Developing a Multi-Factor Model to Identify Winning Growth Stocks

Nicholas Wong
Massachusetts Institute of Technology
nicwjh@mit.edu

Carlo Sun
Massachusetts Institute of Technology
carlosun@mit.edu

Haoxuan Zhang
Massachusetts Institute of Technology
haoxuan@mit.edu

Manas Goyal
Massachusetts Institute of Technology
manas_g@mit.edu

Abstract

We propose a framework for identifying and evaluating outperforming growth stocks through a multi-factor regression model. By leveraging a quarterly sliding window approach, our prediction model dynamically recalibrates its parameters to adapt to changing market conditions. We identify key growth-oriented financial indicators and rank stocks based on predicted forward 12-month returns. Performance is assessed through the construction of a portfolio using standard evaluation metrics such as the Sharpe Ratio and Hit Ratio. Results reveal consistent excess returns in selected portfolios, underscoring the efficacy of the proposed approach. We additionally evaluate factor significance and identify Debt-to-Equity Ratio and Price-To-Book Ratio as robust predictors. The proposed framework offers a data-driven strategy for navigating growth equity.

1 Introduction

Growth stocks are companies expected to grow at an above-average rate compared to other publicly traded companies.

The common characteristics of growth stocks are:

- **High Revenue Growth:** Growth companies exhibit rapid sales growth, often outperforming their industry peers.

- **High Multiples:** Valuation metrics like P/E, and price-to-book (P/B) ratios tend to be higher, reflecting investor expectations for future growth.

- **Low or No Dividends:** Instead of distributing dividends, growth companies reinvest earnings to fuel future growth.

- **Growing Market:** Growth companies are often found in emerging or fast-growing industries like technology, biotech, or renewable energy.

- **Capital Gains Investors:** Growth companies attract investors who are looking to profit from capital gains.

- **High Volatility:** Due to optimism about future performance, these stocks can be more volatile and sensitive to market conditions.

The difference between growth and value stocks can be summarized in the table below.

	Growth Stock	Value Stock
Focus	revenue growth & future earnings	current intrinsic value compared to price
Valuation	high P/E and P/B ratios	low P/E and P/B ratios
Dividends	rarely pays regular dividends	often pays regular dividends
Industry	fast-growing sectors	mature or cyclical sectors
Investors	attract investors seeking capital gains	attract income-oriented or conservative investors
Market	outperform in a bullish market	perform better during economic downturns

2 Characteristics

Growth stocks can offer substantial returns, especially driven by companies that are reinvesting earnings to increase growth. It must be noted that this comes with volatility and therefore, also risk, which is why investors need to assess their risk tolerance, investment horizon and conduct thorough due diligence before engaging in any investments. A diversified portfolio and staying informed about the market allow for better navigation of growth investing, while being able to achieve financial goals.

Growth stocks often exhibit several key traits:

1. **High P/E Ratio**: Investors willing to pay a premium for anticipated future growth – resulting in higher P/E ratios compared to value stocks.

2. **Low or No Dividends**: Profits are typically reinvested to fuel growth, leading to minimal or non-existent dividend payouts.

3. **Innovative Product/Service**: Operate in evolving industries, offering novel products or services that drive growth.

4. **Market Leadership**: Many growth companies hold dominant positions in respective markets, benefiting from economies of scale (EOS) and brand recognition.

5. **Potential for High Returns**: The incentive for early investors lies in the potential for higher payoffs in the future. For example, early investors of Apple and Amazon have reaped large positive payoffs.

Risks and Volatility

1. **Market Sensitivity**: Growth stocks are mostly more volatile than value stocks, as their performance is affected by market sentiment and economic conditions like changes in the interest rate environment. During times of rising interest rates, the present value (PV) of future earnings decreases, leading to a decline in valuations of growth stocks.

2. **Execution Risk**: The high expectations placed on growth companies mean that any failure to meet projected growth rates can result in sharp stock price declines. Operational challenges, increased competition, or regulatory hurdles can hinder a company's growth trajectory.

3. **Valuation Concerns**: Due to their high P/E ratios, growth stocks may be more susceptible to price corrections if the anticipated growth does not materialize. Investors may realize that they have overpaid for future prospects, leading to a reevaluation of the stock's worth.

Historical Performance Trends

1. **Bull Markets**: During periods of economic expansion, growth stocks often outperform the broader market. Optimism stemming from investors and an increase in risk drives up valuations, especially for companies with growth prospects that are promising (like Rivian in 2022 – having a multi-billion valuation, more valuable than legacy car makers like Audi, but not having produced a single car).

2. **Bear Markets**: Conversely, in economic downturns, growth stocks may underperform as investors seek safety in more stable, dividend-paying value stocks. The higher volatility of growth stocks can lead to more pronounced declines during market contractions.

Especially in the tech and healthcare sector, growth stocks have been very prominent, as innovation and capital has been increasing at a high pace. In biotech, for example, Merck's acquisition of Prometheus with a significant three-digit premium in percentages allowed for substantial increases in stock prices.

Key Considerations for Investors

1. **Risk Tolerance**: Investors should look at their ability to withstand volatility and losses. Growth stocks can experience significant price swings, and not all investors are comfortable with this level of risk.

2. **Investment Horizon**: A longer-term perspective can help short-term fluctuations. Investors with a time horizon of 5-10 years might have a better expected outcome.

3. **Diversification**: Diversifying between growth and value stocks should balance potential risks and returns.

4. **Fundamental Analysis**: Conducting thorough DD is important, especially when it comes to evaluating a company's management, competition, financial health and growth strategy.

5. **Market Conditions**: Staying informed about macroeconomic factors, such as interest rates and inflation, can help investors anticipate how growth stocks might perform under different economic scenarios.

3 Challenges

As discussed in previous sections, growth stocks face several challenges that can impact their performance. These challenges include high valuations, cyclic performance, and sensitivity to economic changes. Understanding these challenges is key to navigate the complexities of growth investing and to generate long term returns.

High Valuation

- Growth stocks often trade at a high P/E multiple compared to value stocks, as investors expect a strong future growth reflecting a premium. However, it can lead to vulnerability if these expectations are not met. For example, during the boom in the late 1990s, technology companies saw their valuation increasing drastically. When they failed to meet those forecast, it took them years to recover from their inflated valuations.

Cyclic Performance

- The performance of growth stocks is cyclical, often related to the broader economic cycles. Historically, periods of economic expansion tend to lead to over-performance, and periods of economic recession tend to lead to under-performance. For example, during the Covid-19 pandemic and Russia-Ukraine war, growth stocks were more volatile and experienced losses.

Economic Changes

- **Interest Rate:** Growth stocks are very sensitive to changes in interest rate. A higher interest rate implies a higher discount rate for its future earnings, thus reducing its present value. Moreover, a higher interest rate also puts pressure to growth companies borrowing power, making it difficult for growth companies to implement growth plans.

- **Macroeconomic Factors:** Besides interest rate, macroeconomic factors such as inflation also affect the performance of growth stocks. For example, inflation pressure could cut consumer purchasing power, leading to a lower sale for companies that rely on consumer discretionary spending.

4 Selection of Factors

To select growth stocks, we identify the following factors that correspond to the characteristics of growth stocks.

Revenue Growth

- It is the year to year percentage increase in a company's top line revenue. Revenue growth is a fundamental indicator of business expansion and market demand for the company's products or services. Hence, a high revenue growth is a signal that the company is successfully

gaining market share or selling at premium. Therefore, a high revenue growth usually indicates a growth stock.

EBIT Growth

- It is the year to year percentage increase in a company's operating margin (earning before interest and taxes). EBIT growth demonstrates a company's ability to improve its operation and profitability as it scales up. Hence, a high EBIT growth is a signal that the company is successfully operating efficiently and its core business is thriving. Therefore, a high EBIT growth usually indicates a growth stock.

Debt to Equity Ratio

- It is a measure of financial leverage calculated as total debt over total shareholders equity. A high debt to equity ratio means that the company is taking a lot of debt, compared with equity, mostly used to fund capital expenditure. Growth stocks, such as technology and biotech companies, require significant debt financing to fuel their rapid growth. Therefore, a high debt to equity ratio usually indicates a growth stock.

Price to Book Ratio

- It is the ratio comparing the market value of equity and the book value of equity. The market value is market capitalization, computed as share price multiplied by shares outstanding. The book value can be found in the company's balance sheet. Those two values are in general different, because market value reflects how investors think the future earning potential of the company. A high price to book ratio demonstrates that the market anticipates higher earnings. Therefore, a high price to book ratio usually indicates a growth stock.

Price Earning Ratio

- It is the ratio comparing the market value of equity and the net income, or in other words, the share price and the earnings per share. It is profitability ratio that assesses the relative value of a company's stock. A high price earning ratio reflects investors' optimism about the company's future growth potential and hence are willing to pay more for the return they get. Therefore, a high price earning ratio usually indicates a growth stock.

5   Data

We leverage a dataset based on the Russell 1000 Growth Index, a benchmark comprising large-cap U.S. equities with growth-oriented fundamentals, such as high earnings growth and strong revenue expansion. This dataset is particularly well-suited for our project, as it inherently focuses on companies with attributes traditionally associated with growth investing, making it an ideal foundation for identifying outperforming growth stocks. By working within this growth-oriented universe, we aim to refine and enhance existing methodologies by systematically quantifying key growth metrics, such as EBIT and revenue growth, while also incorporating valuation metrics to account for relative attractiveness. Leveraging the Russell 1000 Growth Index allows us to target high-potential equities within an established growth framework, ensuring alignment with our project's objectives of systematically identifying future outperformers in the growth equity space.

5.1   Preprocessing

We present a brief overview of our methodology for data preprocessing (pivoting from wide to long, feature selection, data imputation, winsorization, feature engineering, normalization) in this section. Detailed implementation details can be found in the appendix at 7.

The dataset contains data for stocks in the Russell 1000 Growth Index for years 2020-2024 and was obtained in a wide-format time series.

Preprocessing included pivoting data into a long-format structure where each row represented a unique combination of stock, metric, and date. The long-format data was then restructured using

pivot_table [1] to create one row per stock per date with stock metrics as individual columns to enable feature engineering and granular analysis.

Only the metrics pertinent to our analysis were retained:

1. EBIT
2. EBITDA
3. Price Earnings Ratio (P/E)
4. Price to Book Ratio
5. Revenue
6. Debt-to-Equity Ratio

Missing values were imputed through feature-specific imputation strategies (backfill, mean imputation). We applied backward filling to time-continuous features like *Last Price*, *Price to Book Ratio*, *Revenue*, *Debt to Equity Ratio* to maintain temporal continuity. For features that lacked time continuity but required stable central tendencies like *EBIT*, *EBITDA*, *Price Earnings Ratio (P/E)*, we used mean imputation instead.

Outliers in these key financial metrics were capped at the 1st and 99th percentiles using winsorization to reduce the influence of extreme outliers.

Growth-oriented features, such as year-over-year percentage changes in EBIT and Revenue, were engineered to quantify growth trends and for use in our predictive model. Forward 12-month returns were also computed for use in our predictive model as a target variable.

Predictive features were then normalized using scikit-learn's *StandardScaler* [2] to have zero mean and unit variance to mitigate bias in regression coefficients due to disparate feature scales.

6 Model

We define our objective to be identifying unidentified growth stocks that will outperform a specified benchmark over a 12-month investment horizon.

We propose a quarterly sliding window linear regression framework for predicting forward 12-month returns, focusing on the evaluation of growth stock outperformance in a quarterly analysis. Our approach involves training a linear regression model over a fixed training window of historical quarters and validating predictive capability on the subsequent quarter. By iteratively training our model on a rolling historical window, our model dynamically recalibrates its parameters based on most recent data to ensure adaptability and robustness.

The linear regression model used to predict the forward 12-month return is formally defined as:

$$y_i = \beta_0 + \beta_1 X_{i,1} + \beta_2 X_{i,2} + \beta_3 X_{i,3} + \beta_4 X_{i,4} + \beta_5 X_{i,5} + \epsilon_i$$

Where:

- y_i: The forward 12-month return for stock i.
- β_0: The intercept term of the regression model.
- β_j: The regression coefficient for the j-th feature ($j = 1, 2, \ldots, 5$).
- $X_{i,j}$: The value of the j-th feature for stock i. Features include:
    - $X_{i,1}$: EBIT Growth
    - $X_{i,2}$: Revenue Growth
    - $X_{i,3}$: Total Debt to Total Equity
    - $X_{i,4}$: Price to Book Ratio
    - $X_{i,5}$: Price Earnings Ratio (P/E)
- ϵ_i: The error term, assumed to be normally distributed with $\epsilon_i \sim \mathcal{N}(0, \sigma^2)$.

The sliding window approach partitions the dataset into overlapping train and test sets defined by quarterly intervals. Each iteration of our sliding window uses data from the most recent four quarters for training and tests it on the subsequent quarter.

The primary objective of this model is to construct top-decile portfolios and evaluate their performance with relevant financial metrics.

6.1 Performance Evaluation Metrics

Using the predicted returns from the trained model, stocks are ranked within each quarter. The top-decile stocks form a portfolio and their performance is compared against a market benchmark (mean return in the universe of stocks in the Russell 1000 Index). We use the following to compare portfolio performance:

Metric	Definition
Mean Excess Return	The average return of the top-decile portfolio minus the benchmark return, calculated across all test periods. This metric indicates the absolute outperformance of the selected portfolio relative to the benchmark.
Sharpe Ratio	A measure of risk-adjusted performance, calculated as the mean excess return divided by the standard deviation of excess returns. A higher Sharpe Ratio indicates better returns for a given level of risk.
Hit Ratio	The percentage of test periods in which the top-decile portfolio outperforms the benchmark. This metric quantifies the consistency of outperformance.

Table 1: Portfolio Evaluation Metrics

For each training window, we also track the R^2 to assess model fit and the significance of individual predictors via t-statistics and p-values.

6.2 Investment Strategies and Results

Using our factor model, top decile portfolios consistently demonstrate higher average returns relative to our benchmark:

Metric	Value
Mean Excess Return	2.14% (annualized)
Sharpe Ratio	0.36 (risk-adjusted return)
Hit Ratio	71.4% (reflecting frequency of outperforming the benchmark)

Table 2: Performance Metrics for Top-Decile Portfolios

We further illustrate the superior performance of the top decile portfolio with a cumulative performance plot, showing the compounded growth of a $1 investment in the top-decile portfolio and the benchmark.
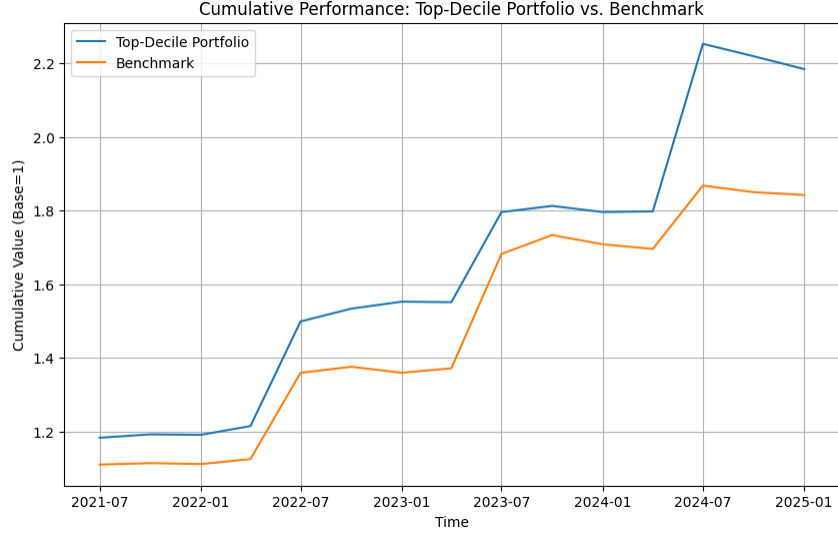
Figure 1: Performance Comparison of Top-Decile Portfolio vs. Benchmark

We see that, despite a 71.4% hit ratio, the ability of the top-decile portfolio to consistently generate positive excess returns on average allows it to perpetually outpace the benchmark.

Using this multi-factor model, we propose the following investment strategy:

1. Apply the multi-factor linear regression model to the growth stocks on a monthly basis adjustment, so that we can get the predicted returns for each individual stock from the model.

2. Based on the predicted returns from previous step, we classify the stocks into 10 groups with an equal number of stocks, and we will long the group with the highest stock values to construct portfolios in each rolling window. We further assume that we will invest equal amount of the money to each individual stock, so that the return of the portfolio is the mean of the individual stock returns

3. In each month, we can calculate the actual profit and loss of the constructed portfolio and compare with the benchmark performance by their returns, and we can also compare the evaluation metrics listed above.

If we relax the long-only constraint, we can further consider a long-short strategy. We long the top-decile stocks and short the bottom 10-decile stocks based on predicted value. Implementing this strategy, we have:

The performance metrics are given by:

Metric	Value
Mean Excess Return	6.7% (annualized)
Sharpe Ratio	0.61 (risk-adjusted return)
Hit Ratio	80% (reflecting frequency of outperforming the benchmark)

Table 3: Performance Metrics for L-S Portfolio

Superior portfolio performance as measured by excess return, risk-adjusted return and hit ratio are achieved by utilizing our investment strategy and relaxing the long only constraint. The relative decile performances are given by the time series plot:
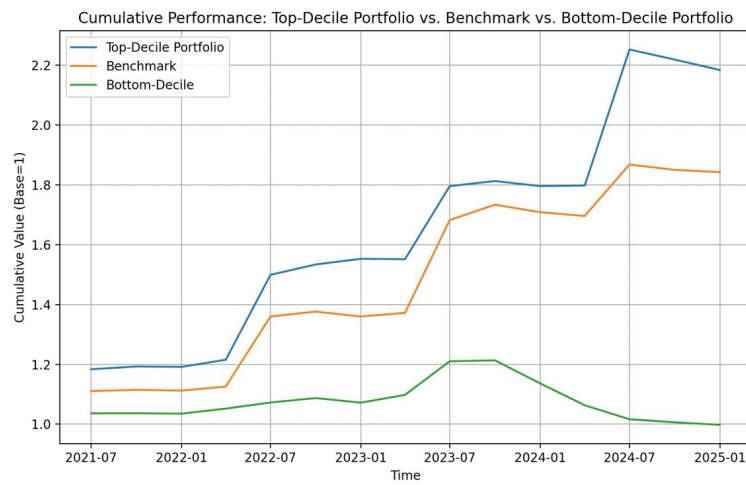
Figure 2: Performance of Top-Decile Portfolio vs. Benchmark vs Bottom Decile Portfolio

7 Findings and Recommendations

Investing in growth stocks is known to be a strategy for investors that want to seek high returns, as these company shares are expected to increase revenues and earnings that outpace the overall market. These companies reinvest earnings into expansion, and mostly positive Net Present Value (NPV) projects, to create new products, and entering new geographical markets or engaging in M&A. The focus on reinvestment rather than a payout policy is intended to drive growth and increase the value of the underlying over time.

In this work, we identify five key factors for growth stock selection (Revenue Growth, EBIT Growth, Debt to Equity Ratio, Price to Book Ratio, Price Earning Ratio) as outlined in 4 and develop a multi-factor model to identify winning growth stocks, utilizing a quarterly sliding window approach to build a top-decile portfolio that seeks to outperform over a 12-month investment horizon (via 12-month forward returns).

We construct a top-decile portfolio using this multi-factor model and find that it consistently outperforms the benchmark, generating superior excess returns with a 71.4% hit ratio.

We further propose an investment strategy, outlined in 6.2, where we construct a long-short portfolio and find that greater excess returns can be obtained by relaxing the long-only constraint. As long-short equity is a standard industry investment strategy, we further show the robustness of this decile-based approach to portfolio construction using our multifactor model.

In terms of factor significance, across all testing windows, we find that consistently significant factors (p-value < 0.05) were Debt-to-Equity Ratio and Price-To-Book Ratio. Debt-to-Equity Ratio was consistently significant across many training windows, suggesting it is a robust predictor of forward 12-month returns. Price-To-Book Ratio is another consistently significant factor, highlighting its importance in identifying undervalued growth stocks. Price-to-earnings ratio was significant in some windows, though it did not have as consistent an influence as the prior two factors. EBIT growth and revenue growth were not consistently significant, suggesting that their predictive power may be limited in this context.

For growth stock identification, we recommend focusing on Debt-to-Equity Ratio and Price-To-Book Ratio as key identifiers as they consistently demonstrated statistical significance and are the most reliable predictors in the framework we implement. Price-to-earnings still has value, particularly in specific contexts where valuation metrics are more relevant. Its use should be considered strategically. Lastly, while we find the EBIT growth and revenue growth are not reliably significant, more robust testing is needed to check if this is simply an artifact of our data. Transformations and interactions may also enhance their impact.

While our implementation of a quarterly sliding window framework to identify outperforming growth stocks provides a robust baseline for performance evaluation, there are opportunities for refinement. For one, we implement this model with a static feature set and feature relevance may vary across economic cycles. Incorporating some form of recursive feature elimination technique to vary the static feature set could improve the model's sensitivity to temporal variations. Additionally, regarding our top-decile portfolio construction strategy, while it is effective, it may not fully utilize the richness of the predicted return distribution. Incorporating multi-decile analyses could offer a more nuanced perspective on model efficacy. We could also validate portfolio construction using alternative ranking criteria, such as volatility-adjusted returns to mitigate potential biases.

References

[1] *pandas.pivot_table; pandas 2.2.3 documentation — pandas.pydata.org.* `https://pandas.pydata.org/docs/reference/api/pandas.pivot_table.html`. [Accessed 06-12-2024].

[2] *StandardScaler — scikit-learn.org.* `https://scikit-learn.org/dev/modules/generated/sklearn.preprocessing.StandardScaler.html`. [Accessed 06-12-2024].