JOURNAL OF UNDERGRADUATE RESEARCH

What Drives Growth: Analyzing Quantitative Factors and their Variation across Sectors

Daniel J. Sweeney
Warrington College of Business (Finance)

Faculty Mentor: David T. Brown, Department of Finance, Insurance and Real Estate

Abstract

Value investing is often considered the antithesis of growth investing. However, core to any value investing strategy is an intrinsic valuation, typically calculated using a DCF analysis. One of the most sensitive DCF assumptions is the estimation of a company's long-term cash flow growth rate. Thus, understanding a company's growth potential is a vital component in any value thesis. This paper attempts to create a quantitative model to help predict a company's long-term cash flow growth rate (using EBITDA growth as a proxy for cash flow growth) and to find the strongest indicators for a company's growth potential by sector. To do so, this study analyzes variables pertaining to operating efficiency, risk metrics, market valuation, corporate investment levels, and the competitive landscape for S&P 500 constituent companies. While all categories contributed at least one statistically significant variable, the market valuation and corporate investment level categories had the highest volume of significant variables. The results show that widely used quantitative metrics can help predict a meaningful portion of a company's five-year EBITDA growth rate when analyzed on a sector-by-sector basis. Furthermore, both the types of variables and predictive strength of the model varies widely across sectors. In practice, analysts should prioritize different ratios, metrics, and quantitative variables based on the target company's sector when estimating the trajectory of a company's long-term growth rate.

Keywords: long-term growth rate, DCF analysis

Introduction

A business's long-term cash flow growth rate is a core piece of its value today given the assumption that a business's financial value is the present value of all future cash flows that it generates for stakeholders. A discounted cash flow analysis (DCF), one of the most widely accepted valuation techniques, is very sensitive to assumptions for a company's terminal value. This terminal value is largely driven by assumptions for long-term cash flow growth. While the DCF and its inputs have been studied extensively, there is limited research and academic methodology to estimate a company's long-term growth (Tengulov et al., 2019). In fact, existing methods of future cash flow growth estimations are rudimentary, and these methodologies

largely lean upon educated guesswork. In general, estimating a company's long-term growth potential often relies heavily upon qualitative judgments.

This project explores quantitative factors that can be used to estimate a company's EBITDA growth rate. First, this study attempts to answer whether quantitative variables can be used to create a model that can predict a meaningful portion of growth. Second, it analyzes whether certain factors are stronger indicators of growth depending on the sector. Finally, the overarching goal is to provide guidance on which metrics analysts should focus on when making predictions about long-term growth potential.

Background

This study is largely inspired by *Valuation and Long-term Growth Expectations* by Angel Tengulov, Josef Zechner, and Jeffrey Zwiebel (Tengulov et al., 2019). Their study identifies the dearth of research on long-term growth estimations and uses pooled cross-sectional regression to analyze potential quantitative predictors for long-term growth. It identifies several factors that have statistically significant impacts on EBITDA growth across the company universe, including book to market ratio, capital expenditures, capital intensity, dividend yield, external financing, firm age, prior growth, industry exits, leverage, and size. Ultimately, their study culminated in the creation of a trading strategy based on their estimations of companies' growth potential. Many of the variables in the initial variable pool selected in this study are influenced by variables that Tengulov et al. analyzed. Tengulov et al. analyzed data from the entirety of US exchange-listed companies and did not filter for sector. Inspired by their study's general framework, this study analyzes a wider pool of variables, focuses on factors' impact for specific sectors, and has a narrower company universe (the S&P 500).

Several other sources provide literature that is instrumental in selecting variables to analyze. Considered a core resource on how to analyze a business's prospects qualitatively, *Common Stocks and Uncommon Profits* by Phillip Fisher helped inspire the variable groups that are analyzed in this study (Fisher, 2003). "How to Estimate the Long-Term Growth Rate in the Discounted Cash Flow Method" by Aaron Rostkowski and Evan Clough discusses processes and variables to consider when estimating growth rates from a legal viewpoint and notes that "none of these factors may be considered by itself" (Rostkowski et al., 2013). "The Economics of Short-Term Performance Obsession" by Alfred Rappaport contrasts variables that are prioritized

in the short-term with variables that matter in the long-term, which inspired the selection of several short-term-focused variables that could potentially have a negative impact on long-term growth (Rappaport, 2005). Finally, several variables are selected to reflect key industry characteristics that impact a company's long-term success as discussed in Porter's Five Competitive Forces framework (Porter, 2008).

Methodology

This study began with gathering a large pool of independent variables to be considered as possible predictors for growth. Then, it narrowed the pool down to ten independent variables for each sector. Finally, this study built a model to estimate growth rates both for the total market and for each sector.

Dependent Variable

Since the definition of "cash flow" varies across companies and time, this study uses EBITDA as a proxy for a company's cash flow. In addition, EBITDA is widely available and positive more frequently than operating income or net income. EBITDA is defined as net sales minus cost of goods sold (COGS) minus selling, general, and administrative expenses (SG&A). EBITDA growth is viewed over five-year periods; this time frame was selected to avoid a short-term analysis, maximize data points, and reflect a common forecast period length in DCFs.

Variable Selection

Based on both traditional valuation metrics and the previously discussed literature review, this study began with a very broad group of potential explanatory variables. The study then ran simple regression analysis between each variable and EBITDA growth. From here, the variable group was narrowed based on R-squared values, data availability, and anticipated overlap between metrics. The full list of variables is shown in Table 1.

Data

All data was downloaded from Compustat North America, which was accessed via WRDS. Specifically, approximately half of the variables were calculated from Compustat's Fundamentals Annual (N = 17), and half were sourced from Compustat's Financial Ratios – Firm Level (N = 20); one variable came from Compustat's Ratings database. The timeframe, largely selected based on data limitations, extends from 1970 to 2020. Data is organized in company-year data points in a panel data format. All variables above and below the 5^{th} and 95^{th}

Table 1. Independent Variables – Initial Group

Variable	Category	Number of Times Selected for Sector Model
Advertising Intensity	Competition	0
Altman's Z Score	Risk Profile	2
CAPEX Ratio (Capex/PPE)	Competition	5
Capital Intensity (PP&E/Total Assets)	Competition	0
Cash Conversion Cycle (Days)	Operating Efficiency	3
Cash to Market Cap	Operating Efficiency	0
Change in Working Capital	Operating Efficiency	4
Credit Rating (S&P)	Risk Profile	3
Current Ratio	Risk Profile	3
Debt to EBITDA	Risk Profile	1
Debt to Equity	Risk Profile	4
Dividend Payout Ratio	Investment	4
Dividend Yield	Investment	4
EV/EBITDA Multiple	Market Valuation	3
Gross Profit Margin	Operating Efficiency	3
Interest Coverage Ratio	Risk Profile	0
Inventory to Current Assets	Operating Efficiency	2
Inventory Turnover	Operating Efficiency	1
Long-Term Debt / Invested Capital	Risk Profile	2
M&A Expense to Market Value	Competition	0
M&A Revenue to Market Value	Competition	0
Market Cap	Competition	8
Price/Book Multiple	Market Valuation	7
Price/Earnings Multiple	Market Valuation	0
Price/Revenue Multiple	Market Valuation	4
Payables Turnover	Operating Efficiency	3
Pre-Tax Profit Margin	Operating Efficiency	3
Quick Ratio	Risk Profile	2
R&D Growth	Investment	3
R&D Intensity (R&D Expense/Sales)	Investment	0
Receivables Turnover	Operating Efficiency	2
Return on Assets	Investment	2
Return on Capital Employed	Investment	2
Return on Equity	Investment	2
Sales / Stockholders Equity	Investment	2
Sales / Working Capital	Investment	0
Shiller's Cyclically Adj. P/E Ratio	Market Valuation	6
Short-Term Debt / Total Debt	Risk Profile	0

percentile were winsorized to the value of the 5th and 95th percentile, respectively. Data points were measured at five-year periods for years ending in "0" or "5" and have both a company key (CUSIP) and year associated with them. The study did not use overlapping data to preserve the independence of observations.

Company Universe

This study analyzes S&P 500 constituents. The S&P 500 was selected as the index in order to analyze growth in relatively mature companies, the majority of which are expected to have positive EBITDA and relatively stable cash flows. In addition, the S&P 500 is an ideal index due to the widespread availability of data, which allows for the analysis of a wide field of independent variables. A company is defined by its CUSIP code, and all data is annual. Sectors are defined by Global Industry Classification ("GIC") codes. Financials, utilities, and real estate companies (GIC Codes 40, 55, and 60, respectively) were omitted from this analysis. In addition, companies must be in the S&P 500 for the entirety of five-year cross sections to have been included. Finally, companies missing greater than 80% of gathered data points were omitted.

Statistical Analysis

In order to narrow the initial variable group (N = 38) to be able to create a meaningful regression analysis, the study ran simple linear regression on each variable for eight sectors. A unique group of ten variables were selected for each sector based on the following criteria: highest adjusted R-squared, no closely related variables (e.g., both dividend yield and dividend payout ratio would not both be selected for a given sector, despite having the two highest adj. R-squared values), and no more than three variables from one variable category.

Once ten variables were selected for each group, pooled cross-sectional times series analysis was used to create a regression equation for each sector (with five-year EBITDA growth as the dependent variable). In the following discussion, a confidence level of 95% is used to define statistically significant variables.

Results

Both the models' predictive strength and the statistically significant variables vary widely across sectors. Furthermore, all sector-specific models have higher predictive capabilities than the total market model. Table 2 demonstrates the independent variables used in each sector and denotes which variables are statistically significant.

Table 2. Pooled OLS Regression Model Variables by Sector

Total Market	Comm. Services	Consumer Discretionary	Consumer Staples	Energy	Health Care	Industrials	Information Technology	Materials
Intercept***	Intercept"	Intercept**	Intercept*	Intercept	Intercept***	Intercept***	Intercept"	Intercept"
CAPEX Ratio	CAPEX Ratio	Credit Rating	Altman's Z Score	Credit Rating	CAPEX Ratio	Altman's Z Score***	CAPEX Ratio	Cash Conv Cycle
Credit Rating	Current Ratio"	Dividend Yield***	CAPEX Ratio**	Debt to EBITDA	Cash Conv Cycle"	Debt to Equity***	Cash Conv Cycle	Current Ratio
Debt to Equity	Dividend Payout Ratio	Debt/Invested Capital	Current Ratio*	Dividend Yield*	Dividend Payout Ratio	Div Payout Ratio**	Dividend Payout Ratio*	Debt to Equity*
Dividend Yield***	EV/EBITDA Multiple	Market Cap***	Debt to Equity*	Debt/Invested Capital	Gross Profit Margin	EV/EBITDA Multiple"	Gross Profit Margin	Dividend Yield*
Market Cap***	Inventory Turnover	P/B Multiple	Gross Profit Margin**	Market Cap***	Inv to Current Assets	Market Cap***	Inv to Current Assets	EV/EBITDA Multiple*
P/B Multiple	P/B Multiple*	Pre-tax Profit Margin	Market Cap***	P/B Multiple*	Market Cap*	P/B Multiple***	Market Cap***	Market Cap***
P/Rev Multiple***	P/Rev Multiple	Return on Assets	P/Rev Multiple**	Pre-tax Profit Margin	P/B Multiple	Payables Turnover*	P/B Multiple***	P/Rev Multiple
Quick Ratio	Return on Assets	Return on Capital*	Payables Turnover	Return on Equity*	R&D Growth	Quick Ratio	R&D Growth	Payables Turnover
R&D Growth	Return on Capital*	Return on Equity	Pre-tax Profit Margin	Sales / Stock. Equity	Shiller's Adj. P/E Ratio***	Shiller's Adj. P/E Ratio	P/B Multiple***	Receivables Turnover
Work. Capital Change	Shiller's Adj. P/E Ratio*	Shiller's Adj. P/E Ratio***	Receivables Turnover***	Shiller's Adj. P/E Ratio	Work. Capital Change	Work. Capital Change"	Work. Capital Change**	Sales / Stock Equity

Note: ***, **, *, and " denote p-values of 0.000, 0.001, 0.01, and 0.05, respectively.

Total Market (S&P 500, less Financials, Real Estate, & Utilities)

When estimating growth rates for the total market, the adj. R-Squared is 6.030%, lower than any sector-specific model. This provides evidence that growth rates can be estimated more accurately when looking at sector specific variables and models. Notably, each of the three significant variables came from a different group, with dividend yield (negative coefficient), market cap (negative), and price to revenue multiple (positive) being part of the investment, competitive landscape, and market valuation groups, respectively. Multiples, dividends, and size are three of the most common metrics for quickly identifying a company's growth prospects, so these results match conventional methods to gauge approximate growth expectations.

Communication Services

Communication Services has an adj. R-squared value of 12.547%. Compared to other sectors, this is in the mid-range and shows that a meaningful portion of EBITDA growth can be predicted with quantitative variables. Again, significant variables came from several different variable groups (market valuation, investment, and operating efficiency). Significant variables

were Shiller's P/E Ratio (positive coefficient), price to book ratio (positive), current ratio (positive), and return on capital employed (negative).

Consumer Discretionary

The Consumer Discretionary model also has a mid-range adj. R-Squared value of 12.715%, a modest but meaningful portion of growth. Significant variables include dividend yield (negative coefficient), market cap (negative), return on capital employed (negative), and Shiller's P/E Ratio (positive). Once again, variables from several different groups were significant: investment, competitive landscape, investment, and market valuation, respectively.

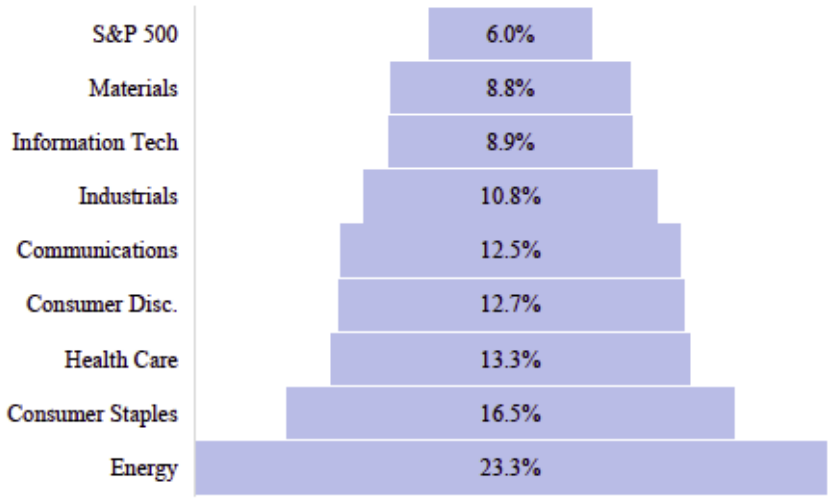

Figure 1. Predictive Strength as Measured by Adj. R-Squared

Consumer Staples

The Consumer Staples model has an adj. R-Squared of 16.522%, the second-highest sector. This confirms that conventional view that Consumer Staples is a relatively predictable sector. Most of the variables selected are statistically significant in this model: capex ratio (positive coefficient), current ratio (positive), debt to equity (negative), gross profit margin (negative), market cap (negative), price to revenue multiple (positive), and receivables turnover (positive). The coefficients' signs for all of the above variables match expectations, with the exception of gross profit margin.

Energy

Energy has the highest adj. R-squared value of any sector at 23.346%. Nevertheless, just four variables are statistically significant. Furthermore, the intercept was not significant, contrary to the results from all of the other sectors. Significant variables are dividend yield (negative coefficient), market cap (negative), price to book multiple (positive), and return on equity (positive). These variables appear closely related to variables that are significant in the S&P 500 model.

Health Care

The Health Care model has a relatively strong adj. R-Squared of 13.258%. This model has the lowest number of significant variables, with only three passing the 95% confidence level: cash conversion cycle (negative coefficient), market cap (negative), and Shiller's cyclically adj. P/E ratio (positive). This is the only variable group where a variable from the investment group is not significant.

Industrials

As an industry that is traditionally considered relatively predictable and stable, the Industrials sector has a very high number of significant variables (eight) despite a relatively low adj. R-squared value of 10.812%. Variables tied to capital structure or balance sheet strength were particularly important in this sector. With the exception of Altman's Z-score (negative coefficient), where a lower value indicates increased likelihood of bankruptcy, the signs of the variables' coefficients are as expected: change in working capital (positive), debt to equity (negative), dividend payout ratio (negative), EV/EBITDA multiple (positive), market cap (negative), price to book multiple (positive), and payables turnover (positive).

Information Technology

One of the weakest predictors, the Information Technology model has an adj. R-squared of 8.936%. Traditional metrics' weak relative strength at predicting technology's growth is expected given the impact of innovation and technological advancements in this sector. Despite weak strength, several variables still have significance: change in working capital (negative coefficient), dividend payout ratio (negative), market cap (negative), and price to book multiple (positive).

Materials

The Materials sector model posts the lowest predictive capabilities with an adj. R-squared of 8.795%, only slightly above the S&P 500 model. In addition, Materials has only four significant variables: debt to equity (negative coefficient), dividend yield (negative), EV/EBITDA multiple (positive), and market cap (negative). This sector's results are surprising, as Materials is typically considered a low-growth industry. However, cyclicality from price fluctuations likely makes this sectors' long-term growth more difficult to predict.

Conclusion

The results of this study suggest three key findings for consideration when constructing forward-looking growth predictions. First, the sector-specific models are more reliable than the total market model. For example, the Energy sector's model has an adj. R-Squared of nearly 4x the S&P 500 model's. Second, sector-specific models' predicative abilities vary widely across industries. Traditionally cyclical or volatile sectors such as Materials and Information Technology are the most difficult to predict, while traditionally stable sectors such as Consumer Staples and Energy are the most predictable. Third, relevant explanatory variables differ widely in type and strength across sectors, with Industrials having as many as eight significant variables and Health Care only having three.

Several variables stood out as particularly important indicators of growth potential. Company size (as measured by market cap) is the single most common significant variable, appearing in every model except for Communication Services. The statistical significance of market valuation multiples (e.g., price to revenue and enterprise value / EBITDA) in every model supports the idea that market sentiment is a reliable gauge of large cap companies' growth potential. Finally, a firm's commitment to paying profits to shareholders through dividends is a strong indicator of slowing growth: dividend yield or dividend payout ratio has a statistically significant and negative coefficient in six sectors.

A significant number of company-year data points are missing variables, making this unbalanced panel data. Due to missing data points, the results are difficult to back-test, as missing independent variables make out-of-sample analysis difficult. Out-of-sample analysis on company-year data points that do have all necessary independent variables shows high error (using both root mean square error and mean absolute error) for all models. This suggests that the models should not be used to determine the EBITDA growth rate for a specific company, but

rather that the models should be used to show analysts both which variables and which sectors have higher predictive strengths.

As the focus of this study was on finding which sectors are more predictable and which variables are relevant to which sectors, it focused on independent variables that are theoretically relevant to all sectors. Futures studies should consider adding sector-specific variables (such as EV/Capacity or EV/Subscribers). Further analysis could also analyze what factors cause some sectors to be more predictable than others (e.g. earnings or price volatility).

Acknowledgements

I would like to thank Dr. David Brown for his mentorship and guidance throughout the project and Peter McKay for his support with the data collection process.

References

Fisher, P. A. (2003). *Common stocks and uncommon profits and other writings by Philip A. Fisher*. Wiley.

Porter, M. E. (2008, January). The five competitive forces that shape strategy. *Harvard Business Review*. https://hbr.org/2008/01/the-five-competitive-forces-that-shape-strategy

Rappaport, A. (2005). The Economics of Short-Term Performance Obsession. *Financial Analysts Journal, 61*(3), 65-79. Retrieved January 30, 2021, from http://www.jstor.org/stable/4480672

Rostkowski, A., & Clough, E. (2013). How to estimate the long-term growth rate in the discounted cash flow method [PDF]. *Insights - Willamette Management Associates, Spring*. http://www.willamette.com/insights_journal/13/spring_2013_2.pdf

Tengulov, A., Zechner, J., & Zwiebel, J. (2019, November 18). *Valuation and long-term growth expectations* [PDF]. https://dx.doi.org/10.2139/ssrn.3488902