deviation	-,405**	1,0	0,0	0,0
Historic CAGR	0,1	0,0	1,0	,301**
Forecast CAGR	0,0	0,0	,301**	1,0

* Significant at the p = 0.05 level. ** Significant at the p < 0.01 level

As evident, there are clear correlations between historic volatility and the growth deviation of analysts. With a correlation coefficient of-0,4 ($p < 0.01$), it is shown that the more volatile the historic revenue has been, the more conservatively do analysts estimate future revenues – and expectedly so. As trends are harder to identify, analysts seem to act with a principle of caution in their forecast estimates. Further, historic revenue growth is positively correlated with forecasted growth, with a correlation coefficient of 0,3 ($p < 0.01$) and a R^2 of 0,318, indicating that analysts are indeed looking to historic revenue performance when forecasting. Looking at the plot below shows that most of the forecasted revenue CAGR's are somewhere between 0-5%, with a few outliers with very large CAGR rates.

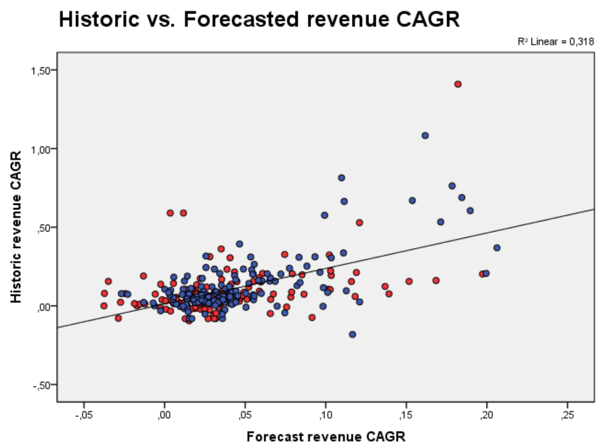


Figure 9: Historic to forecasted revenue CAGR

To investigate whether revenue is of importance in terms of target price accuracy, we conduct a Pearson's correlation test against our accuracy variables. The correlation results are shown below in Table 16 below.

Table 16: Revenue correlations with accuracy variables

	TPA			TPFE			STPFE			EIR		
Pearson's Correlation	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3
RSD Historic Revenue	0,03	0,01	0,01	-0,07	-0,06	-0,07	0,10	,163**	,190**	-0,05	-0,05	-0,10
Growth Deviation	0,01	-0,01	-0,05	0,01	0,08	0,09	-0,06	-0,08	-0,10	-0,02	0,04	0,10
Historic CAGR	0,07	0,01	-0,01	0,01	0,07	0,12	-0,02	0,09	0,02	-0,01	0,04	0,01
Forecast CAGR	-0,06	-0,07	-0,06	,155*	0,12	,132*	0,02	0,01	-0,02	,140*	0,12	0,09

* Significant at the p = 0.05 level. ** Significant at the p < 0.01 level

As evident, forecasted CAGR is positively correlated with year1 and year3 TPFE, indicating that increased CAGR in revenue forecasts yield more aggressive target price estimates – which is not surprising, as a larger revenue will, all else equal, yield higher target prices, if other valuation input levels are held constant. What is surprising is that forecasted revenue CAGR is not correlated with STPFE, which seems completely random when plotting against revenue forecast CAGR. This indicates that TPFE is only correlated because the errors go from conservative estimates with low CAGR to aggressive estimates with high CAGR. A larger CAGR could therefore be justified, as the absolute forecast errors do not necessarily increase.

Further, the relative standard deviation of historic revenue is positively correlated with STPFE at all years, with year 2 and 3 being statistically significant at $p<0.01$. This indicates that it gets harder to accurately forecast target prices when the company has had volatile revenues historically. Intuitively, this makes sense, as increased volatility makes it difficult to spot trends and thereby determine the correct direction the company is headed. Surprisingly however – and perhaps a bit reassuring – is that increased volatility leads decreases in the level of positive bias, as seen by the negative correlation with EIR. This indicates that while analysts generally find it more difficult to forecast companies with higher sales volatility, they tend to value them more conservatively. These correlations are, however, not statistically significant.

Forecast CAGR is, not surprisingly, positively correlated with excess implied returns, meaning that your target prices are more positively biased as forecast CAGR increases. What might be surprising is that the correlation lowers with year 2 and 3 EIR, indicating that forecast CAGR's are generally at a reasonable level in our sample, as the market seems to adjust prices accordingly, and thus lower the effect of CAGR on positive bias.

To test the possible impacts of the abovementioned variables, two hypotheses were formulated:

> *$H6_0$: Historic revenue volatility does not influence analysts' accuracy*
>
> *$H6_a$: Historic revenue volatility does influence analysts' accuracy*

Based on the above results, we reject the null hypothesis. Historic revenue volatility does influence the accuracy of analysts. We found that, while the TPA variable was largely unaffected, larger volatility increased standardized target price forecast errors.

> *$H7_0$: The deviation of forecasted revenue growth to historical growth does not influence analysts' accuracy*
>
> *$H7_a$: The deviation of forecasted revenue growth to historical growth does influence analysts' accuracy*

We found no statistically significant correlations between growth deviations and accuracy, and we therefore fail to reject the null hypothesis. The analysis could indicate that our analysts in general are providing quite sober revenue forecasts, as also indicated by the decrease in EIR and forecast CAGR correlation in year two and three.

NOPAT-margins

The next variable of interest is the NOPAT-margins. As shown in the introduction of this section, the NOPAT-margin has, on average, been forecasted to increase slightly, with somewhat stabilizing margins occurring in year four and onwards. To measure the impact of NOPAT-margins, we computed four variables, as outlined in section 4.4.1. These are the relative standard deviation of NOPAT-margin, the historic and forecast average NOPAT-margin, and the deviation between the two on a per-case basis. When looking at the correlation of our four variables, the following become apparent:

Table 17: Correlations between NOPAT-margin variables

Pearson's Correlation	RSD Historic NOPAT-margin	NOPAT-Margin Deviation	Historic average NOPAT-margin	Forecast average NOPAT-margin
RSD Historic NOPAT-margin	1	,307**	0,0	-0,1
NOPAT-Margin Deviation	,307**	1	0,0	-0,1
Historic average NOPAT-margin	0,0	0,0	1	,319**
Forecast average NOPAT-margin	-0,1	-0,1	,319**	1

* Significant at the p = 0.05 level. ** Significant at the p < 0.01 level

Interestingly, historic NOPAT-margin volatility and NOPAT-deviation is positively correlated, meaning that increases in NOPAT volatility historically leads to a larger NOPAT-margin in forecasts compared to historical levels. This is quite interesting, as we found the opposite to be true in the case of revenues – that revenue growth in the explicit forecast period was negatively correlated with increased historical volatility. This could however be explained in that analysts use more recent margins as benchmarks for forecast levels when volatility is large, and thus, in companies with historical margin increases, the later years of the historic period are being used. If we instead test the correlation between last historical years' NOPAT-margin to the forecast average, we get a correlation coefficient of 0.556, significant at $p < 0.01$ level, indicating that analysts, in general, seem to forecast NOPAT margins closer to recent levels than to historical averages.

In order to uncover any potential impact of our NOPAT-margin variables on our accuracy measures, we have computed and presented the Pearson correlation for each variable, as shown below.

Table 18: Pearson correlation of NOPAT-margin variables and accuracy measures

	TPA			TPFE			STPFE			EIR		
Pearson's Correlation	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3
RSD Historic NOPAT-margin	-0,05	-0,02	-0,02	-0,01	-0,02	-0,03	0,02	0,00	0,00	-0,05	-0,05	-0,06
NOPAT-Margin Deviation	0,01	0,00	0,02	0,01	-0,01	0,03	0,00	-0,03	-0,05	0,03	0,02	0,13*
Historic average NOPAT-margin	0,01	0,03	0,05	0,00	-0,02	-0,05	0,01	-0,02	0,03	-0,02	-0,03	-0,01
Forecast average NOPAT-margin	-0,03	-0,04	-0,02	-0,03	-0,04	-0,06	0,08	0,08	0,09	-0,04	-0,03	-0,05

* Significant at the p = 0.05 level. ** Significant at the p < 0.01 level

As depicted in Table 18, very little correlations exist between our accuracy and bias measurements and the NOPAT variables. TPA seems slightly negatively correlated with historic NOPAT-margin volatility – as does EIR. Surprisingly, STPFE is largely unaffected by changes in NOPAT-margin volatility, and only slightly negatively correlated with NOPAT-margin deviation.

The margin deviation is slightly positively correlated with EIR in all years, with statistical significance at p = 0,05 level in year three EIR, indicating that analysts who forecast margins above historical levels exhibit more positive bias. Contrary to our expectations, the forecasted average NOPAT-margin is negatively correlated with EIR, indicating that larger margins do not create more positive bias – actually the opposite – although not statistically significant.

$H8_0$: Historic NOPAT-margin volatility does not influence analysts' accuracy

$H8_a$: Historic NOPAT-margin volatility does influence analysts' accuracy

Based on the above results, we fail to reject the null hypothesis, as there are no statistically significant correlations between relative standard deviation of historical NOPAT and the analysts' accuracy. Analysts therefore seem quite unaffected by volatility in NOPAT-margin, which is surprising, given the sensitivity to revenue volatility.

$H9_0$: The deviation of forecasted NOPAT-margin to historical NOPAT-margin does not influence analysts' accuracy

$H9_a$: The deviation of forecasted NOPAT-margin to historical NOPAT-margin does influence analysts' accuracy

As evident on the above tests, we fail to reject the null hypothesis, as there are no statistically significant correlations between the margin deviation and accuracy. In general, these findings indicate that analysts seem to forecast NOPAT margins closer to recent levels than to historical averages.

Asset turnover

As shown in section 5.2.2 the ATO is rather conservatively forecasted, with slight decreases in the later years of the forecast period. To measure the impact of asset turnover we computed four variables, as outlined in section 4.2.1 These consist of the volatility of ATO in the historic period, the historic and forecast average ATO, and the deviation between the two on a per-case basis. When looking at the correlation of our four variables, the following become apparent:

Table 19: Pearson correlation of ATO variables

Pearson's Correlation	RSD Historic ATO	ATO Deviation	Historic Average ATO	Forecast Average ATO
RSD Historic ATO	1	0,05	-0,02	-0,07
ATO Deviation	0,05	1	-0,02	0,00
Historic Average ATO	-0,02	-0,02	1	,577**
Forecast Average ATO	-0,07	0,00	,577**	1

* Significant at the p = 0.05 level. ** Significant at the p < 0.01 level

In Table 19 it is shown that more volatility in historic ATO, the more conservative forecasted ATO estimates have been made – although not significant, and with very low correlation coefficients. Not surprisingly, the historical ATO is highly positively correlated with forecasted ATO. Analysts seem generally conservative on this measure. When testing the ATO variables on our independent variables, we derive the following:

Table 20: Pearson correlation between ATO variables and accuracy variables

	TPA			TPFE			STPFE			EIR		
Pearson's Correlation	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3
RSD Historic ATO	-0,04	-0,02	-0,06	0,07	0,03	-0,03	,146*	0,11	0,13	0,09	0,05	-0,01
ATO Deviation	0,01	0,03	0,01	0,03	0,00	0,00	-0,02	0,06	0,07	0,02	0,02	0,04
Historic Average ATO	0,06	0,05	0,04	0,01	0,03	0,03	-0,01	0,00	-0,02	-0,03	-0,01	0,00
Forecast Average ATO	0,02	0,06	0,11	0,02	0,02	0,00	-0,03	-0,07	-0,09	-0,02	-0,02	-0,01

* Significant at the p = 0.05 level. ** Significant at the p < 0.01 level

As shown in the above table, the ATO volatility is negatively correlated with TPA in all years, indicating that less volatile firms are easier to predict. TPFE is positively correlated with historic ATO volatility in year one and two, showing that increased historical volatility leads to slightly more aggressive forecast estimates. We find that year one STPFE is positively correlated with ATO volatility, with a correlation coefficient of 0,146, significant at p = 0.05 level. This indicates that increased historical ATO volatility influences absolute target price forecast errors. As shown in Figure 8, the deviation of historical and forecasted ATO is rather conservative on average, thus making the impact on accuracy limited, as also found in the above test.

$H10_0$: Historic ATO volatility does not influence analysts' accuracy

$H10_a$: Historic ATO volatility does influence analysts' accuracy

Based on the correlations test, we reject our null hypothesis, as we found that historical ATO volatility is positively correlated to STPFE, with statistical significance at p = 0.05 level. This indicates that higher volatility in ATO increases forecasting errors. Further, it seems that higher RSD increases bias, with positive one-year EIR correlation of 0.09, although this finding is not significant.

$H11_0$: The deviation of forecasted ATO to historical ATO does not influence analysts' accuracy

$H11_a$: The deviation of forecasted ATO to historical ATO does influence analysts' accuracy

We fail to reject the null hypothesis as we found ATO deviation to be statistically insignificant on all accuracy and bias measures. Further, the correlation coefficients were very low, with 0.07 correlation on three-year STPFE being the highest. This is not surprising when considering the conservative nature of the average forecasts of ATO, as evident from Figure 8.

Market capitalization

The last financial variable of interest is the company market capitalization at the time of valuation. Our hypothesis is that larger firms will be easier to accurately forecast, as the information available is more plentiful, which is consistent with existing literature.

As depicted in section 4.4.1, the numeric market capitalization within our sample ranges from 10 million USD in Brøndbyernes IF to 1.034 billion USD in Apple Inc., and therefore, the market capitalization was log-transformed at a base of 10. The intuition is that a market capitalization of 10 to 20 USDm would not double the information availability, and thus not have a great impact on accuracy. With a log-transformed variable, we seek to flatten out the exponential nature of the market cap variable in our dataset.

In addition to the market capitalization variable, we also seek to uncover whether the price-earnings ratio (P/E) at the time of valuation can help explain differences in accuracy or bias, as previous literature has shown that increased analyst coverage leads to higher P/E ratios.

Table 21: Pearson correlations of logMcap and P/E ratio at T_0 to accuracy variables

	TPA			TPFE			STPFE			EIR		
Pearson's Correlation	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3
P/E Ratio at t0	-,035	-,063	-,035	,026	,047	,044	-,220**	-,215**	-,206**	,110	,113	,117
Log(mcap)	,068	,087	,072	-,310**	-,221**	-,150*	-,168**	-,112	-,122	-,252**	-,213**	-,140*

* Significant at the p = 0.05 level. ** Significant at the p < 0.01 level

In Table 21 we find that the P/E ratio is significantly negatively correlated with STPFE in years 1-3, showing that a larger P/E ratio makes it easier for analysts to avoid forecasting errors. This might be due to larger analyst coverage for high P/E stocks, which would be consistent with existing literature. The intuition is that larger analyst coverage increases investor interest, thus making shares trade higher than fundamental values. If this is indeed the case, it would explain some of the positive correlation with TPFE and EIR, indicating that higher P/E ratios at the time of valuation leads to more aggressive target prices, and more positive bias.

Looking at the logMcap variable, we find that it is negatively correlated with EIR, indicating that larger firms have less positively biased target prices published. Intuitively this makes sense, as there are, all else equal, more coverage on larger companies, and thus more consensus estimates to sanity check your own target price against, which is also consistent with existing literature. Further, we find that logMcap is highly negatively correlated with TPFE, meaning that analysts generally are publish more conservative price estimates on large companies in comparison to smaller companies. The same intuition as above seems evident, and the results are thus not surprising. Looking at absolute target price forecast errors, we find that increased logMcap also significantly decreases errors.

Another possible explanation between the large effect of logMcap on the accuracy of analysts could be that analysts find it harder to justify larger growth rates and increased value creation in larger companies. However, when plotting the forecasted revenue CAGR, the average forecasted NOPAT-margin and the average forecasted ATO levels against market capitalization, this argument seems less valid.

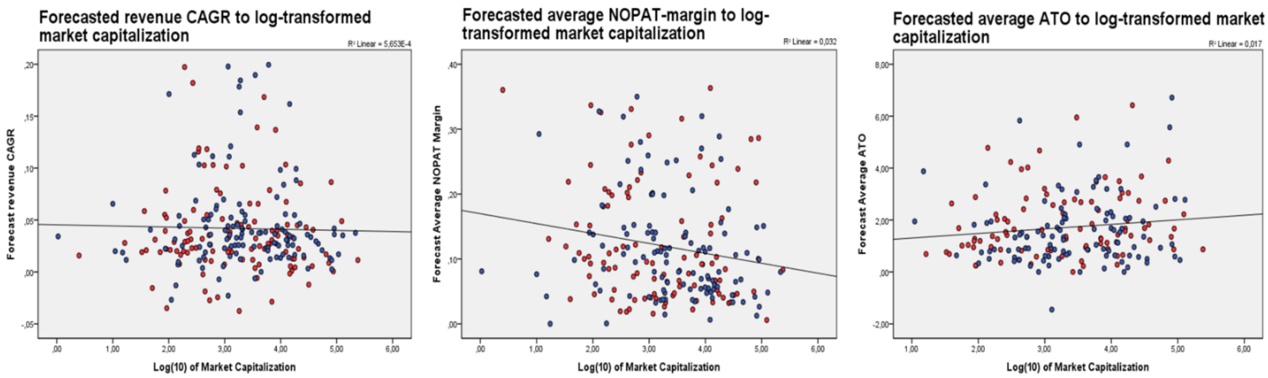


Figure 10: Forecasted revenue, NOPAT and ATO to Log-transformed market capitalization

In Figure 10 we see, quite interestingly, that the revenue growth forecasts seem rather random when plotting against market capitalization, and thus there is no apparent connection. The forecasted NOPAT-margin is however negatively correlated, with a correlation coefficient of -0,17 (p = 0,01) with market cap, showing that smaller firms are forecasted at higher NOPAT margins, possibly explaining some of the reason behind the overprediction as evident from TPFE and EIR. ATO is however increasing with logMcap.

$H12_0$: The company's market capitalization at T_0 does not influence analysts' accuracy

$H12_a$: The company's market capitalization at T_0 does influence analysts' accuracy

In conclusion, we reject our null hypothesis, as there are clear and significant negative correlation between forecast errors and market cap, indicating that the size of the firm is relevant. This is especially true for TPFE and EIR, showing that larger firms are less likely to have overpredicted target prices published – however, on STPFE the effect is less noticeable, and only significant in one-year STPFE. These findings are, as mentioned, consistent with existing literature on the topic. We also found indications that higher P/E ratios lead to lower forecast errors, and more conservative and aggressive forecasts.

5.2.3 Methodological variables

As previously argued, we have listed a set of theorized explanatory variables. In previous sections, the categories of analyst characteristics and financial factors were covered. In this section, we will test whether the methodological choices made by analysts have an impact on target price accuracy and bias, seeking to answer the hypotheses outlined in section 3.3.

Multiples

The first variable is regarding multiple usage in analysts' valuations. As depicted in section 2.4, relative valuation using multiples is commonly used in practice. In our thesis, we have limited the valuation model to primarily focus on DCF analyses – however, a lot of analysts in our sample also utilize multiples for peer comparison – in many cases as a sanity check post valuation, but we also hypothesize that the use of multiples may help shape the assumptions used by the analyst, and thus interfere with the valuations and target prices proposed. We will therefore seek to test whether the amount of multiples used has an impact on accuracy, and if some multiples allow for more accurate estimates than others.

Table 22: The number of multiples used and its effect on accuracy

		TPA			TPFE			STPFE			EIR		
No. Multiples	n	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3
0	121	0,51	0,63	0,71	0,21	0,09	0,00	0,42	0,50	0,54	0,30	0,18	0,03
1	29	0,45	0,66	0,72	0,32	0,21	0,08	0,43	0,47	0,48	0,46	0,34	0,22
2	74	0,49	0,59	0,69	0,32	0,17	0,07	0,47	0,52	0,56	0,52	0,36	0,23
3	53	0,57	0,66	0,68	0,26	0,08	-0,17	0,39	0,49	0,73	0,31	0,14	-0,18
4	31	0,48	0,65	0,71	0,25	0,10	-0,01	0,43	0,51	0,59	0,34	0,19	0,01
5+	13	0,67	0,89	0,89	0,49	0,36	0,26	0,49	0,62	0,78	0,57	0,44	0,23

* Significant at the p = 0.05 level. ** Significant at the p < 0.01 level

In Table 22, the number of multiples used has been linked to accuracy. As evident, the majority of analysts in our sample has used no multiples, and thus relied entirely on intrinsic valuation. For analysts using multiples, two multiples are the most common amount. As evident, multiple usage is difficult to link with accuracy – TPA is rather random on the amount of multiples used. In terms of TPFE and EIR, we find the lowest values for those using no multiples, yet no significance is found. Through a one-way ANOVA test, we find no significant differences in accuracy or bias between groups of number of multiples used. If we instead look to the actual multiples used, we derive the following.

Table 23: Regression coefficients of multiple usage on accuracy

		TPA			TPFE			STPFE			EIR		
		Exponentiated Coefficients			*Beta Coefficients*			*Beta Coefficients*			*Beta Coefficients*		
OLS Regression	n	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3
Constant		2,14	2,23	1,06	0,31**	0,18**	0,03	0,44**	0,55**	0,63**	0,47**	0,32**	1,17
EV/REVENUE	87	0,64	0,68	0,69	-0,11	-0,18	-0,20	-0,06	0,03	0,09	-0,13	-0,15	-0,24
EV/EBITDA	153	1,36	1,38	1,36	-0,07	-0,10	-0,06	-0,03	0,01	-0,02	-0,14	-0,12	0,07
EV/EBIT	89	1,01	1,30	1,13	0,11	0,29*	0,34*	0,03	-0,09	-0,14	0,13	0,26*	0,36*
EV/NOPAT	12	0,66	0,66	0,87	0,03	-0,06	-0,13	-0,10	-0,14	-0,07	0,01	-0,82	-0,16
P/E	103	1,08	1,17	1,11	0,03	-0,12	-0,01	0,02	-0,08	-0,10	0,04	0,02	0,02
P/BVE	59	0,89	0,89	0,89	-0,24	0,05	0,04	-0,02	0,03	0,08	-0,03	0,04	0,07
R^2		*0,018*	*0,022*	*0,013*	*0,023*	*0,047*	*0,032*	*0,02*	*0,021*	*0,015*	*0,018*	*0,021*	*0,023*

* Significant at the p = 0.05 level. ** Significant at the p < 0.01 level

As the multiples can be used interchangeably by analysts, and with some using more than one multiple, we opted for an OLS regression analysis, with binary logistic regression, and thus exponentiated coefficients (odds ratios) for the binary TPA measures, and linear regression for the remaining numeric measures. From Table 23 it is evident that, when looking at the usage of individual multiples, EV/Income Statement multiples are most frequently used, with EV/EBITDA being the most common, closely followed by price-earnings.

As evident from TPA, the odds of hitting the target price are larger for analysts using EV/EBTIDA, EV/EBIT and P/E ratios, yet without statistical significance, and with a very small R^2 of approximately 2% for all TPA years. The odds of hitting the target price are notably low on EV/REVENUE and EV/NOPAT multiples, yet not statistically significant. The large difference in odds on EV/NOPAT and EV/EBIT is interesting, as the two multiples only differ on tax assumptions – the explanation should therefore be found in the low sample size of EV/NOPAT multiples, indicating that the two groups are not identical.

As seen, TPFE is significant in year two and three on EV/EBIT, with beta coefficients of 0,29 and 0,34 respectively. This indicates that analysts using EV/EBIT multiples are prone to issue more aggressive target price forecast, but as shown in the STPFE coefficients, not necessarily have larger absolute forecast errors. Looking at EIR, we also find significance in EV/EBIT multiple usage, leading to more positive bias – although, again, with a very low degree of explanation in R^2 at approximately 2%.

If we look beside the multiples and rather focus on the regression intercepts, there seem to be significance on all independent variables except TPA. To investigate this further, we split our sample into a new dummy-variable, coded by grouping those analysts who did used multiples against those who did not. To compare these groups, an independent T-test was performed, with outputs seen below.

Table 24: T-test results of analyst accuracy grouped by multiple usage.

		TPA			TPFE			STPFE			EIR		
Multiple use	n	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3
Has used multiples	198	0,55	0,67	0,70	0,31	0,18	0,03	0,42	0,52	0,64	0,48	0,36	0,19
Has not used multiples	123	0,51	0,63	0,71	0,21	0,09	0,00	0,42	0,50	0,54	0,30	0,18	0,03
Difference		0,03	0,04	-0,01	0,10	0,09	0,03	0,00	0,01	0,10	0,18*	0,18*	0,16

* Significant at the p = 0.05 level. ** Significant at the p < 0.01 level

As evident in Table 24, when dummy-coding the multiple variable into whether the analyst has used or has not used multiples, we derive some interesting findings. There seem to be little difference on TPA, but TPFE is slightly lower for those who have not used multiples, which is coherent with the findings from Table 22. We find that the number of multiples does not aid in target price accuracy, but rather the opposite - that analysts using multiples along with the DCF model is more prone to overshooting their target price estimates.

On both one-year and two-year EIR, we find statistically significant differences between groups at the $p < 0.01$ level, with analysts using multiples having larger positive bias, with a mean deviation of 18% for both years. This is quite an interesting finding, as using multiples is often considered as a sanity check of the DCF valuation against peers. This begs the question whether relative valuation using multiples and intrinsic valuation using DCF is essentially useful to combine, or whether they should be kept separate. One possible explanation is that the analysts might not have used the multiples as sanity checks, but rather used them to adjust their forecasting assumptions, and thus deviate from coherency in forecasts. Further, we do not know whether the analysts in our sample have used appropriate peers, with similarity in growth, risk and margins, but we can conclude that multiple usage along with the DCF creates more positive bias in target price estimates than the use of DCF alone.

$H13_0$: The amount of multiples used in valuation does not influence accuracy

$H13_a$: The amount of multiples used in valuation does influence accuracy

Through the above tests, we fail to reject the null hypothesis. While we did see some variance in accuracy, we failed to find significant differences between groups in our ANOVA test. We did however find significance in our dummy-coded variable, showing that the use of multiples, quite surprisingly, create more positively biased target price estimates than if multiples were not used. Therefore, we cannot create a direct link between the number of multiples used and accuracy, but merely a link between the act of using multiples altogether.

Other models than DCF

The next explanatory variable of interest is whether analysts have used other valuation models than the DCF. As stated previously, all analyzed valuations have, as a minimum, used the DCF model. However, some analysts have utilized multiple models. In theory, all present value models should be internally consistent, and thus yield the same value estimates, given equal assumptions – however, the use of multiple models might require assumptions to be put under larger scrutiny, and thus yield more accurate results. We will test this using OLS regression on our independent variables, as shown below.

Table 25: OLS regression results on the use of different valuation models and accuracy.

		TPA			TPFE			STPFE			EIR		
		Exponentiated Coefficients			*Beta Coefficients*			*Beta Coefficients*			*Beta Coefficients*		
OLS Regression	n	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3
Constant		1,19	1,61	2,87	0,13	-0,12	-0,50*	0,46*	0,69*	1,02**	0,39*	0,12	-0,26
Discounted cash flow	196	0,54	0,81	0,81	0,17	0,29*	0,55**	-0,02	-0,14	-0,39**	0,09	0,23	0,45*
Economic value added	120	1,09	1,08	1,02	-0,09	-0,09	-0,06	-0,06	-0,09	-0,08	-0,16	-0,12	-0,12
Residual income	53	0,96	1,17	0,82	0,12	0,24*	0,22	-0,01	-0,14	-0,20	0,03	0,09	0,06
R^2		*0,009*	*0,002*	*0,002*	*0,03*	*0,034*	*0,033*	*0,011*	*0,022*	*0,037*	*0,014*	*0,012*	*0,017*

* Significant at the p = 0.05 level. ** Significant at the p < 0.01 level

As evident from Table 25, the model choice is non-significant in determining accuracy on our TPA measure. The coefficients indicate that the utilization of a DCF model lowers the odds of hitting the target price in all years, while the utilization of the EVA model increases the odds slightly. These coefficients are however non-significant and accompanied by a very low R^2 for each year.

The use of the DCF model and RI model is seen to greatly increase TPFE, with two and three-year TPFE being significant on DCF, and two-year TPFE being significant on RI. This indicates that analysts using these models publish more aggressive target price forecasts. These results are however accompanied by a very low R^2, meaning that, while the effect of using DCF and RI models are significant on TPFE, these do not explain a lot of the variance in TPFE, and thus other variables are of importance as well. In terms of STPFE however, the effect of DCF model usage is statistically significantly on STPFE in year 3, with negative coefficients in all three years. This indicates, together with the TPFE results, that analysts using DCF are more prone to issue higher target prices (as TPFE is increasing with DCF use), but they are not necessarily having larger forecast errors. Remember that TPFE is a relative variable, becoming negative in conservative scenarios and positive in aggressive scenarios.

For EIR, analysts using DCF seem to be a slight bit more positively biased, whereas EVA users are negatively biased – although none are significant. Overall, the RI analysts seem to be most consistent with market values, with somewhat average odds of hitting TPA, low STPFE and very low EIR influence. As evident, we derive statistically significant intercepts in our regressions, with similar explanations as those from Table 23. Therefore, we group the number of models used into those solely using one model, and those who have used multiple valuation models. From our sample, 111 analysts used two models, and only three analysts used three models, so a more detailed grouping on the amount of models used was not applicable. The results are compared in an independent T-test, with results presented below.

Table 26: T-test results on accuracy between groups using one or more than one model.

		TPA			TPFE			STPFE			EIR		
Number of models	n	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3
One	129	0,56	0,66	0,73	0,28	0,21	0,10	0,37	0,41	0,50	0,44	0,37	0,26
More than one	114	0,51	0,66	0,72	0,31	0,15	-0,01	0,44	0,54	0,67	0,48	0,32	0,15
Difference		0,05	0,00	0,01	-0,03	0,05	0,11	-0,07*	-0,13**	-0,16**	-0,03	0,05	0,11

* Significant at the p = 0.05 level. ** Significant at the p < 0.01 level

Interestingly, we find that the analysts using only one model seem to be slightly better in hitting target prices on TPA level, but also more prone to publish more aggressive estimates in TPFE in year 2 and 3. On STPFE results, we find that using just one model is superior in all years, with less standardized forecast errors, significant at p = 0.05 level in year 1 and p = 0.01 level in year two and three. This shows that analysts generate more inaccurate estimates when using multiple valuation models. This is interesting, as it is rather contrary to the concept of internal consistency between models, namely that all present value models should give equal values, given equal assumptions. Perhaps analysts use various models as a sanity check of value drivers and assumptions, as the inputs of a DCF varies to the inputs of a residual income model, which includes the book-value of equity. In terms of analyst bias, there was not found any significant differences between groups.

$H14_0$: Analysts' model choice does not influence accuracy

$H14_a$: Analysts' model choice does influence accuracy

To conclude on the above findings, we reject our null hypotheses, as analysts' model choice does influence accuracy. We found that analysts using DCF-models alone were more prone to publish aggressive and biased estimates and had larger standardized target price forecast errors in all years further, we found that using multiple valuation models yielded significantly more accurate target prices measured on STPFE in all years, and as such, both the choice of model and the number of models used has influence on accuracy.

Forecast horizon

Another relevant variable in terms of methodology in valuations is the length of the explicit forecast period. As evident from previous literature, there is a lot of contradicting debates in regard to forecast periods – naturally, any event gets harder to predict with increases in time. However, for short forecast periods, you are heavily reliant on your terminal value assumptions, as the terminal value makes up a larger part of your estimated target price. For this reason, we seek to test the effect of forecast horizon lengths. To do this, we have grouped the estimates based on this length, with a cap after 10 years of forecast. Many analysts forecast more than 10 years in our sample, with one analyst forecasting 69 years into the future – however, as the frequency of analysts using more than 10 years of forecast horizon is very limited, and very sporadic, we decided to group analysts above 10 years into the final category, with a combined count of 30 valuations. The results from the ANOVA test are presented below.

Table 27: One-way ANOVA test on differences in accuracy by forecast horizon

		TPA			TPFE			STPFE			EIR		
Years in Forecast	n	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3
3	4	0,75	0,75	0,75	0,26	-0,20	-0,49	0,29	0,59	0,90	0,53	0,08	-0,22
4	18	0,61	0,78	0,83	0,28	-0,05	-0,11	0,37	0,44	0,47	0,25	0,10	0,04
5	110	0,61	0,72	0,75	0,29	0,19	0,11	0,41	0,52	0,57	0,40	0,29	0,13
6	31	0,55	0,61	0,65	0,25	0,23	0,13	0,41	0,47	0,59	0,41	0,40	0,33
7	21	0,48	0,71	0,76	0,06	-0,18	-0,44	0,34	0,65	0,91	0,15	-0,05	-0,31
8	17	0,65	0,71	0,82	0,18	0,11	-0,15	0,38	0,45	0,61	0,34	0,27	0,01
9	22	0,41	0,50	0,64	0,47	0,33	0,16	0,56	0,58	0,75	0,69	0,54	0,39
10	65	0,48	0,60	0,66	0,29	0,19	0,01	0,43	0,51	0,61	0,46	0,35	0,14
>10	30	0,40	0,57	0,63	0,24	0,12	-0,08	0,45	0,47	0,54	0,47	0,31	0,12

* Significant at the p = 0.05 level. ** Significant at the p < 0.01 level

As shown, the most common length of forecast is five years, closely followed by 10 years. As evident from the above, there seems to be no systematic differences in the accuracy based on number of years. When testing the groups with Turkey Post-Hoc tests, no one group stood out in comparison to the others, indicating insignificant differences, which seems to contradict prior research – namely that target price accuracy decreases with time. On the TPA variable, the means indicate that target prices are hit more often when analysts forecast 3-5 years in comparison to longer periods, but this is not significant.

$H15_0$: The length of forecast period does not influence accuracy

$H15_a$: The length of forecast period does influence accuracy

Based on the above findings, we fail to reject our null hypothesis. Through our ANOVA analysis, we found no significant differences in the length of forecast period on neither accuracy nor bias measures.

Historic period

Having tested the forecast horizons influence on target price accuracy, the historic period is naturally also of interest. There is a limited amount of literature on the impact of the length of the historical period in terms of valuation accuracy. Intuitively, one could argue that a longer historical period would yield more accurate estimates, as the analyst has a larger information base and is better able to spot trends in value drivers. We therefore seek to uncover whether the length of the historical period does indeed matter in terms of accuracy.

As with the length of the forecast period, a one-way ANOVA was performed on our accuracy and bias variables, grouping the length of the historical period from three years as the minimum used in our sample, to eight years as the maximum. The results are presented in the table below.

Table 28: ANOVA table of historical period to accuracy measures

		TPA			TPFE			STPFE			EIR		
Years in history	n	Year 1	Year 2	Year 3	Year 1**	Year 2	Year 3*	Year 1	Year 2	Year 3**	Year 1*	Year 2	Year 3*
3	34	0,53	0,68	0,71	0,35	-0,04	-0,47	0,49	0,62	0,92	0,58	0,17	-0,47
4	30	0,67	0,73	0,77	0,22	0,06	-0,22	0,49	0,65	0,84	0,52	0,36	0,08
5	130	0,50	0,65	0,70	0,33	0,24	0,14	0,41	0,46	0,48	0,47	0,38	0,26
6	65	0,52	0,62	0,68	0,32	0,18	0,06	0,40	0,48	0,60	0,46	0,37	0,25
7	27	0,63	0,74	0,81	0,01	0,04	0,12	0,38	0,53	0,59	0,12	0,16	0,20
8	35	0,54	0,60	0,66	0,18	0,04	-0,15	0,41	0,55	0,69	0,15	0,01	-0,18

* Significant at the p = 0.05 level. ** Significant at the p < 0.01 level

As evident in Table 28, the most common number of years in the historical period is five years. The results in TPA seems rather random, with four years and seven years being the most accurate in terms of hitting target prices, yet without statistically significant differences between groups. One and three-year TPFE have significant differences between groups, where seven-year history is by far creating the least aggressive target price forecasts on TPFE. On STPFE however, there seem to be no significant differences between groups, except in year three, where the three and four-year groups are vastly more aggressive. On EIR, we find significant differences on one-year and three-year EIR, with the three-year historical group being the most positively biased. On both STPFE year one and EIR year one, we find a somewhat linear relationship between the number of historic years and the degree of forecast errors and bias, indicating that more years in history leads to more accurate target price estimates.

In order to explore the abovementioned differences, a Turkey Post-Hoc test is performed, which reveals that one-year TPFE is significantly different on analysts using 7 years of history compared to those using 5 or 6 years of history, with mean differences of -0,32 and 0,31 respectively, and p=0,01 and p=0,03. Contrary to these findings, the differences within-groups of "three-year TPFE" shows that analysts using 3 years of history has much lower TPFE than those using 5 years, with a mean difference of -0,6 and p=0,03.

On three-year STPFE however, both analysts using 3 years and 4 years of history have higher standardized errors than those using 5 years, with mean differences of 0,43 and 0,36 respectively, significant at p=0,02 and p=0,04. Finally, three-year EIR shows much lower EIR scores for analysts using 3 years of history than those using 5 years of history, with mean differences of -0,73, significant at p = 0,03.

While the above results are significant on specific years, it is difficult to say for certain whether increases in historic period length does increase accuracy. The results indicate that, on TPFE, 7 years of history is vastly superior, indicating that less aggressive forecast estimates are published when considering longer periods of historical financials – however, on STPFE, the differences are less pronounced, showing that analysts still create somewhat similar absolute errors. To see whether we can explain a linear relationship between length of historic period and accuracy, we run a Pearson correlation test on our historic period against accuracy measures. Here, we find that length of historic period is negatively correlated at p=0,016 level against one-year TPFE, with a correlation coefficient of -0,13. The same goes for one-year EIR, with a correlation coefficient of -0,18 at p = 0,001. This shows that the length of the historic period does matter in terms of TPFE and EIR – namely that you reduce the tendency to publish aggressive estimates and reduce the amount of positive bias with added years of history – at least to the seven-year mark, as evident from Table 28. These results are to some degree not surprising – as argued previously, more historical data would, all else equal, be a better basis for formulating coherent value driver assumptions.

$H16_0$: The length of historical period does not influence accuracy

$H16_a$: The length of historical period does influence accuracy

Based on the above tests, we reject the null hypothesis. The length of the historical period does matter. We found that both the aggressiveness and bias of target price estimates decrease with added historical years. Further, we found that three-year standardized forecast years were significantly lower when using five years of historical financials in comparison to only using three or four years.

Steady state assumptions

When scouting the literature on proper DCF assumptions, one frequent assumption is that of steady state. In order for the terminal value to be computed on sound financial metrics, the company must have reached a steady state level of operations, as argued by Damodaran. While much debate exists on the assumptions needed for a company to be in steady state, we have computed three variables based on some of these assumptions, as outlined in section 4.4.1. The frequency of steady state satisfaction based on our binary (assumption met, yes/no) variables is shown below.

Table 29: Frequency of satisfaction of steady state assumptions

	Satisfied	Not Satisfied	*% Satisfied*
SS_1	180	141	*56,1%*
SS_2	65	256	*20,2%*
SS_3	146	175	*45,5%*
All SS	34	287	*10,6%*

Where SS_1 checks whether the revenue growth of the last forecast year is equal to terminal value growth (g), SS_2 checks whether the revenue growth in the second last forecast year is equal to terminal value growth (g) and SS_3 checks whether the ROIC margin in the terminal year is stable when compared to the last forecast year. As evident, the majority of analysts do satisfy the first criteria of stable revenue growth between TY and TY_{-1}. Only 20% of analysts ensure two consecutive years of stable revenue growth prior to the terminal year. Lastly, slightly less than half of the analysts satisfies the stable ROIC assumption between TY and TY_{-1}.

When testing the effect of steady state assumptions being satisfied through an independent T-test, we found no significant differences on neither of our three steady state criteria on accuracy measures. If we instead of looking at steady state assumptions at a binary (satisfied/not satisfied) level, but instead look at deviations between terminal year and last year of forecast on a DuPont decomposed level, we derive the following.

Table 30: Deviations on terminal year value drivers vs last year of forecast

	TPA			TPFE			STPFE			EIR		
Pearson Correlation	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3
Growth deviation, TY vs TY_{-1}	-0,08	-0,04	0,05	0,07	0,00	-0,09	-0,04	-0,10	0,02	0,11*	0,05	-0,04
NOPAT-Margin deviation, TY vs TY_{-1}	0,02	0,09	0,07	0,02	0,04	0,03	0,05	-0,02	-0,03	0,07	0,07	0,07
ATO deviation, TY vs TY_{-1}	0,02	0,08	0,07	-0,05	-0,06	-0,03	0,00	0,05	0,03	-0,04	-0,06	-0,05
ROIC deviation, TY vs TY_{-1}	-0,07	-0,03	-0,02	0,11*	0,12*	0,10	0,17**	0,08	0,06	0,13*	0,14*	0,13*

* Significant at the p = 0.05 level. ** Significant at the p < 0.01 level
TY = terminal year

As evident from Table 30, NOPAT-margin deviation and ATO deviation alone have no significant correlation with neither of our accuracy measures. When combining them to ROIC deviation, we do however find significant effects on both TPFE, STPFE and EIR.

On TPFE we find significant positive correlations in year one and two, indicating that analysts' target price estimates become more aggressive when ROIC in the terminal year becomes larger than in the year before. Intuitively this makes sense, as a larger ROIC in terminal year will, all else equal, give higher estimates, as either NOPAT margins are higher, revenue is higher or invested capital is lower. However, positive correlations are also found for STPFE in year 1, indicating that absolute errors are likewise larger, thus concluding that deviations in NOPAT-margin and ATO in isolation are not great predictors of forecast accuracy, but when combined into ROIC, there is clear positive correlations, which is consistent with existing literature.

Lastly, we find that EIR is positively correlated with the deviation in ROIC in all three years, indicating that larger ROIC deviations are a source of additional positive bias. The same is true for revenue growth in one-year EIR. These findings are however not surprising, as algebraically the intuition is clear – a larger ROIC will inevitably lead to higher forecast estimates. However, when comparing to the findings in .

Table 29, it is quite interesting that merely half of the analysts in our sample satisfies the steady state assumption of having stable ROIC levels between the terminal year and last forecast year – especially given the significant positive correlations.

$H17_0$: Satisfying steady state assumptions does not influences accuracy

H17a: Satisfying steady state assumptions influences accuracy

Based on the above analysis, we fail to reject the null hypothesis, as our simple, binary variable, showed no significant effects on target price accuracy. However, when extending the analysis into ROIC, and the levers of ROIC, we found that larger ROIC in the terminal year compared to the year before does increase absolute forecast errors, the level of aggressiveness in estimates and the amount of positive bias. Thus, although our null hypothesis is, strictly speaking, rejected, we still found steady state to be of importance in determining accuracy and bias.

Growth sanity check

Another key element of the DCF model assumptions is that of the constant growth rate, g. Damodaran (2002) propose a simple sanity check that he coins "obeying the growth cap", namely that the terminal growth rate cannot exceed the risk-free rate, as the risk-free rate is a simple proxy for the nominal growth rate of the economy. If terminal growth is above the risk-free rate, you assume, in a nutshell, that the company will outgrow the economy in perpetuity.

To test the effect of this assumption on target price accuracy, we computed a simple binary variable on g < rf, with the following results:

Table 31: T-test of growth sanity check on accuracy measures

		TPA			TPFE			STPFE			EIR		
Growth Sanity Check	n	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3
Passed	121	0,51	0,64	0,72	0,33	0,22	0,06	0,43	0,47	0,59	0,51	0,39	0,23
Not Passed	122	0,56	0,67	0,73	0,27	0,14	0,02	0,39	0,50	0,59	0,41	0,29	0,17
Difference		-0,04	-0,03	-0,01	0,06	0,08	0,04	0,03	-0,03	0,00	0,10	0,11	0,07

* Significant at the p = 0.05 level. ** Significant at the p < 0.01 level

As evident from the above T-test results, there is very little difference between those who pass the growth sanity check, namely that the terminal year growth cannot be higher than the risk-free rate, and those who does not. Surprisingly however are the proportions, where half have passed, and half have not passed. This indicates that the sanity check, as proposed by Damodaran, is either not very known among analysts, or not very respected. When analyzing the deviation between risk-free rates and terminal growth rates, we derive the following:

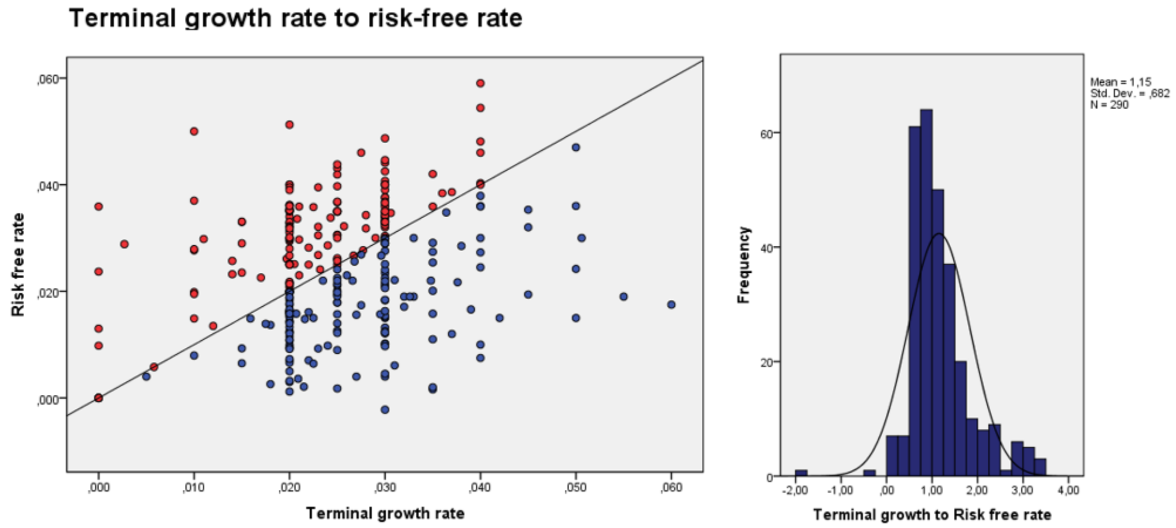


Figure 11: Scatter plot and histogram of terminal growth rate to risk-free rate

As shown above, the distribution of terminal growth rates seems rather random, with high frequencies around even numbers (2%, 3% and 4%). From the histogram however, we see that, while the majority of analysts have used a terminal value close to the risk-free rate, there is a rather significant proportion of analysts in the right tail, with terminal growth rates exceeding risk-free rates by a factor of 1.5x (80th percentile) and above. If we instead of testing the passing or failing of the growth sanity check, instead focus on the outer tail we derive the following:

Table 32: T-test of outer tail deviations between terminal growth rate and risk-free rate

		TPA			TPFE			STPFE			EIR		
TV_g to rf	n	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3
TV_g > 1.5x rf	48	0,48	0,56	0,60	0,26	0,03	-0,16	0,45	0,59	0,79	0,47	0,24	0,05
TV_g < 1.5x rf	176	0,53	0,67	0,74	0,33	0,25	0,12	0,41	0,44	0,54	0,48	0,41	0,27
Difference		-0,05	-0,11	-0,14*	-0,07	-0,22*	-0,28*	0,04*	0,15**	0,26**	-0,01	-0,16	-0,22

* Significant at the p = 0.05 level. ** Significant at the p < 0.01 level

In Table 32 we find large differences in accuracies when grouping terminal value growth on the outer tail. Analysts with terminal value growth in excess 1.5x the risk-free rates are, not surprisingly, less prone to hitting their target prices, with three-year TPA being statistically significant at the p = 0,05 level. On TPFE however, both two-year and three-year TPFE are significantly smaller for this group, indicating that a terminal year growth rate in excess of 1.5x the risk-free rate produces less aggressive target price estimates. These findings are rather surprising, as the intuition would be quite the opposite – that larger growth rates in the terminal year would, all else equal, lead to more aggressive estimates. On STPFE, we find that, although the estimates may be more conservative, they come with significantly larger forecast errors. These findings indicate that while the growth sanity check in itself might not be of significant importance, the outer tails, and thus the most extreme cases of terminal growth deviation do have an effect on forecast errors and estimation aggressiveness.

$H18_0$: Satisfying the growth sanity check does not influences accuracy

$H18_a$: Satisfying the growth sanity check does influences accuracy

Based on the above analysis, we fail to reject our null hypothesis. There were no statistically significant differences in accuracy or bias between groups on the growth sanity check variable. However, we found that grouping analysts on outer tails provided statistically significant results. Surprisingly, the analysts with the largest terminal growth deviations had the most conservative estimates in terms of TPFE, but also the largest absolute forecast errors. While the sanity check is not a perfect "rule of thumb" measure, it indicates the importance of careful consideration when determining terminal growth rates in comparison to the risk-free rate.

Change in WACC

As previously argued in section 2.4, the literature suggests that analysts should consider changes in capital structure during forecasts, and in turn, the changes made to WACC from these changes in capital structure. In our sample, only 6,2% of valuations include changes in WACC during the explicit forecast period – however, only 11 analysts of 161(the sum of analysts who forecasted both NIBD and equity in forecast period) have the same level of leverage throughout the forecast period, indicating that either analysts are not aware of the implications of changing leverage, or simply do not respect this relationship. To test this impact, we grouped analysts' valuations based on whether they changed their WACC rates during the explicit forecast period or not.

Table 33: T-test of analyst groups who changed WACC during forecast

		TPA			TPFE			STPFE			EIR		
WACC	n	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3
Changed	13	0,54	0,54	0,54	0,31	0,15	0,02	0,38	0,54	0,71	0,45	0,29	0,16
Constant	230	0,53	0,67	0,73	0,30	0,18	0,04	0,41	0,48	0,58	0,46	0,34	0,20
Difference		0,00	-0,13	-0,20	0,01	-0,04	-0,02	-0,03	0,06	0,13	-0,01	-0,06	-0,04

* Significant at the p = 0.05 level. ** Significant at the p < 0.01 level

As shown above, the group of analysts with changing WACC during forecasts are less likely to hit target prices in year two and three and publish target prices with larger standardized forecast errors in year two and three. However, caution should be taken on the interpretive validity of these results, as the group is very small with a n of 13. Neither of the accuracy measures show statistically significant differences.

$H19_0$: Changing WACC during forecast does not influences accuracy

$H19_a$: Changing WACC during forecast does influences accuracy

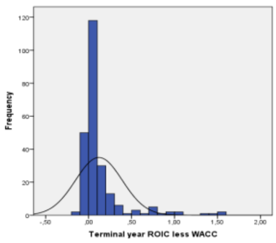
Based on the above results, we fail to reject the null hypothesis. The change of WACC was not shown to have any influence on analysts' accuracies. However, the number of analysts in our sample with changing WACC is very small, and as such the results are hard to interpret with any degree of certainty. The results do not infer that a dynamic WACC is not of importance – but merely that the analysts in our sample either does not know of the impact of leverage on WACC, which we find highly unlikely, or simply do not respect this rule of thumb.

WACC to ROIC

The last explanatory variable of interest in our analysis is the relationship between WACC and terminal year ROIC. As previously argued, having a ROIC larger than WACC in perpetuity essentially means that the analyst is believing the company to sustain competitive advantages for eternity. Contrary, a ROIC lower than WACC implies that the company will be destroying value in eternity. These concepts cement the importance of forecasting ROIC accurately – especially so in the terminal year of the DCF.

In our sample, 145 of 183 analysts have forecasted ROIC to be larger than WACC, with a mean difference of 5,47%. To find whether this difference has an impact on forecasting accuracy or potential bias, a T-test was computed, grouping differences at the zero-level, thereby deriving two groups of analysts – those who believe ROIC will be forever larger than WACC, and those believing the opposite. Those with terminal year ROIC equal to WACC (n = 13) were excluded from this test.

Table 34: T-test of ROIC and WACC groups. Pearson correlation on ROIC less WACC range



		TPA			TPFE			STPFE			EIR		
ROIC minus WACC	n	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3
> 0	145	0,50	0,63	0,72	0,28	0,14	-0,02	0,41	0,45	0,57	0,48	0,34	0,18
< 0	38	0,66	0,76	0,79	0,32	0,22	0,12	0,40	0,50	0,59	0,33	0,23	0,13
Difference		-0,16	-0,13	-0,07	-0,04	-0,08	-0,14	0,01	-0,05	-0,02	0,15	0,11	0,05

* Significant at the p = 0.05 level. ** Significant at the p < 0.01 level

Pearson Correlation	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3
ROIC minus WACC	0,01	-0,05	-0,08	-0,10	-0,05	-0,01	-0,12*	-0,11	-0,05	0,05	0,02	0,02

* Significant at the p = 0.05 level. ** Significant at the p < 0.01 level

As evident from the above table and histogram, there is a wide variety in the forecasting of ROIC. In general, the most optimistic analysts – namely those with ROIC-WACC > 0 are less likely to hit their target prices, with mean differences of 16% in one-year TPA. On TPFE, the same group is more conservative, with mean differences of 4%, 8% and 14% on one-to-three-year TPFE respectively. However, the differences in STPFE are less pronounced, indicating that, while they are more conservative in their estimates, they are not necessarily more accurate.

To test whether linear relationships exist, we computed a Pearson correlation test on the ROIC-WACC variable against our accuracy measures. In general, the variable shows negative correlation on TPA and TPFE, indicating that a larger difference between ROIC and WACC decreases accuracy and increases conservatism. Surprisingly, on one-year STPFE, we derive a negative correlation of -0,12, which is statistically significant at p = 0,05 level, showing that larger differences significantly decrease the absolute forecasting errors.

> *$H20_0$: The relationship between WACC and ROIC in terminal year does not influence accuracy*
>
> *$H20_a$: The relationship between WACC and ROIC in terminal year does influence accuracy*

As evident from the above results, we reject our null hypothesis, as we found statistically significant correlations between the relationship between WACC and ROIC and the target price accuracy. Specifically, we found that larger differences between ROIC and WACC decrease the absolute target price forecast errors.

5.3 Sub-conclusion of findings

In the following section, a brief sub-conclusion of our analysis findings will be presented. In Appendix I, a complete overview of our hypotheses is given, along with the test results from our analysis.

5.3.1 Primary research question

Our primary research question was formulated to uncover whether analysts' target prices were accurate or biased. It is difficult to establish a certain threshold indicating whether analysts are accurate or not, but with only 53,4% of analyst hitting the target price within a 12-month period, we found it reasonable to conclude that analyst are fairly inaccurate, with mean sample TPA differing significantly from zero.

In relation to TPFE and EIR, we found sample means of 27% and 48% respectively, steadily decreasing with increased horizon, due to the overshooting and undershooting cancelling each other out. TPFE was initially smaller (i.e., in the first year), which is due to target prices generally being higher than the actual share prices at the valuation date. This is confirmed by the fact that the distribution of buy/sell recommendations is skewed toward buy recommendations.

STPFE was found to be 23% in the first year and increasing with the horizon. Intuitively this finding was as expected, as it becomes inherently more difficult to forecast in the long run, because of increased uncertainty.

In relation to whether buy and sell recommendation have the same accuracy, we found that they differed on a significant level in relation to TPA, TPFE and EIR in all years, however no significance was found in relation to STPFE. We also, as expected, found buy recommendations to be positive biased and aggressive, and sell recommendations to be negatively biased and conservative. Based on the above findings, we reject the null hypothesis that buy and sell recommendations are equally accurate.

5.3.2 Secondary research question

In this section, the findings related to our secondary research question will be presented.

Analyst characteristics

In terms of gender, we found no significant differences in target price accuracy between men and women. Nonetheless we found female analysts to be more prone to publishing aggressive target prices, especially on two and three-year TPFE, where the mean female analyst showed target price forecast errors of 34% and 26% respectively, whereas the TPFE of male analysts was 15% and 0% for two and three-year TPFE. In relation to bias, female analysts were shown to exhibit more positive bias in all three years, with statistically significant differences on two and three-year EIR. This contradicts previous literature that female analysts are more conservative. In the STPFE measure, male analysts have larger absolute forecast errors in year two and three, although not statistically significant. Based on the above results, we do however fail to reject the null hypothesis that female analysts' accuracy is equal to male analysts' accuracy.

On differences between study programmes, we found that FIR analysts were most frequently hitting their target prices on all three years, closely followed by AUD and AEF analysts. Using an ANOVA test on TPFE, we found that differences among study programmes were significant. For TPFE, the AEF analysts were the most conservative. The most aggressive group of analysts was HD_AFM, with TPFE reaching 40% in year one. Both HD-programmes were above the total means on one and two-year TPFE. FIR analysts had the lowest absolute forecast errors measured on STPFE, followed by ASC and FIN analysts.

One-year STPFE also showed statistical significance between groups. FIR analysts showed the lowest absolute forecast errors measured on STPFE, followed by ASC and FIN analysts. On EIR, all groups are seen to be positively biased on average, with only the FIR, FSM and AEF analysts being negatively biased on three-year EIR. When comparing on one-year EIR measures, AEF is the least biased group of analysts, with excess implied returns of only 8%. Based on the above results, we reject our null hypothesis that analyst study programme does not influence accuracy, as we found statistically significant differences between groups on the STPFE measure.

Between individuals and groups, we found a minor difference on TPA, with groups of analysts being slightly worse in hitting their target prices in our sample. Further, the groups of analysts were shown to be slightly more conservative on their TPFE measure and had slightly smaller forecast errors on the STPFE measure. On EIR we found the largest differences – namely that groups of analysts generally were less positively biased, with mean EIR differences of 8-12% in one-to-three-year EIR. None of our tests yielded statistically significant results, and we therefore fail to reject the null hypothesis that groups and individuals have the same accuracy.

Financial factors

We found that historic revenue volatility does indeed influence the accuracy of analysts. While the TPA variable was largely unaffected, larger volatility increased standardized target price forecast errors. We therefore rejected the null hypothesis that historical revenue volatility does not influence accuracy. On the other hand, we found no statistically significant correlations between revenue growth deviations and accuracy, and we therefore fail to reject the null hypothesis that the deviation of forecasted revenue growth to historical revenue growth does not influence accuracy.

We found no statistically significant correlations between historical NOPAT volatility and the analysts' accuracy. Analysts therefore seem unaffected by volatility in NOPAT-margin, which is surprising, given the sensitivity to revenue volatility. We therefore fail to reject the null hypothesis that historical NOPAT volatility does not influence accuracy. We found no statistically significant correlations between the NOPAT-margin deviation and accuracy. In general, we found indications that analysts seem to forecast NOPAT margins closer to recent levels than to historical averages. We therefore fail to reject the null hypothesis that the deviation between historic NOPAT-margin and forecasted NOPAT-margin does not influence accuracy.

We found the correlation between historic ATO volatility to be significantly correlated with STPFE, which indicates that higher volatility in ATO increases absolute forecast errors. We therefore reject the null hypothesis that historic ATO volatility does not influence analysts' accuracy. We also found that RSD was correlated with one-year EIR, but this was not significant. Further, we failed to reject the null hypothesis that the deviation of forecasted ATO to historical ATO does not influence analysts' accuracy. The correlation coefficients were very low, with 0.07 correlation on three-year STPFE being the highest. This is not surprising when considering the conservative nature of the average forecasts as evident from Figure 8.

The correlation between market cap and forecast errors was significant, and we can therefore reject the null hypothesis, that market cap does not influence accuracy. We found that TPFE and EIR decreased with higher market cap. The relationship with STPFE was less noticeable as the correlation was only significant in the first year. The findings were consistent with the existing literature.

Methodological factors

In relation to multiples, we failed to reject our null hypothesis that the amount of multiples does not affect the accuracy, as we could not establish a significant link between the amount of multiples used and accuracy. However, we found that the mere use of multiples created significantly positively biased target price estimates.

Regarding present value models, we found significant differences between analysts' model choice and accuracy, and we therefore reject the null-hypothysis that analysts model choice does not influence accuracy. We found that using more than one valuation model yielded significantly more accurate target prices in all years. We further found indications that analysts' only using the DCF model exhibited positive bias, as the mean difference in EIR in year 3 was significant.

We found no significant relationship between the length of the forecasted period and accuracy or bias. We can therefore not reject the null hypothesis that the length of the forecast period does not influence accuracy. This finding also indicated that the relationship between the terminal value and EV, does not affect accuracy as well, as this relationship is closely related to the forecast period.

The relationship between the length of the historical period and the accuracy was found to be significant, and we could therefore reject the null hypothesis that the length of the historic period did not influence accuracy. We found that both aggressiveness and bias measures decreased with additional historical years. We further found that using five years of historical financials had significantly lower STPFE than using three or four years of historical financials.

We found no significance in relation to the steady state assumption and accuracy, we therefore could not reject the null hypothesis that the steady state assumption does not influence accuracy. However, we found that a higher ROIC in the terminal year compared to the year before, was associated with higher levels of aggressiveness, positive bias and STPFE.

In relation to WACC we could not reject the null hypothesis that a change of WACC does not influence accuracy. We could not establish a link between change of WACC and forecast accuracy. However, the number of analysts that changed WACC during the forecast was very scarce, and therefore, it was difficult to obtain any definitive answer on this subject.

Quite surprisingly, we found that higher "ROIC minus WACC" decreased the absolute forecast error, and we therefore rejected the null hypothesis that differences in ROIC and WACC does not influences accuracy.

To summarize, we found that the study programme of analysts' affected accuracy. An increase in historical revenue volatility decreased accuracy, as did historical ATO volatility. Company size in terms of market capitalization was found to be positively correlated with accuracy, as was the length of the forecast period and the analysts' model choice. Finally, we found that the relationship between WACC and ROIC in the terminal year does influence accuracy.

6.0 Discussion

In this section, we will seek to provide discussions of our findings in light of existing literature, and lastly, seek to highlight the implications of our results – both to academics and practitioners alike.

6.1 Discussion of findings

Having tested and presented the results of our analysis in section 5.0, we will now proceed to highlight interesting findings as topics for discussion in comparison to the literature outlined in section 2.0. As previously defined, we will split our discussion into the findings on analyst characteristics, financial variables and methodological variables.

As found in the results section of our thesis, analysts target prices are indeed inaccurate, which is largely consistent with existing literature. With only 54% of our sample being able to predict target prices in a one-year timespan, the results support existing literature – for instance, Imam et al. (2013) found overall target price accuracy to be 49% in their study, and Bradshaw et al. (2013) found that only 38% of target prices were met on a one-year horizon. In that regard, our sample of analysts have generally performed well in comparison to existing literature – yet still with considerable inaccuracies and forecast errors. When subsetting each accuracy measure based on the year of publication, we found that analysts are systematically inaccurate and biased, with only one of 13 years having average negative bias. This goes well in line with the findings of Lim (2001), who argued that analysts are systematically publishing positively biased target price estimates to sustain their flow of information from corporates. Similar viewpoints and conclusions regarding the association between information flow and systematic bias has been made throughout literature (Dechow, Hutton, & Sloan, 2000; Hong & Kubik, 2003; Lin & McNichols, 1998; McNichols & O'Brien, 1997). What is interesting though, is that by using students as proxies for analysts, this fear of information cut-off should be neutralized – yet we still see systematic positive bias in our results. While this does not rule out the explanation posed by previous researchers, it does leave room for further research on the causes of systematic bias.

A surprising finding from the analysis of our primary research question was the large difference in accuracy when comparing buy and sell recommendations. We found significant difference on TPA, TPFE and EIR in all years, suggesting that analysts issuing sell recommendations are more often hitting their target prices, and that they exhibit less positive bias. The issue of bias is not surprising, as it is to be expected from a sell recommendation – however, the target price accuracy results suggest that our TPA measure, as inspired by Sayed (2015), might be favoring conservative estimates, as it is negatively correlated with TPFE and EIR. Another interesting finding was that two and three-year STPFE measures were significantly negatively correlated with EIR in all years, indicating that larger target price forecasting errors leads to less positive biases on two- and three-year horizons. This shows that STPFE increases in year two and three are mostly due to stocks increasing beyond the target price from T_0 in the buy scenario and decreasing below the target price in the sell scenario.

6.1.1 Analyst characteristics

As evident from the results in section 5.2.1, the target price accuracy between genders is a split topic, with conclusions leading in both directions. We found statistically significant differences between male and female analysts in both aggressiveness of target price estimates, and the level of bias associated with those estimates, with female analysts publishing significantly more aggressive estimates than their male counterparts and having more positive bias on our EIR measure. These results are rather contradictory in comparison to existing literature. Bosquet, de Goeij, & Smedts (2014) found female analysts to be 40% less likely to offer optimistic/aggressive investment advice, which is not directly applicable to our study, as we have not investigated the proportions of aggressive or conservative estimates, but rather the mean forecast error of target prices issued. Green, Jagadeesh, & Tang (2009) found absolute forecast errors to be larger for female analysts. For the female analysts in our sample, this metric is quite the opposite – while the female analysts are found to issue more aggressive estimates, they are in fact more accurate, however not statistically significant.

One possible explanation for the deviation of our results to existing literature may be the found in the sample size constraints of this study – namely that we fail to control for other, possibly interfering variables. Bosquet. et al. (2014) used, for instance, the riskiness of companies covered and the task complexity as controlling variables in their study. Commonplace psychological research suggests that men are generally less risk averse than women (Charness & Gneezy, 2012), which could potentially be an interfering factor in our study as well – namely that the men might have chosen to cover more risky companies, and thus issue less aggressive target prices as a result of trying to mitigate these risks. These hypotheses are beyond the scope of our thesis, yet still of interest as possible explanations to our results.

As our sample is somewhat unique in nature given the information gathered regarding study programmes, naturally we also sought to test whether differences in accuracy were present across programmes. As consistent with the findings from Hope & Fang (2020), we also found significant differences between groups of different educational background. As our sample comprise of CBS students of various programmes, our results are inherently difficult to apply on a grander scale – however, on a fundamental level, the results from our tests showcase that educational profile is not insignificant in determining analyst accuracy.

Hope & Fang (2020) also studied the effects of analyst teams versus individual analysts, and found that analyst teams, in general, publish more accurate forecast estimates. The intuition here is that groups help keep each other at bay, thus lowering biases and forecast errors. Contrary to these findings, we found no significant differences between the performance of individual analysts versus their group peers. However, the reliability of this result is somewhat limited by the fact that the groups of analysts in sample consists of a maximum of two analysts, thus impacting the overall generalizability of these findings, as analyst teams might be larger in practice. Further, one could hypothesize that the team dynamics are different between practitioners and students, and as such our findings on this matter is not to be considered entirely contradictory to those of Hope & Fang (2020), because they nonetheless find team dynamics to matter.

6.1.2 Financial variables

As for the results on financial variables in section 5.2.2, we found, first and foremost, that the analysts in our sample were quite optimistic in terms of forecasting ROIC. From Figure 8, we saw quite rapid ROIC increases as the financials transitioned from being historic to being forecasted. The same became evident on NOPAT-margin level, with rather steady increases in the initial years of the explicit forecast period, after which it seemed to neutralize from year four and onwards. On the contrary, analysts seemed to lower their revenue growth expectations in comparison to historical averages. This forecasting behavior is rather interesting, as it indicates that while analysts do exhibit a level of conservativeness regarding sales growth, they seem to believe that the profitability of companies can be vastly increased from historical standards. Further, they seem to believe, from the increased ROIC levels, that the company can increase the utilization of its asset base and create larger amounts of value than what is historically evident. In the following section we will discuss these observations effect on accuracy and bias.

Revenue, NOPAT-margin and ATO

As evident from the results, historic volatility is a big concern in terms of forecasting accuracy and target price errors. We found significant correlation between relative standard deviation of company revenues and target price forecasting errors, showing that analysts are less accurate on historically volatile companies. This is somewhat in line with Kerl (2011), who finds that higher stock volatility increases forecast errors. Further, correlation tests on historic volatility and revenue forecast levels revealed that analysts apply more conservative growth estimates to companies of higher volatility. This finding is rather contradictory to prior research, as Das et al. (1998) found that increased volatility leads to more optimistic forecasts – the intuition being that increased volatility makes firms more difficult to predict, which leads to increased optimism about the ability to increase operational capabilities. A similar finding was however also found in our results, but on the NOPAT-margin variable, where analysts were found to increase forecasted margin levels in comparison to historical levels (larger margin deviation) for firms with increased volatility. Our hypothesis on this result is that analysts tend to focus on recent years' performance when forecasting volatile firms, rather than longer historic periods, and thus deviate more from historical averages. When testing this hypothesis, we found larger correlation coefficients between the last year of historical NOPAT-margin and the forecasted average NOPAT-margin, leading us to believe that this forecasting behavior is the primary driver between the aforementioned correlations.

We found, rather surprisingly, that forecasted revenue CAGR was not correlated with STPFE. Rather, when plotting the two, we found somewhat random relationships, indicating that large forecast CAGR rates are not necessarily hurting absolute forecast error rates, but rather tilt estimates from being conservative to aggressive. We found forecasted CAGR levels to be less correlated with year two and three EIR, further indicating that the growth rates forecasted in our sample are generally at reasonable levels compared to market expectations, as the biases are becoming less positive as the horizon length increases. Quite surprisingly, we found no significant correlations between revenue CAGR deviation and our accuracy measures. This was unexpected, as we hypothesized that analysts deviating more from historical average growth rates would be less accurate.

We found no significant correlations between neither of our NOPAT-margin variables and our accuracy variables, with very small correlation coefficients. As evident from Figure 8, this finding was rather surprising to us, as average forecasted NOPAT levels are generally increasing rapidly after the last realized financial year, leading us to hypothesize that NOPAT-margin deviation would be an explanatory variable in the aggressiveness and bias of target price estimates.

On our last DuPont-related financial variable, Asset Turnover (ATO), we found similar results as in the NOPAT-margin tests. As shown, the ATO levels are, on average, rather conservatively estimated in our sample, and thus we did not expect ATO levels to increase the incremental predictability of forecast errors by large margins. We did find a slight, yet significant positive correlation between historic ATO volatility and STPFE in year one, indicating that more volatile ATO levels historically lead to larger absolute forecast errors. However, when testing the deviations between historical ATO levels and forecasted ATO levels, no significant correlation was present.

Market capitalization

The final financial variable under scrutiny was the market capitalization of companies valued in our sample. The previous literature is, as discussed in our review, rather torn on this topic.

Through our tests, we found logMcap to be highly negatively correlated with TPFE, meaning that analysts generally are less likely to overshoot target prices on large companies in comparison to smaller companies. We also found evidence that analysts covering larger companies had smaller absolute forecast errors, as STPFE was significantly negatively correlated with logMcap. These findings are consistent with the findings of Lim (2001), who concluded that firm size was negatively correlated with forecast bias, and Falkenstein (1996), who found share price prediction errors to be inversely related to firm size and liquidity. However, Bonini et al. (2010) found that the magnitude of forecasting errors increased with firm size, which is contradictory to our findings. From literature, market capitalization has often been used as a proxy for information availability under the hypothesis that larger firms have richer information environments. If we accept this hypothesis as being true, then our findings are somewhat expected, as richer information environments would allow for more in-depth analysis of the company operations, and thus less degrees of uncertainty about future earnings potential.

When plotting the forecasted revenue CAGR, forecast NOPAT-margin and forecasted ATO levels to logMcap, we found revenue CAGR levels to be completely randomly distributed compared to market capitalization. However, we found NOPAT margins to be negatively correlated, and ATO levels to be positively correlated with firm size. In light of the DuPont decomposition, these findings are rather interesting. Soliman (2008) showed that decomposing company financials into DuPont variables provided incremental insights into the company operations, and Fairfield & Lombardi (2001) showed that disintegrating historic ΔROIC into ΔPM and ΔATO provided incremental predictive power on forecasting ROIC levels. Our findings on this decomposition alone yielded rather contradictory results, as we found no explanation power in neither NOPAT-margin nor ATO levels at all. However, when introducing the effects of firm size, it seems that the differences become more apparent. In hindsight, it would have been interesting to test the correlations derived from DuPont decomposed financials, compared to correlations derived by keeping ROIC as-is, while simultaneously controlling for market capitalization. Our hypothesis for such a test would be that we would derive greater explanatory abilities from this decomposition, and thus arrive at more consistent results to those of Soliman and Fairfield & Lombardi.

P/E ratio

Besides the market capitalization of companies valued, Doukas, Kim & Pantzalis (2005) found companies with larger analyst coverage were overvalued, and thus trading above their fundamental values. We therefore sought to uncover whether companies trading high on P/E values were more difficult to accurately estimate. To try and test this on our sample, we computed the price-earnings ratio at T_0, which we tested against forecast errors and accuracies. Our results are however limited by the lack of information regarding industry averages from which the companies operate. We do not know whether a P/E of 12 is large without considering the industry characteristics of the firm at hand. However, our results were somewhat unanimous, with significant negative correlation between P/E ratio as T_0 and standardized target price forecast errors, concluding that target price errors are smaller for companies trading higher on P/E at the time of valuation. While this fact does not speak to whether *overvalued* companies are easier to predict, it does however show that companies valued higher on P/E ratio are easier to predict. If we consider these results in the light of Doukas et al. (2005), one possible explanation may be that a larger P/E ratio increases analyst coverage, and increased analyst coverage increases the information environment of the firm. However, such conclusions are vastly outside the scope of this thesis, and thus we can only hypothesize about these relationships.

6.1.3 Methodological variables

As with the other aspects of our analysis, the methodological choices made by analysts proved to contain some interesting findings, which will be discussed below.

Model choice

Through analysis of valuation models and multiple usage, we found that analysts using multiples along with intrinsic present value models were prone to publish more aggressive estimates and exhibit more positive bias. This finding is rather interesting on many levels. First and foremost, it questions the reasoning behind using multiples when also seeking to measure intrinsic values. As is well established, multiple valuation differ fundamentally from present value models, as multiple valuation is based on the notion that markets are efficient in its valuation of peer groups. Without this fundamental belief, the act of doing valuation with multiples would be rather obsolete, as you would be basing your valuation on a mis-valued peer group. On the contrary, when using present value models, you assume that the value of a company is the present value of future cash flows, and as such, you have a fundamental belief that the company is somewhat incorrectly priced by the market, as you embark on the task of finding its *true* value. Therefore, the main ideas of relative valuation and intrinsic valuation seem to be in conflict. This begs the question of why analysts use multiples and intrinsic value models simultaneously, as evident from our findings, using both leads to more positively biased estimates. The answer to this question is outside the scope of this paper, but the relevance of the question still stands – do analysts use multiples as a way of sanity-checking their intrinsic value estimates? And if so, are they efficient in establishing a suitable peer group? From our results, it seems that analysts are either not very proficient in establishing a comparable peer group, or too influenced of the peer group in forming their valuation model assumptions, thus ultimately hurting their target price estimates. Further, in the light of the findings of Petersen, Plenborg, & Kinserdal (2017), who found that roughly 90% of practitioners were using multiples, the above findings would be highly relevant, as multiple usage along with present value models seems to increase bias.

When considering the above results in light of literature, there seem to be some debate on the use of multiples versus present value models. Our results are rather consistent with those of Sayed (2015), who found that the discounted cash flow model yielded more accurate target prices than those derived from using multiples. Imam et al. (2013) found, contrary to our results, that cash flow-based valuation models in conjunction with multiples improve forecast accuracy. And lastly, Asquith et al. (2002) found no relationship between valuation model choice and target price forecast accuracy at all. Our results thus deviate from the findings of Imam et al. and Asquith et al., further cementing the inconsistency in findings on this matter. One possible, and rather pragmatic explanation could be that the model choice is only relevant to consider insofar the assumptions and input of the model are error-free, and thus without any risk of interference.

Regarding the model input, we found some interesting findings on the utilization of more than one present value valuation model. As evident from Table 26, we found that analysts using more than one present value model had significantly larger standardized target price forecast errors, which goes against the concept of internal consistency between models, as argued by Lundholm & O'Keefe (2001). The very concept of present value models yielding the same results, given the same inputs and assumptions, has long been tested and ascertained by academics, and as such, we did not expect to find significant differences between analysts using more than one model.

Forecast horizon

Another interesting finding within our methodological results is that of impact of the forecast and historical horizon. Surprisingly, we found no evidence that the length of forecast horizon matters in terms of accuracy or bias, which is inconsistent to the findings of Richardson et al. (1999), who found that inaccuracies increase with larger forecast horizons. Present value models are often criticized for their large reliance on the terminal value, as that makes up a large part of the total enterprise value (Platt et al. 2009; Green et al., 2016) – and therefore, intuitively, shorter forecast periods require more effort in guaranteeing realistic terminal value assumptions, as shorter forecast horizons, all else equal, lead to larger reliance on the terminal value. However, an opposing argument is that forecast periods should be kept short, for the reasons also proposed by Richardson et al. (1999) – namely that errors accumulate over longer horizons, causing higher inaccuracies.

Historic period

While the results in our study do not succeed in showing significant differences between forecast lengths, we do find that there is great utility in increasing the number of years of history being analyzed. As we saw in section 5.2.3, both the aggressiveness of target price estimates and the standardized forecast errors were significantly lower for analysts using five years of historical financials in comparison to those only using three or four years of financials. These results were rather expected, as it ties somewhat into the notion of information environments, as discussed previously. Longer historical periods would, all else equal, give you a fairer basis for judging the operating performance of the company.

Steady state

When considering the steady state assumptions of analysts, a few interesting results became apparent. First and foremost, we tested the impact of a binary steady state assumption satisfied/not satisfied variable on our accuracy measures. Here, we found no significant differences, which was rather surprising. For instance, we hypothesized that analysts having stable ROIC margins between last forecast year and terminal year would perform more accurate value estimates. Consistent with Levin & Olsson (2000), we also hypothesized that analysts failing to forecast two consecutive years of stable growth and margins would be less accurate, but these hypotheses were not fulfilled. Rather, we found that only 20% of the analysts in our sample forecasted two consecutive years of stable growth and margins, and only 45% forecasted stable ROIC margins in two consecutive years. Neither of which had any impact on our accuracy measures.

ROIC

However, while our binary variable showed no significant differences, we found that ROIC deviation between forecast average and terminal year was significantly positively correlated with both TPFE, STPFE and EIR measures, showing that larger ROIC levels in the terminal year compared to forecast average ROIC yielded both more aggressive estimates, more positive bias and larger absolute target price errors. Tying the above findings back to the DuPont decomposition, as discussed earlier, is rather interesting – as ATO and NOPAT-margin deviation alone did not show significant correlations, but when combining them to ROIC, the correlations were rather large and significant, inconsistent with previous literature by Soliman (2008) and Fairfield & Lombardi (2001). These findings indicate that ROIC as a combined metric is of utmost importance, but the levers of ROIC may be more freely forecasted, and thus the characteristics of the company can change during the forecast period through to the terminal period, without significant effect on accuracy or bias, if the return on invested capital does not deviate significantly from forecasted averages. Intuitively, this also makes sense, as ROIC is a measure of overall value creation, and if ROIC stays constant, the combination of ATO and NOPAT-margin is irrelevant.

WACC

While literature is somewhat divided on the computation of WACC, where some believe that a target capital structure should be used in determining the weights of cost of debt and cost of equity, and others propose that WACC should be computed using an iterative method (Larkin, 2011), we sought to test the impact of changing WACC during forecasts on our accuracy measures. Surprisingly, only 13 analysts of 243 had changes in WACC during the forecast period, and we failed to find significant differences between these groups. The inference of these results may be limited by the very low count of analysts having changed WACC – however, we still found it rather surprising, as the firm leverage was very rarely constant during forecast periods, and thus we expected more analysts to have dynamic WACC levels during explicit forecast periods.

Terminal growth rate

The last variables under scrutiny in our results all relate to the terminal value. As previously argued, great consideration should be taken when computing terminal value assumptions, as the terminal value comprise a substantial part of the enterprise value in a DCF model. From literature we found that terminal value growth should seldom be larger than the risk-free rate, as the risk-free rate can proxy the nominal GDP rate of the economy in which the company operates (Damodaran, 2002). Therefore, we sought to test this relationship, but we found no significant differences between groups forecasting a larger terminal growth rate than risk-free rate. However, we also sought to test the outer tails of deviation, finding that terminal value growth in excess of 1.5x the risk-free rate provided lower chances of hitting target prices, but also more conservative estimates, however accompanied by larger target price forecast errors. These results are very surprising, as the intuition would be the opposite – namely that an aggressively forecasted terminal growth level would lead to more aggressive estimates. These correlations remain unexplained, but they would definitely be interesting to uncover in another study. One possible explanation of the incoherence might be in the limitations of our research design – namely that we fail to control for the impact of other variables.

WACC vs ROIC in terminal period

From our review of the literature, it became evident that the relationship between WACC and ROIC in the terminal year is important, as you, on one hand, believe that the company will create excess value forever if ROIC > WACC, and on the other hand believe that the company will destroy value forever if WACC > ROIC. In our sample, 145 of 183 analysts forecasted ROIC to be larger than WACC, and thus believing excess value to be created in perpetuity. We sought to test differences on accuracy by splitting the sample at the zero-deviation level, thus artificially forming a group of analysts who believed in excess value creation, and one who believed in value destruction. The results showed no significant differences in accuracy or bias between groups, which was rather surprising, as we had a clear hypothesis that bias would be larger and target prices more aggressive on those forecasting larger ROIC than WACC in the terminal year. We proceeded in testing the correlation between ROIC minus WACC on our accuracy measures to test the possibility of correlations beyond a binary splitting of groups. Here, we found a significant negative correlation between STPFE and ROIC minus WACC deviation, showing that larger positive differences significantly decrease the absolute forecasting errors, which, again, was surprising. This becomes even more peculiar in relation to another finding from the analysis, namely that higher ROIC in the terminal year compared to the year before, increased positive bias and absolute forecast error.

6.2 Implications of results

Based on the above discussions, few implications for both practice and academia become evident.

First and foremost, we found consistent positive bias to exist without the interference of analyst-corporate relationships, as otherwise suggested as main causes by previous literature (Dechow, Hutton, & Sloan, 2000; Hong & Kubik, 2003; Lin & McNichols, 1998; McNichols & O'Brien, 1997), suggesting that bias may be apparent for other reasons than the fear of information cut-off. Students might not partake in a principal-agent problem in the same way professional analysts do, but they might still exhibit other biases. You would assume students wanting to get the best grade possible, which could potentially lead them to cross-check their target prices with professional analysts' target prices, in order to not deviate too much, and therefore get a less "controversial" result. In this way, students could also exhibit herding behavior, just like professional analyst do. As this information is not readily available in the analyzed projects, you would most likely have to conduct interviews to further investigate this issue. A more likely explanation for the positive bias is probably the fact that students have chosen the company they want to cover themselves. You could speculate, that they generally have chosen companies that they found interesting, which may to a larger degree be companies of higher growth rates.

The strength of our conclusions generally suffers from the lack of investigative control variables, for example gender, educational background, buy/sell recommendations, and size of the companies (market cap), which were all found to be significant in relation to accuracy. We might also control for the industry of the companies, as it is quite heavily skewed toward for example brewers, luxury goods and airlines, as these have proven exceptionally popular among students, with a combined 30% of our total dataset belonging to these industries (see Appendix G), and one-third of our dataset consisting of the same 10 companies (see Appendix H). On the positive side, our dataset spans around a decade, thus minimizing the effect of general market conditions within these industries. The fundamental reason behind these issues is, as previously mentioned, limitations in our dataset. A larger sample would enable us to better control for these variables.

Below are listed the results which we have found difficulty explaining, because they were either opposed to existing literature or common intuition, and thus interesting to pursue for further research:

- Higher ROIC minus WACC led to lower forecast errors, but higher ROIC in the terminal year compared to the year before led to higher forecast errors. This relationship is difficult to assess and could be a subject for further study.
- Growth in excess of 1.5x the of the risk-free rate in the terminal period led to lower TPA, but more conservative target prices. The "sanity checks" of terminal growth rates as previously laid out by literature therefore seems more nuanced.
- Students were still positively biased, even though they do not partake in a principal-agent problem like professional analyst. This finding indicates that the bias of analysts goes beyond the explanations derived from previous research.
- Increased historic revenue volatility negatively affects target price accuracy, but no such relationship was found regarding historic ATO or NOPAT-margin volatility. This finding spark further interest in the forecasting behavior surrounding ATO and NOPAT-margins.
- There was found no relationship between forecast horizon length and accuracy, but longer historical periods were associated with more accurate target price estimates. This indicates that analysts may, in practice, shorten their forecast horizons as to secure more accurate forecasting abilities, but simultaneously utilize more historical data as a basis for their forecasts. Naturally, further studies need to be made to ascertain this relationship.

The above issues should serve as a reminder to the general applicability of our results. Our results can hardly be used as a checklist for analysts to propose the most accurate forecasts, but can rather act as a guideline to what analysts should consider when estimating target prices of companies. For instance, we do not recommend increasing the difference between ROIC and WACC in the terminal year without reason, in order to decrease forecast error, but it is obvious that this relationship influences the accuracy and is therefore something to consider when forecasting. It is also certain, that students are inaccurate and positively biased, but it is to a large extent still a mystery what drives students' or analysts' forecast accuracy. Generally, investors should rely less on single-analyst recommendations, as we found target price errors to have wide standard deviations, signaling that pooling analyst recommendations would provide better estimates in general.

7.0 Conclusion

To investigate the accuracy and bias of student analysts target prices, two research questions were formulated, namely:

i. "Are target prices derived from student valuations accurate and unbiased?"
ii. "Which factors influence the accuracy or bias derived from student valuations?"

To answer the above research questions, a thorough literature review was conducted, with the purpose of establishing a solid methodological foundation as well as several hypotheses based on theorized explanatory variables. Through data collection on analyst-specific, financial, and methodological variables from 321 student valuations on publicly traded companies, we were able to test the hypotheses presented in section 3.0.

In conclusion, we found student analysts to be both inaccurate and biased. With one-year TPA of 54%, our results are consistent with existing literature. Also, in line with existing literature, we found analysts to be consistently positively biased, which was further confirmed when analyzing the accuracy and bias measures per publication year, as only one of 13 years contained mean negative bias.

We also found significant differences in accuracy and bias when dividing the dataset into buy and sell recommendations. Sell recommendations were less positively biased than buy recommendations, which confirmed our intuition. We further found sell recommendations to be more accurate, measured across all variables.

We found female analysts to be significantly more aggressive and having more positive bias in their target prices, contrary to the findings of previous literature. We failed to find significant differences in accuracy between male and female analysts.

We could not find any significant relationship between group size and accuracy or bias, which was contradicting to existing literature. As our sample only consisted of individuals and groups of two analysts, these results are however less comparable to existing literature, as previous studies included larger groups.

We found that higher historic revenue and ATO volatility negatively affected accuracy. We found no such effect in relation to NOPAT-margin volatility. The deviation between historical and forecasted revenue, NOPAT-margin and ATO, showed no significant association with accuracy or bias measures either.

We found a clear relationship between market cap at T_0 and accuracy. A higher market cap led to less positive bias, and more accurate target prices. We also found higher target price accuracies for companies trading at higher P/E ratios at the date of valuation, consistent with previous literature on analyst coverage and richness of information environments surrounding larger firms.

We did not find significant relationships between the number of multiples used and accuracy. However, we found that analysts using multiples in addition to present-value models increased the level of positive bias exhibited. This relationship was rather surprising, as multiples are often used in practical settings, and would thus be interesting for further studies.

We found no significant relationship between the length of the forecast period and accuracy. However, we found a significant positive relationship between the historical period presented in the valuations and accuracy, with the use of five historic years to be significantly more accurate than using four and three years.

No significant relationship was found between satisfying the steady state assumption and accuracy. Further, no significant relationship was found between satisfying our growth sanity check (i.e., setting the terminal growth rate equal to the risk-free rate) and accuracy. Changing WACC during the forecast showed no significant relationship with accuracy either, contrary to recommendations presented in literature. However, we found a significant relationship between WACC and ROIC in the terminal year. Quite surprisingly we found that higher "WACC minus ROIC" in the terminal period led to greater accuracy and less target price forecasting errors.

While our conclusion regarding our primary research question is arguably robust, we realize that the inferential strength of conclusions regarding our secondary research question generally suffers from the lack of investigative control variables, for example gender, educational background, buy/sell recommendations and firm size. We might also want to control for specific industries and companies. The lack of control variables was primarily due to limitations in the size of our dataset when controlling for these variables. We therefore propose that future research conducted be more thorough in utilizing control variables, as naturally, a multitude of variables might affect analysts' accuracy and bias.

8.0 Bibliography

Abarbanell, J. S. (1991). Do analysts’ earnings forecasts incorporate information in prior stock price changes? *Journal of Accounting and Economics, vol. 14*, 147-165.

Agresti, A., & Franklin, C. (2013). *Statistics: The Art and Science of Learning from Data, Third edition.* Pearson.

Amiram, D., Owens, E., & Rozenbaum, O. (2016). Do information releases increase or decrease information asymmetry? New evidence from analyst forecast announcements. *Journal of Accounting and Economics*, 121-138.

Asquith, P., Mikhail, M. B., & Au, A. S. (2002, october). INFORMATION CONTENT OF EQUITY ANALYST REPORTS. *Working Paper 9246 NATIONAL BUREAU OF ECONOMIC RESEARCH*.

Asquith, P., Mikhail, M. B., & Au, A. S. (n.d.). INFORMATION CONTENT OF EQUITY ANALYST REPORTS. *Working Paper 9246, NATIONAL BUREAU OF ECONOMIC RESEARCH*.

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

Banghøj, J., Jensen, M., Nielsen, T. T., & Plenborg, T. (2009). Nøgletal fra analyserapporter. *INSTITUT FOR REGNSKAB OG REVISION, Handelshøjskolen i København*, 1-14.

Bjerregaard-Nielsen, A. (2015). STOCK ANALYSTS’ FORECASTING ACCURACY - STILL A PUZZLE. *Master thesis, Copenhagen Business School*.

Bonini, S., & Kerl, A. (2012). Subjective valuation and target price accuracy. *Bocconi University & University of Giessen*.

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

Bosquet, K., de Goeij, P., & Smedts, K. (2014). Gender heterogeneity in the sell-side analyst recommendation issuing process. *Finance research letters, Vol. 11*, 104-111.

Bradshaw, M., & Brown, L. (2002). The Use of Target Prices to Justify Sell-Side Analysts' Stock recommendations. *Accounting Horizons, Vol 16.*, 27-41.

Bradshaw, M., Brown, L., & Huang, K. (2013). Do sell-side analysts exhibit differential target price forecasting ability? *Review of accounting studies, Vol. 18*, 930-955.

Brown, L., & Mohd, E. (2003). The Predictive Value of Analyst Characteristics. *Journal of Accounting, Auditing and Finance, Vol. 18*, 625-647.

Brown, L., Foster, G. P., & Noreen, E. (1985). Security Analyst Multi-Year Earnings Forecasts and the Capital Market. *American Accounting Association.* Sarasota, FL.

Campbell, J. L., Quincy, C., Osserman, J., & Pedersen, O. K. (2013). Coding in-depth semistructured interviews: Problems of unitization and intercoder reliability and agreement. *Sociological Methods & Research, Volume 42*, 294-320.

Cassia, L., Plati, A., & Vismara, S. (2007). EQUITY VALUATION USING DCF: A THEORETICAL . *Investment Management and Financial Innovations, Volume 4, Issue 1*, pp. 91-107.

Chan, L. C., Karceski, J., & Lakonishok, J. (2003). ANALYSTS' CONFLICT OF INTEREST AND BIASES IN EARNINGS FORECAST. *National Bureau of Economic Research*.

Charness, G., & Gneezy, U. (2012). Strong Evidence for Gender Differences in Risk Taking. *Journal of Economic Behavior & Organization*, 50-58.

Cheng, C. S., Chu, K. K., & Ohlson, J. (2020). Analyst forecasts: sales and profit margins. *Review of Accounting Studies (2020), 25*, pp. 54-83.

Damodaran, A. (2002). *Investment Valuation.* New York: John Wiley & Sons.

Damodaran, A. (2006). Growth Rates and Terminal Value, New York University - Stern School of Business . *http://people.stern.nyu.edu/adamodar/pdfiles/ovhds/dam2ed/growthandtermvalue.pdf* .

Das, S., Levine, C. B., & Sivaramakrishnan. (1998). Earnings predictability and Bias in Analysts' Earnings Forecast. *THE ACCOUNTING REVIEW, Vol. 73, No. 2*, pp. 227-294.

Dechow, P., Hutton, A., & Sloan, R. (2000). The relationship between analysts' forecasts of long-term earnings growth and stock price performance following equity offerings. *Contemporary Accounting Research 17*, 1-32.

Doukas, J. A., Kim, C. F., & Pantzalis, C. (2005). The Two Faces of Analyst Coverage. *Journal of Financial Management*, 100-124.

Fairfield, P., & Teri Lombardi, Y. (2001). Using Asset Turnover and Profit Margin to Forecast Changes in Profitability. *Review of Accounting Studies volume*, 371-385.

Falkenstein, E. (1996). Preferences for stock characteristics as revealed by mutual fund. *The Journal of Finance, Vol. 51*, 111-135.

Feng, G. C. (2014). Intercoder reliability indices: Disuse, misuse, and abuse. *Quality & Quantity, Vol 48*, 1803-1815.

Gleason, C. A., Johnson, W. B., & Li, H. (2011, august). Valuation Model Use and the Price Target Performance of Sell-Side Equity Analysts. *Contemporary Accounting Research 30(1)*.

Green, C., Jagadeesh, N., & Tang, Y. (2009). Gender and Job Performance: Evidence from Wall Street. *Financial Analysts Journal, Vol. 65*.

Green, J., Hand, J. R., & Zhang, X. F. (2016). Errors and questionable judgments in analysts' DCF models. *Review of Accounting Studies* , pp. 596–632.

Hilary, G., & Hsu, C. (2013). Analyst Forecast Consistency. *The Journal of Finance*, 271-297.

Hong, H., & Kubik, J. (2003). Analyzing the Analysts: Career Concerns and Biased Earnings Forecasts. *The Journal of Finance*.

Hope, O.-K., & Fang, B. (2020). Analyst Teams. *Review of Accounting Studies*.

Huang, R., Krishnan, M., Shon, J., & Zhou, P. (2017). Who Herds? Who Doesn't? Estimates of Analysts' Herding Propensity in Forecast Earnings. *Contemporary Accounting Research Vol. 34 No. 1*, pp. 374-399.

Ikromov, N., & Yavas, A. (2012). Cash Flow Volatility, Prices and Price Volatility: An Experimental study. *The journal of real estate finance and economics, 2012-01, Vol.4 4*, pp. 203-229.

Imam, S., Chan, J., & Ali Shah, S. Z. (2013). Equity valuation models and target price accuracy in Europe: Evidence from equity reports. *International Review of Financial Analysis*, 9-19.

Kerl, A. G. (2011). Target Price Accuracy . *Verband der Hochschullehrer für Betriebswirtschaft e.V.*, 74-94.

Kolbe, R. H., & Burnett, M. S. (1991). Content-analysis research: An examination of applications with directives for improving research reliability and objectivity. *Journal of Consumer Research, Vol. 18*, 243-250.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, Vol. 33*, 159-174.

Larkin, P. J. (2011, november). To Iterate Or Not To Iterate? Using The WACC in Equity Valuation. *Journal of Business & Economics Research, Volume 9, Number 11*.

Levin, J., & Olsson, P. (2000). Terminal value techniques in equity valuation –implications of the steady state assumption. *Working paper, Stockholm School of Economics*.

Lim, T. (2001). Rationality and Analysts' Forecast Bias. *The Journal of Finance*, 369-385.

Lin, H., & McNichols, M. (1998). 1998. Underwriting relationships, analysts' earnings forecasts and investment recommendations. *Journal of Accounting and Economics, vol. 25*, 101-127.

Lombard, M., Snyder-Duch, J., & Bracken,, C. C. (2002). Content analysis in mass communication: Assessment and reporting of intercoder reliability. *Human Communication Research, Volume 28*, 587-604.

Lundholm, R., & O'Keefe, T. (2001). Reconciling Value Estimates from the Discounted Cash Flow Model and the Residual Income Model. *Contemporary Accounting Research, Vol. 18*, 331-335.

McKinsey & Company. (2017). *Reinventing Equity Research As a Profit-Making Business.* London & New York: McKinsey.

McNichols, M., & O'Brien, P. (1997). Self-selection and analyst coverage. *Journal of Accounting Research, Vol 35*, 167-199.

Neuendorf, K. A. (2002). *The content analysis guidebook.* Sage.

Nielsen, J. M. (2011, August). *Dansk Aktionærforening*. Retrieved from shareholders.dk: https://www.shareholders.dk/arkiv/11421/vi-anklager-pandora-skandalen-narrede-de-smaa

O'Connor, C., & Joffe, H. (2020). Intercoder Reliability in Qualitative Research: Debates and Practicale Guidelines. *International Journal of Qualitative Methods*, 1-13.

Palepu, K., Bernard, V., & Healy, P. (1996). Business analysis and Valuation: Using Financial Statements. *Cincinatti, Ohio: South-Western College Publishing*.

Patel, H. (2018). Target Price Achievement and Target Price Accuracy Models: An Analysis of Advisory Firms' Recommendation for the Indian Banking Stocks. *Global Business Review*.

Penman, S. (1997). A Synthesis of Equity Valuation Techniques and the TerminaL Value Calculation for the Dividend Discount Model. *Review of Accounting Studies, Vol. 2, No. 4*, 303-323.

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

Platt, H., Demirkan, S., & Platt, M. (2009, November - December). PERILOUS FORECASTS: IMPLICATIONS OF RELIANCE ON THE TERMINAL VALUE. *Corporate Finance Review, Vol. 14, No. 3*.

Popping, R. (1988). On agreement indices for nominal data. In W. E. Saris, & G. N. Irmtraud, *Sociometric research: Volume 1, data collection and scaling* (pp. 90-105). New York: St. Martin's Press.

Richardson, S., Teoh, S., & Wysocki, P. D. (1999). Tracking Analysts' Forecasts Over the Annual Earnings Horizon: Are Analysts' Forecasts Optimistic or Pessimistic? *Working Paper*.

Sayed, A. (2015). Should Analysts Go by the Book? Valuation Models and Target Price Accuracy in an Emerging Marke. *Global Business Review*, 832-844.

Sjølin, S. (2021, maj 14). *Børsbarometer: Euforisk aktiemarked gør børsdrømme til virkelighed.* Retrieved from Berlingske.dk: https://www.berlingske.dk/aktier/boersbarometer-euforisk-aktiemarked-goer-boersdroemme-til-virkelighed

Soliman, M. T. (2008). The Use of DuPont Analysis by Market Participants. *The Accounting Review*, 823-853.

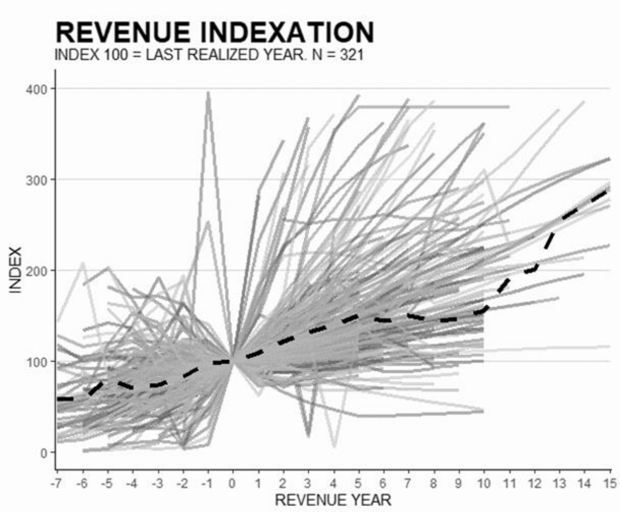
Stickel, S. (1990). Predicting individual analyst earnings forecasts. *Journal of Accounting Research, vol 28*, 409-417.

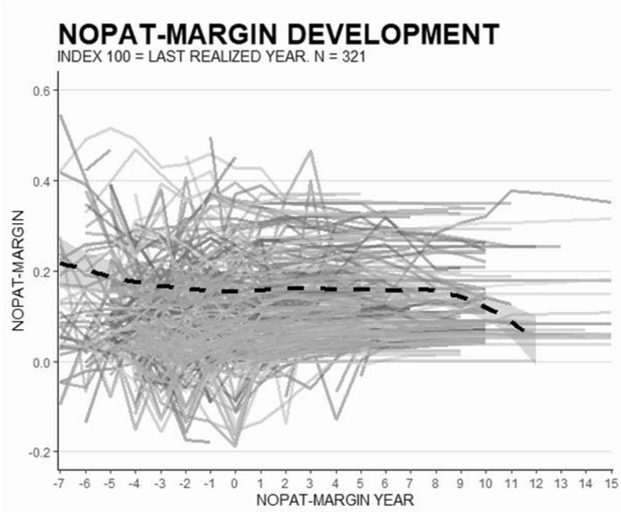
Stickel, S. (1992). Reputation and Performance Among Security Analysts. *The Journal of Finance*.

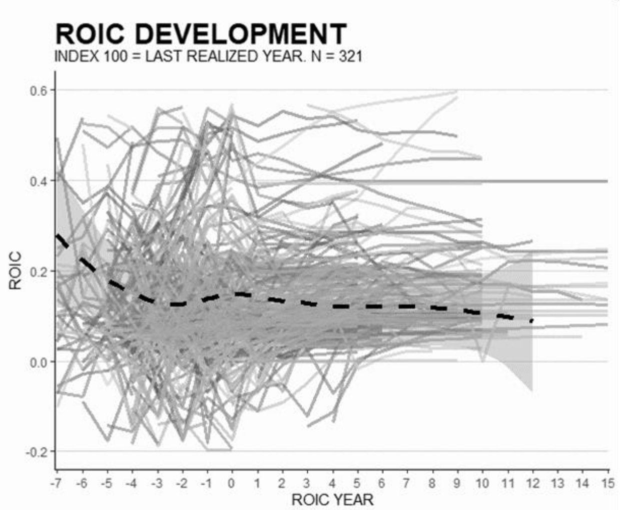
Vriborg Petersen, C., & Plenborg, T. (2008). The implementation and application of firm valuation models. *Department of Accounting and Auditing, Copenhagen Business School*.

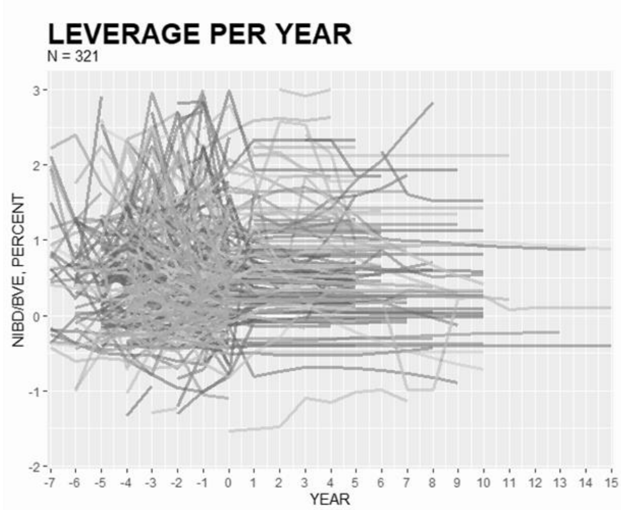
9.0 Appendices

Appendix A









Appendix B

Codebook Variables Summary

Thesis information

- Title of thesis
- # pages (manually calculated)
- # pages (scraped)
- # of characters
- Number of study programs
- Study program of author(s)

- Gender(s) of author(s)

- Number of pages strategic analysis
- Number of pages financial analysis

General Company Information

- Company name
- Industry
- CAP IQ Company name
- Year of establishment
- Is listed? (1/0)
- Ticker if listed
- Year of IPO

Methodology

EV - income statement (1/0)

Used together with peer benchmark? (1/0)

- EV/EBITDA
- EV/EBIT
- EV/Sales
- EV/NOPAT
- EV/FCFF
- Other (note specific mulitiple in "other" sheet)

EV - balance sheet (1/0)

Used together with peer benchmark? (1/0)

- Other (note specific mulitiple in "other" sheet)

Equity - income statement (1/0)

Used together with peer benchmark? (1/0)

- P/E
- Other (note specific mulitiple in "other" sheet)

Equity - balance sheet (1/0)

Used together with peer benchmark? (1/0)

- P/BVE
- Other (note specific mulitiple in "other" sheet)

Present value (1/0)

- EV- DCF
- EV - EVA
- EV - APV
- Equity - DCFE
- EV - RI
- Other (note specific mulitiple in "other" sheet)

Liquidation value (1/0)

Sensitivity analysis (1/0)

Gordons growth model (1/0)
Value driver model (1/0)
Exit mupltiple method (1/0)
Steady state growth satisfied (1/0/"N/A")

Budget control section (1/0)

Financial data

- Currency
- First Year in Forecast
- Terminal year

Revenue adjusted (1/0)
- Revenue figures, year -7 to 15

EBIT adjusted (1/0)
- EBIT figures, year -7 to 15

NOPAT adjusted (1/0)
- NOPAT figures, year -7 to 15

Net income adjusted (1/0)
- Net Income figures, year -7 to 15

Assets adjusted (1/0)
- Asset figures, year -7 to 15

Equity adjusted (1/0)
- Equity figures, year -7 to 15

Invested Capital adjusted (1/0)
- Invested Capital figures, year -7 to 15

Net debt adjusted (1/0)
- Net debt figures, year -7 to 15

FCFF adjusted (1/0)
- FCFF figures, year -7 to 15

Valuation

- PV of FCFF in explicit forecast
- PV of terminal value

- Enterprise value
- Equity value
- Shareprice currency
- Implied shareprice
- Implied shareprice date
- EV or EQV adjusted (1/0)
- Actual shareprice noted in thesis
- Actual shareprice date

- WACC rate
- Change of WACC during forecast (1/0)?
- Terminal growth rate
- Risk free rate

- Calculation mistake (1/0)
- Notes

Appendix C

Variable	Explanation
ID	ID of valuation
NO_ANALYSTS	Number of analysts
STUDY_PROG	Study programme of analyst
GENDER	Gender (0 = male, 1 = female)
PAGES	Number of pages
STRAT_FIN_RATIO	Number of pages in strategic analysis / number of pages in financial analysis
EV_EBITDA	Has used EV/EBITDA multiple with peers
EV_EBIT	Has used EV/EBIT multiple with peers
EV_SALES	Has used EV/Sales multiple with peers
EV_NOPAT	Has used EV/NOPAT multiple with peers
EV_FCFF	Has used EV/FCFF multiple with peers
PE_multiple	Has used P/E multiple with peers
P_BVE	Has used P/BVE multiple with peers
Other_EVInc	Has used other EV/Income Statement multiples
AOM	Amount of multiples used
MU	Has used multiples
Number_models_used	Number of valuation models used
EV_DCF	Has used DCF
EV_EVA	Has used EVA
EV_APV	Has used APV
EV_RI	Has used Residual Income
LV	Has used Liquidation value
SA	Has made sensitivity analysis
GGM	Has used Gordons Growth Model
VDM	Has used Value Driver model
EMM	Has used exit multiple method
FORECAST_PERIOD	Years in forecast period
HIST_PERIOD	Years in historic period
Log_mcap	Log-transformed (10) market capitalization
Est_year	Company established year
IPO_year	Year of IPO
PUB_YEAR	Year of valuation publication
TV_of_EV	Terminal value share of enterprise value
WACC	WACC
WACC_chg	Has WACC changed during forecast?
G	Terminal growth rate
RF	Risk-free rate
GSC	Satisfied growth sanity check?
SS1	Checks whether the revenue growth of the last forecast year is equal to terminal value
SS2	Checks whether the growth in revenue in the second last forecast year is equal to terminal
SS3	Check whether the EBIT margin in the last forecast year is equal to the second last forecast
SS_ALL	Satisfies all steady state assumptions
RSD_REV_HIST	Relative standard deviation of historic revenue
RSD_REV_FC	Relative standard deviation of forecasted revenue
RSD_NOPAT_HIST	Relative standard deviation of historic NOPAT-margin
RSD_NOPAT_FC	Relative standard deviation of forecasted NOPAT-margin
RSD_ATO_HIST	Relative standard deviation of historic ATO
RSD_ATO_FC	Relative standard deviation of forecasted ATO
RSD_ROIC_HIST	Relative standard deviation of historic ROIC
RSD_ROIC_FC	Relative standard deviation of forecasted ROIC
LAST_HIST_YR_NOPAT	Last historic year of NOPAT-margin
GDHFC	The deviation in growth between historical CAGR and forecast CAGR. (FORECAST_CAGR
GDTYFC	The deviation in growth between last year growth and forecast CAGR. (LAST YEAR
GDHFC_PP	Same as above, but presented in percentage points (for those where CAGR is negative)
GDHTY_PP	Same as above, but presented in percentage points (for those where CAGR is negative)
NMDHFC	NOPAT-margin deviation, forecast vs. History

NMDTYFC	NOPAT-margin deviation, forecast vs. Terminal year
NMDHFC_PP	NOPAT-margin deviation, forecast vs. History (percentage points)
NMDTYFC_PP	NOPAT-margin deviation, forecast vs. Terminal year (percentage points)
ATODHFC	ATO deviation, forecast vs. History
ATODTYFC	ATO deviation, forecast vs. Terminal year
ATODHFC_PP	ATO deviation, forecast vs. History (percentage points)
ATODTYFC_PP	ATO deviation, forecast vs. Terminal year (percentage points)
ROICDFCH	ROIC forecast avg / ROIC hist avg -1
ROICDTYH	ROIC term year / ROIC hist average - 1
ROICDTYFC	ROIC term year / ROIC fc average -1
REVCH	Historic revenue CAGR
REVCFC	Forecast revenue CAGR
NMH	Historic average NOPAT-margin
NMFC	Forecast average NOPAT-margin
ATOH	Historic average ATO
ATOFC	Forecast average ATO
ROICHIST	Historic average ROIC
ROICFC	Forecast average ROIC
PE	Price/earnings ratio at T_0
REC	Recommendation (Buy/sell)
TPA1	Target price accuray year 1
TPA2	Target price accuray year 2
TPA3	Target price accuray year 3
TPFE_1	Target price forecast error year 1
TPFE_2	Target price forecast error year 2
TPFE_3	Target price forecast error year 3
STPFE_1	Standardized target price forecast error year 1
STPFE_2	Standardized target price forecast error year 2
STPFE_3	Standardized target price forecast error year 3
EIR_1	Excess implied return year 1
EIR_2	Excess implied return year 2
EIR_3	Excess implied return year 3

Appendix D

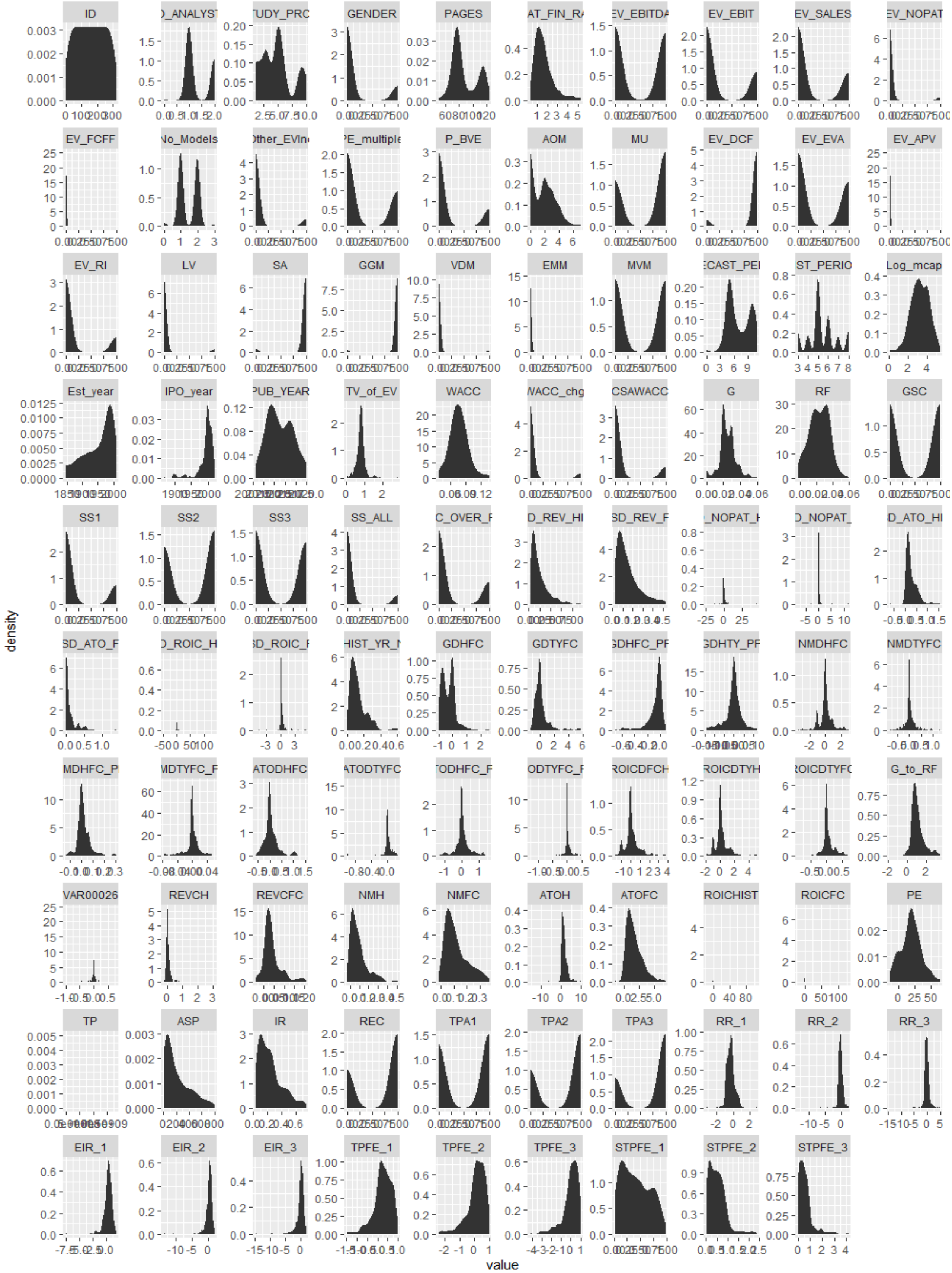
PROJECT	BINARY VARIABLES OBSERVATIONS	BINARY VARIABLES DIFFERENCES	BINARY VARIABLES ICR SCORE	FINANCE VARIABLES OBSERVATIONS	FINANCE VARIABLES DIFFERENCES	FINANCE VARIABLES ICR SCORE	ALL VARIABLES OBSERVATIONS	ALL VARIABLES DIFFERENCES	ALL VARIABLES ICR SCORE
Code1	79	7	0,91	452	147	0,67	531	154	0,71
Code2	77	2	0,97	326	29	0,91	403	31	0,92
Code3	160	5	0,97	985	169	0,83	1145	174	0,84
Code4	164	5	0,97	598	118	0,80	762	123	0,84
Code5	164	10	0,94	500	187	0,63	664	197	0,70
Code6	164	5	0,97	538	146	0,73	702	151	0,78
Code7	165	5	0,97	771	197	0,74	936	202	0,78
Code8	161	22	0,86	455	120	0,74	616	142	0,76
Code9*	164	10	0,94	310	211	0,32	474	221	0,52
Code10	165	18	0,89	546	84	0,85	711	102	0,86
Total	**1463**	**89**	**0,94**	**5481**	**1408**	**0,74**	**6944**	**1497**	**0,78**

PROJECT	BINARY VARIABLES OBSERVATIONS	BINARY VARIABLES DIFFERENCES	BINARY VARIABLES ICR SCORE	FINANCE VARIABLES OBSERVATIONS	FINANCE VARIABLES DIFFERENCES	FINANCE VARIABLES ICR SCORE	ALL VARIABLES OBSERVATIONS	ALL VARIABLES DIFFERENCES	ALL VARIABLES ICR SCORE
Code11	162	19	0,88	389	84	0,78	551	104	0,81
Code12	165	11	0,93	604	164	0,73	769	175	0,77
Code13	165	12	0,93	574	98	0,83	739	110	0,85
Code14	164	7	0,96	759	56	0,93	923	63	0,93
Code15	164	6	0,96	645	78	0,88	809	87	0,89
Code16	165	11	0,93	514	41	0,92	679	52	0,92
Total	**985**	**66**	**0,93**	**3485**	**521**	**0,85**	**4470**	**591**	**0,87**
COMBINED	**2448**	**155**	**0,94**	**8966**	**1929**	**0,78**	**11414**	**2088**	**0,82**

INITIAL ICR TEST (N=10)	ICR SCORE
BINARY VARIABLES	0,94
FINANCE VARIABLES	0,74
ALL VARIABLES	**0,78**

ON-GOING ICR TEST (N=6)	ICR SCORE
BINARY VARIABLES	0,93
FINANCE VARIABLES	0,85
ALL VARIABLES	**0,87**

Appendix E


ID
_ANALYST
TUDY_PRC
GENDER
PAGES
AT_FIN_RA
EV_EBITDA
EV_EBIT
EV_SALES
EV_NOPAT
EV_FCFF
No_Models
Other_EVIn
PE_multiple
P_BVE
AOM
MU
EV_DCF
EV_EVA
EV_APV
EV_RI
LV
SA
GGM
VDM
EMM
MVM
ECAST_PEI
ST_PERIO
Log_mcap
Est_year
IPO_year
PUB_YEAR
TV_of_EV
WACC
WACC_chg
CSAWACC
G
RF
GSC
SS1
SS2
SS3
SS_ALL
C_OVER_F
D_REV_HI
SD_REV_F
_NOPAT_I
D_NOPAT_
D_ATO_HI
SD_ATO_F
D_ROIC_H
SD_ROIC_F
HIST_YR_N
GDHFC
GDTYFC
GDHFC_PF
GDHTY_PF
NMDHFC
NMDTYFC
MDHFC_P
MDTYFC_F
ATODHFC
ATODTYFC
TODHFC_F
ODTYFC_F
ROICDFCH
ROICDTYH
ROICDTYFC
G_to_RF
VAR00026
REVCH
REVCFC
NMH
NMFC
ATOH
ATOFC
ROICHIST
ROICFC
PE
TP
ASP
IR
REC
TPA1
TPA2
TPA3
RR_1
RR_2
RR_3
EIR_1
EIR_2
EIR_3
TPFE_1
TPFE_2
TPFE_3
STPFE_1
STPFE_2
STPFE_3
density
value

Appendix F

Currency	Actual date	date +1	Exchange	Rate
EUR	22-06-2009	23-06-2009	USDEUR	0.72
NOK	28-02-2020	29-02-2020	USDNOK	9.40
USD	31-12-2019	01-01-2020	USDUSD	1.00
SEK	31-03-2020	01-04-2020	USDSEK	9.91
NOK	01-05-2015	02-05-2015	USDNOK	7.59
NOK	08-05-2015	09-05-2015	USDNOK	7.46
EUR	01-01-2012	02-01-2012	USDEUR	0.77
USD	27-12-2013	28-12-2013	USDUSD	1.00
DKK	08-04-2016	09-04-2016	USDDKK	6.53
DKK	17-11-2015	16-11-2015	USDDKK	6.21
NOK	17-07-2013	18-07-2013	USDNOK	5.99
NOK	31-12-2009	01-01-2010	USDNOK	5.78
NOK	01-01-2009	02-01-2009	USDNOK	6.94
EUR	31-12-2014	01-01-2015	USDEUR	0.83
NOK	18-02-2013	19-02-2013	USDNOK	5.55
DKK	30-09-2009	01-10-2009	USDDKK	5.08
SEK	18-03-2016	19-03-2016	USDSEK	8.23
DKK	30-04-2018	01-05-2018	USDDKK	6.17
DKK	16-08-2014	17-08-2014	USDDKK	5.57
DKK	01-06-2010	02-06-2010	USDDKK	6.08
DKK	01-11-2016	02-11-2016	USDDKK	6.73
USD	01-04-2014	02-04-2014	USDUSD	1.00
NOK	17-04-2015	18-04-2015	USDNOK	7.83
DKK	31-12-2016	01-01-2017	USDDKK	7.07
DKK	31-12-2009	01-01-2010	USDDKK	5.19
EUR	07-07-2016	08-07-2016	USDEUR	0.90
NOK	16-04-2015	17-04-2015	USDNOK	7.75
SEK	07-04-2014	08-04-2014	USDSEK	6.52
DKK	01-03-2013	02-03-2013	USDDKK	5.73
DKK	22-08-2018	23-08-2018	USDDKK	6.44
DKK	31-12-2017	01-01-2018	USDDKK	6.20
NOK	22-08-2008	23-08-2008	USDNOK	5.37
NOK	01-05-2017	02-05-2017	USDNOK	8.58
DKK	27-02-2015	28-02-2015	USDDKK	6.67
USD	31-12-2017	01-01-2018	USDUSD	1.00
DKK	29-06-2011	30-06-2011	USDDKK	5.15
DKK	08-06-2009	09-06-2009	USDDKK	5.35
DKK	00-01-1900	01-01-1900	USDDKK	N/A
DKK	09-02-2017	10-02-2017	USDDKK	6.97
DKK	09-02-2017	10-02-2017	USDDKK	6.97
USD	01-04-2012	02-04-2012	USDUSD	1.00
DKK	01-05-2017	02-05-2017	USDDKK	6.82
NOK	19-03-2019	20-03-2019	USDNOK	8.53
DKK	10-03-2010	11-03-2010	USDDKK	5.45
EUR	27-04-2016	28-04-2016	USDEUR	0.88
USD	12-12-2015	13-12-2015	USDUSD	1.00
DKK	31-12-2016	01-01-2017	USDDKK	7.07
DKK	25-10-2016	26-10-2016	USDDKK	6.83
USD	31-10-2014	01-11-2014	USDUSD	1.00
DKK	02-12-2018	03-12-2018	USDDKK	6.58
DKK	22-07-2017	23-07-2017	USDDKK	6.37
DKK	01-05-2011	02-05-2011	USDDKK	5.04
DKK	14-03-2014	15-03-2014	USDDKK	5.36
DKK	02-05-2018	03-05-2018	USDDKK	6.23
DKK	04-02-2011	05-02-2011	USDDKK	5.49
DKK	24-04-2017	25-04-2017	USDDKK	6.85
DKK	31-01-2012	01-02-2012	USDDKK	5.68
KRW	01-03-2019	02-03-2019	USDKRW	1126.66
SEK	31-10-2016	01-11-2016	USDSEK	9.04
SEK	01-05-2013	02-05-2013	USDSEK	6.47
DKK	24-02-2016	25-02-2016	USDDKK	6.78
CHF	29-02-2016	01-03-2016	USDCHF	1.00
DKK	30-09-2013	01-10-2013	USDDKK	5.51
EUR	24-09-2014	25-09-2014	USDEUR	0.78

Currency	Actual date	date +1	Exchange	Rate
DKK	28-02-2011	01-03-2011	USDDKK	5.39
EUR	03-03-2009	02-03-2009	USDEUR	0.83
SEK	01-06-2018	02-06-2018	USDSEK	8.82
ISK	28-11-2018	29-11-2018	USDISK	123.74
DKK	05-02-2014	06-02-2014	USDDKK	5.51
NOK	27-05-2014	28-05-2014	USDNOK	5.95
DKK	15-08-2012	16-08-2012	USDDKK	6.06
DKK	30-04-2015	01-05-2015	USDDKK	6.66
DKK	04-04-2012	05-04-2012	USDDKK	5.66
DKK	18-01-2012	19-01-2012	USDDKK	5.78
DKK	31-03-2013	01-04-2013	USDDKK	5.82
DKK	01-10-2013	02-10-2013	USDDKK	5.51
DKK	04-12-2012	05-12-2012	USDDKK	5.70
DKK	10-01-2014	11-01-2014	USDDKK	5.46
DKK	14-05-2009	15-05-2009	USDDKK	5.46
DKK	05-01-2011	06-01-2011	USDDKK	5.66
DKK	15-11-2010	16-11-2010	USDDKK	5.49
DKK	30-06-2011	01-07-2011	USDDKK	5.14
DKK	28-06-2013	29-06-2013	USDDKK	5.73
DKK	02-02-2011	03-02-2011	USDDKK	5.39
DKK	31-07-2015	01-08-2015	USDDKK	6.79
DKK	30-04-2014	01-05-2014	USDDKK	5.38
DKK	16-05-2016	17-05-2016	USDDKK	6.57
DKK	31-01-2015	01-02-2015	USDDKK	6.58
DKK	24-08-2011	25-08-2011	USDDKK	5.17
DKK	01-09-2014	02-09-2014	USDDKK	5.67
JPY	31-08-2012	01-09-2012	USDJPY	78.38
DKK	31-12-2011	01-01-2012	USDDKK	5.73
DKK	14-07-2015	15-07-2015	USDDKK	6.78
DKK	00-01-1900	01-01-1900	USDDKK	N/A
EUR	31-12-2018	01-01-2019	USDEUR	0.87
SEK	12-05-2010	13-05-2010	USDSEK	7.54
DKK	00-01-1900	01-01-1900	USDDKK	N/A
DKK	20-11-2012	21-11-2012	USDDKK	5.82
DKK	15-03-2015	16-03-2015	USDDKK	7.12
DKK	31-12-2009	01-01-2010	USDDKK	5.19
DKK	17-08-2009	18-08-2009	USDDKK	5.27
DKK	01-06-2012	02-06-2012	USDDKK	5.97
DKK	26-06-2013	27-06-2013	USDDKK	5.73
DKK	30-06-2012	01-07-2012	USDDKK	5.87
DKK	31-07-2012	01-08-2012	USDDKK	6.05
DKK	01-08-2011	02-08-2011	USDDKK	5.22
DKK	01-06-2010	02-06-2010	USDDKK	6.08
DKK	22-03-2012	23-03-2012	USDDKK	5.63
DKK	23-03-2009	24-03-2009	USDDKK	5.47
DKK	16-03-2017	17-03-2017	USDDKK	6.90
DKK	13-05-2011	14-05-2011	USDDKK	5.28
DKK	01-07-2013	02-07-2013	USDDKK	5.71
DKK	30-06-2011	01-07-2011	USDDKK	5.14
DKK	31-12-2012	01-01-2013	USDDKK	5.65
DKK	30-06-2008	01-07-2008	USDDKK	4.73
DKK	03-02-2012	04-02-2012	USDDKK	5.65
DKK	01-09-2010	02-09-2010	USDDKK	5.81
DKK	30-06-2008	01-07-2008	USDDKK	4.73
DKK	00-01-1900	01-01-1900	USDDKK	N/A
DKK	31-12-2009	01-01-2010	USDDKK	5.19
DKK	26-06-2009	28-06-2009	USDDKK	5.30
DKK	05-05-2008	06-05-2008	USDDKK	4.82
DKK	28-12-2016	31-12-2016	USDDKK	7.14
EUR	02-09-2013	03-09-2013	USDEUR	0.76
USD	31-12-2018	01-01-2019	USDUSD	1.00
DKK	10-03-2011	11-03-2011	USDDKK	5.40
USD	14-09-2016	15-09-2016	USDUSD	1.00
NOK	01-04-2016	02-04-2016	USDNOK	8.30

DKK	01-04-2016	02-04-2016	USDDKK	6.54	JPY	01-01-2014	02-01-2014	USDJPY	105.33
SEK	31-03-2012	01-04-2012	USDSEK	6.61	USD	01-01-2014	02-01-2014	USDUSD	1.00
DKK	04-05-2016	05-05-2016	USDDKK	6.47	EUR	01-01-2014	02-01-2014	USDEUR	0.73
EUR	22-11-2016	23-11-2016	USDEUR	0.94	EUR	01-01-2014	02-01-2014	USDEUR	0.73
NOK	31-12-2011	01-01-2012	USDNOK	5.97	USD	31-03-2019	01-04-2019	USDUSD	1.00
SEK	16-04-2015	17-04-2015	USDSEK	8.59	DKK	31-12-2015	01-01-2016	USDDKK	6.87
DKK	30-12-2016	31-12-2016	USDDKK	7.06	DKK	31-12-2015	01-01-2016	USDDKK	6.87
NOK	31-05-2012	01-06-2012	USDNOK	6.11	DKK	09-11-2011	10-11-2011	USDDKK	5.50
NOK	31-12-2014	01-01-2015	USDNOK	7.45	DKK	22-02-2018	23-02-2018	USDDKK	6.04
EUR	30-12-2016	31-12-2016	USDEUR	0.95	USD	31-12-2012	01-01-2013	USDUSD	1.00
USD	01-04-2016	02-04-2016	USDUSD	1.00	DKK	31-03-2018	01-04-2018	USDDKK	6.05
DKK	22-02-2012	23-02-2012	USDDKK	5.61	DKK	06-02-2013	07-02-2013	USDDKK	5.52
GBP	17-11-2015	18-11-2015	USDGBP	0.66	DKK	11-02-2015	12-02-2015	USDDKK	6.58
USD	25-07-2013	26-07-2013	USDUSD	1.00	DKK	31-03-2017	01-04-2017	USDDKK	6.98
USD	31-03-2017	01-04-2017	USDUSD	1.00	DKK	05-04-2013	06-04-2013	USDDKK	5.73
SEK	29-04-2016	30-04-2016	USDSEK	8.03	DKK	31-12-2011	01-01-2012	USDDKK	5.73
NOK	01-01-2011	04-01-2011	USDNOK	5.82	DKK	18-08-2011	19-08-2011	USDDKK	5.20
NOK	19-04-2018	20-04-2018	USDNOK	7.79	DKK	02-09-2013	03-09-2013	USDDKK	5.65
EUR	23-03-2016	24-03-2016	USDEUR	0.89	DKK	07-05-2014	08-05-2014	USDDKK	5.37
DKK	31-10-2013	01-11-2013	USDDKK	5.49	DKK	08-02-2012	09-02-2012	USDDKK	5.61
ISK	01-09-2019	02-09-2019	USDISK	126.01	DKK	17-08-2011	18-08-2011	USDDKK	5.16
ISK	09-01-1900	10-01-1900	USDISK	N/A	DKK	31-03-2013	01-04-2013	USDDKK	5.82
EUR	31-12-2017	01-01-2018	USDEUR	0.83	DKK	15-03-2017	16-03-2017	USDDKK	6.92
NOK	01-04-2016	02-04-2016	USDNOK	8.30	DKK	21-05-2014	22-05-2014	USDDKK	5.45
NOK	01-06-2017	02-06-2017	USDNOK	8.45	SEK	11-04-2013	12-04-2013	USDSEK	6.34
NOK	31-12-2013	01-01-2014	USDNOK	6.06	DKK	31-12-2019	01-01-2020	USDDKK	6.66
DKK	01-09-2010	02-09-2010	USDDKK	5.81	DKK	12-10-2012	13-10-2012	USDDKK	5.76
DKK	01-10-2013	02-10-2013	USDDKK	5.51	DKK	31-08-2015	01-09-2015	USDDKK	6.65
NOK	31-12-2010	01-01-2011	USDNOK	5.81	DKK	12-10-2011	13-10-2011	USDDKK	5.40
NOK	04-12-2013	05-12-2013	USDNOK	6.15	DKK	31-12-2012	01-01-2013	USDDKK	5.65
NOK	31-12-2019	01-01-2020	USDNOK	8.65	USD	13-03-2013	14-03-2013	USDUSD	1.00
NOK	18-04-2013	19-04-2013	USDNOK	5.82	DKK	08-04-2009	09-04-2009	USDDKK	5.62
NOK	31-03-2013	01-04-2013	USDNOK	5.85	DKK	01-05-2009	02-05-2009	USDDKK	5.61
NOK	12-04-2016	13-04-2016	USDNOK	8.17	DKK	31-12-2011		USDDKK	5.73
NOK	10-03-2020	11-03-2020	USDNOK	9.59	DKK	01-06-2010	02-06-2010	USDDKK	6.08
NOK	01-05-2018	02-05-2018	USDNOK	8.09	DKK	19-08-2010	20-08-2010	USDDKK	5.81
NOK	31-12-2010	01-01-2011	USDNOK	5.81	DKK	29-03-2013	30-03-2013	USDDKK	5.82
NOK	31-03-2016	01-04-2016	USDNOK	8.27	EUR	17-10-2011	18-10-2011	USDEUR	0.73
DKK	05-02-2015	06-02-2015	USDDKK	6.49	eur	31-12-2008	01-01-2009	USDEUR	0.72
DKK	24-04-2017	25-04-2017	USDDKK	6.85	DKK	30-12-2018	31-12-2018	USDDKK	6.53
DKK	31-03-2016	01-04-2016	USDDKK	6.55	DKK	05-05-2011	06-05-2011	USDDKK	5.13
NOK	09-02-2016	10-02-2016	USDNOK	8.58	SEK	30-06-2009	01-07-2009	USDSEK	7.71
DKK	21-10-2016	22-10-2016	USDDKK	6.83	DKK	01-07-2014	02-07-2014	USDDKK	5.45
DKK	01-06-2016	02-06-2016	USDDKK	6.65	NOK	31-05-2013	01-06-2013	USDNOK	5.87
SEK	02-05-2011	03-05-2011	USDSEK	6.02	USD	01-05-2013	02-05-2013	USDUSD	1.00
CZK	30-06-2010	01-07-2010	USDCZK	20.94	USD	23-03-2018	24-03-2018	USDUSD	1.00
USD	13-05-2019	14-05-2019	USDUSD	1.00	SEK	06-04-2017	07-04-2017	USDSEK	9.02
DKK	01-04-2020	02-04-2020	USDDKK	6.81	DKK	22-02-2016	23-02-2016	USDDKK	6.77
DKK	08-05-2013	09-05-2013	USDDKK	5.66	DKK	07-02-2018	08-02-2018	USDDKK	6.07
DKK	31-12-2017	01-01-2018	USDDKK	6.20	NOK	20-03-2014	21-03-2014	USDNOK	6.06
DKK	01-01-2009	02-01-2009	USDDKK	5.32	USD	31-12-2019	01-01-2020	USDUSD	1.00
NOK	12-05-2015	13-05-2015	USDNOK	7.48	USD	20-12-2016	21-12-2016	USDUSD	1.00
NOK	30-04-2013	01-05-2013	USDNOK	5.77	USD	31-12-2015	01-01-2016	USDUSD	1.00
NOK	23-08-2018	24-08-2018	USDNOK	8.39	DKK	04-03-2009	05-03-2009	USDDKK	5.90
SEK	15-02-2013	16-02-2013	USDSEK	6.32	EUR	07-11-2012	08-11-2012	USDEUR	0.78
NOK	01-02-2017	02-02-2017	USDNOK	8.23	DKK	30-03-2012	31-03-2012	USDDKK	5.58
NOK	31-03-2017	01-04-2017	USDNOK	8.56	DKK	13-05-2011	14-05-2011	USDDKK	5.28
USD	31-12-2017	01-01-2018	USDUSD	1.00	DKK	31-01-2012	01-02-2012	USDDKK	5.68
NOK	16-04-2013	17-04-2013	USDNOK	5.73	DKK	17-03-2011	18-03-2011	USDDKK	5.32
NOK	27-04-2015	28-04-2015	USDNOK	7.73	DKK	15-03-2012	16-03-2012	USDDKK	5.69
NOK	04-01-2017	05-01-2017	USDNOK	8.58	CHF	31-12-2011	01-01-2012	USDCHF	0.94
NOK	24-02-2017	25-02-2017	USDNOK	8.38	DKK	28-02-2011	01-03-2011	USDDKK	5.39
NOK	31-12-2015	01-01-2016	USDNOK	8.86	DKK	29-04-2011	30-04-2011	USDDKK	5.03
USD	30-06-2016	01-07-2016	USDUSD	1.00	RMB	05-09-2010	02-09-2010	USDRMB	N/A
USD	23-05-2016	24-05-2016	USDUSD	1.00	DKK	31-08-2011	01-09-2011	USDDKK	5.18
USD	31-03-2014	01-04-2014	USDUSD	1.00	DKK	30-12-2011	31-12-2011	USDDKK	5.74
USD	01-05-2016	02-05-2016	USDUSD	1.00	DKK	09-06-2011	10-06-2011	USDDKK	5.14

DKK	08-04-2011	09-04-2011	USDDKK	5.15
EUR	19-04-2011	20-04-2011	USDEUR	0.70
DKK	31-12-2011	01-01-2012	USDDKK	5.73
DKK	30-04-2012	01-05-2012	USDDKK	5.62
DKK	24-08-2011	25-08-2011	USDDKK	5.17
DKK	01-03-2011	02-03-2011	USDDKK	5.41
EUR	01-05-2011	02-05-2011	USDEUR	0.68
DKK	18-04-2012	19-04-2012	USDDKK	5.67
DKK	08-04-2011	09-04-2011	USDDKK	5.15
EUR	14-09-2011	15-09-2011	USDEUR	0.73
DKK	01-04-2021	02-04-2021	USDDKK	6.32
DKK	01-05-2012	02-05-2012	USDDKK	5.62

DKK	09-02-2012	10-02-2012	USDDKK	5.60
DKK	30-09-2011	01-10-2011	USDDKK	5.56
DKK	30-09-2011	01-10-2011	USDDKK	5.56
DKK	30-09-2011	01-10-2011	USDDKK	5.56
DKK	01-06-2012	02-06-2012	USDDKK	5.97
DKK	13-05-2011	14-05-2011	USDDKK	5.28
DKK	07-01-2011	08-01-2011	USDDKK	5.77
NOK	29-03-2011	28-03-2011	USDNOK	8.34
DKK	31-12-2012	01-01-2013	USDDKK	5.65
DKK	21-09-2011	22-09-2011	USDDKK	5.48

Appendix G

Industries	N	% of sample
Brewers	28	9%
Apparel, Accessories and Luxury Goods	24	7%
Heavy Electrical Equipment	22	7%
Airlines	22	7%
Packaged Foods and Meats	19	6%
Pharmaceuticals	15	5%
Consumer Electronics	12	4%
Automobile Manufacturers	10	3%
Oil and Gas Equipment and Services	7	2%
Health Care Equipment	7	2%
Top 10 industries	**166**	**52%**
Other	155	48%
Total	321	100,0%

Appendix H

Company name	N	% of sample
Vestas Wind Systems A/S	22	7%
Carlsberg A/S	17	5%
PANDORA A/S	15	5%
Norwegian Air Shuttle ASA	11	3%
Novo Nordisk A/S	10	3%
Bang & Olufsen A/S	9	3%
IC Companys A/S	7	2%
Matas A/S	6	2%
Royal Unibrew A/S	5	2%
Coloplast A/S	4	1%
Top 10 companies	**106**	**33%**
Other	215	67%
Total	321	100,0%

Appendix I

No.	Hypothesis	Result
1	*$H1_0$: Analyst target prices are accurate and unbiased* *$H1_a$: Analyst target prices are inaccurate and biased*	*Reject H0*
2	*$H2_0$: Accuracy on buy recommendations = Accuracy on sell recommendations* *$H2_a$: Accuracy on buy recommendations ≠ Accuracy on sell recommendations*	*Reject H0*
3	*$H3_0$: Male analysts' accuracy = Female analysts' accuracy* *$H3_a$: Male analysts' accuracy ≠ Female analysts' accuracy*	*Fail to reject H0*
4	*$H4_0$: The study programme of the analyst does not influence the analysts' accuracy* *$H4_a$: The study programme of the analyst does influence the analysts' accuracy*	*Reject H0*
5	*$H5_0$: Analyst groups' accuracy = individual analysts' accuracy* *$H5_a$: Analyst groups' accuracy ≠ individual analysts' accuracy*	*Fail to reject H0*
6	*$H6_0$: Historic revenue volatility does not influence analysts' accuracy* *$H6_a$: Historic revenue volatility does influence analysts' accuracy*	*Reject H0*
7	*$H7_0$: The deviation of forecasted revenue growth to historical growth does not influence analysts' accuracy* *$H7_a$: The deviation of forecasted revenue growth to historical growth does influence analysts' accuracy*	*Fail to reject H0*
8	*$H8_0$: Historic NOPAT margin volatility does not influence analysts' accuracy* *$H8_a$: Historic NOPAT margin volatility does influence analysts' accuracy*	*Fail to reject H0*
9	*$H9_0$: The deviation of forecasted NOPAT margin to historical NOPAT margin does not influence analysts' accuracy* *$H9_a$: The deviation of forecasted NOPAT margin to historical NOPAT margin does influence analysts' accuracy*	*Fail to reject H0*
10	*$H10_0$: Historic ATO volatility does not influence analysts' accuracy* *$H10_a$: Historic ATO volatility does influence analysts' accuracy*	*Reject H0*
11	*$H11_0$: The deviation of forecasted ATO to historical ATO does not influence analysts' accuracy* *$H11_a$: The deviation of forecasted ATO to historical ATO does influence analysts' accuracy*	*Fail to reject H0*
12	*$H12_0$: The company's market capitalization at t0 does not influence analysts' accuracy* *$H12_a$: The company's market capitalization at t0 does influence analysts' accuracy*	*Reject H0*
13	*$H13_0$: The amount of multiples used in valuation does not influence accuracy* *$H13_a$: The amount of multiples used in valuation does influence accuracy*	*Fail to reject H0*
14	*$H14_0$: Analysts' model choice does not influence accuracy* *$H14_a$: Analysts' model choice does influence accuracy*	*Reject H0*
15	*$H15_0$: The length of forecast period does not influence accuracy* *$H15_a$: The length of forecast period does influence accuracy*	*Fail to reject H0*
16	*$H16_0$: The length of historical period does not influence accuracy* *$H16_a$: The length of historical period does influence accuracy*	*Reject H0*
17	*$H17_0$: Satisfying steady-state assumptions does not influences accuracy* *$H17_a$: Satisfying steady-state assumptions influences accuracy*	*Fail to reject H0*
18	*$H18_0$: Satisfying the growth sanity check does not influences accuracy* *$H18_a$: Satisfying the growth sanity check does influences accuracy*	*Fail to reject H0*
19	*$H19_0$: Changing WACC during forecast does not influences accuracy* *$H19_a$: Changing WACC during forecast does influences accuracy*	*Fail to reject H0*
20	*$H20_0$: The relationship between WACC and ROIC in terminal year does not influence accuracy* *$H20_a$: The relationship between WACC and ROIC in terminal year does influence accuracy*	*Reject H0*