Data Bootcamp Final Project presentation

Edmon Tong

Monte Carlo ROIC-based DCF modelling

Introduction

In this project, we intend to create a 10 year horizon dcf model for valuation of PEPSICO, using Monte Carlo simulations and ROIC_based dcf. We extract PEPSICO's past 75 years of financial data, 1950-2024 and insert it to a Monte Carlo layer for simulation. We then take the monte carlo layer simulated financials (Ebit margin, ROIC, Reinvestment Rate) and plug in ROIC based dcf model which uses ROIC*Reinvestment Rate for revenue growth rate. As we simulate 5000 times, we obtain 5000 distinct valuations simulating the company undergoing various circumstances for the following 10 years, and their potential equity value under such circumstances.

We desire to obtain a PEPSICO valuation that is more accurate and accounts for various scenarios compared to a traditional, simple DCF For our conclusion. In our findings, we present the mean & median equity value, mean & median per share price, and a 90% confidence interval valuation. We also compare our findings to a simple dcf's valuation, as well as current equity value and per share price.

Data Description

The data we use in this project all derive from WRDS, Compustat database, “a vast database offering standardized fundamental financial, statistical, and market data for active and inactive public companies globally, drawing from official SEC filings like 10-Ks and 10-Qs”.

For this project, we extract: SALE, total revenue; EBIT, operating profit before interest & taxes; DP, depreciation; CAPX, capex; TXT, tax expense; AT, total assets; CHE, cash balance ; ACT / LCT, short term debt, long term debt; DLTT / DLC, debt components; CSHO: shares outstanding; PRCC_F: price per share.

We organize these data into a dataframe which we name "comp" and fill its nan values in PRCC_F with mean. We then use these extracted raw data to compute for data we require for monte carlo simulation and ROIC-based dcf modeling – Revenue; Ebit margin; ROIC; Reinvestment; RR; NOPAT. Along with Revenue, fiscal year, tax rate, ebit, invested capital, we organize these data together into a dataframe we name "driver", and plug these drivers into our model.

datadate	fyear	sale	ebit	ebit_margin	tax_rate	nopat	invested_capital	roic	reinvestment	rr
1951-12-31	1951	46.8	3.97	0.084829	0.392947	2.41	26.8	0.089925	7.3	3.029046
1952-12-31	1952	53.3	7.77	0.145779	0.508366	3.82	25.4	0.150394	4.9	1.282723
1953-12-31	1953	66.1	11.14	0.168533	0.530521	5.23	26.7	0.19588	5.08	0.971319
1954-12-31	1954	74.2	12.99	0.175067	0.499615	6.5	30.7	0.211726	7.31	1.124615
1955-12-31	1955	89.0	18.36	0.206292	0.513072	8.94	37.7	0.237135	8.24	0.9217
1956-12-31	1956	97.0	17.05	0.175773	0.527859	8.05	48.9	0.164622	8.57	1.064596
1957-12-31	1957	120.3	18.95	0.157523	0.533509	8.84	61.3	0.144209	13.54	1.531674
1958-12-31	1958	136.8	24.32	0.177778	0.47903	12.67	66.2	0.19139	12.1	0.955012

(portion of "drivers" dataframe)

Models and Methods

In this project, we coded for our own Monte Carlo ROIC-based dcf model. Since ROIC-based dcf requires ROIC, RR, Ebit margin for calculation of free cashflow, we simulates these three data with our monte carlo layer by using numpy to apply random amounts of standard deviation to such data's historic mean based on 75 years of historical data extracted, simulating the company undergoing various scenarios and impact to their equity value in the following 10 years. For our assumptions, we assume that our forecast is built upon the latest reported revenue; WACC = 8%, and terminal growth rate of 2%. These assumptions can be easily altered.

We run 5000 simulations, providing us with 5000 distinct ROIC-based dcf valuations.

Results and Interpretation

Our model generates the according predictions for PEPSICO's valuation:

<u>For Equity valuation</u>

Mean EV (millions): 147,225.05; Median EV (millions): 141,240.09;

5% EV (millions): 84,401.39; 95% EV (millions): 232,477.30

<u>For Per chare price valuation</u>

Mean price ($/share): 79.26; Median price ($/share): 74.89;

5% price ($/share): 33.47; 95% price ($/share): 141.39

In comparison, we create and run a *simple dcf model* which assumes constant growth rate on free cashflow. This simple model generates predictions of:

<u>For Equity valuation</u>

Equity value (millions): 162,740.76

<u>For Per chare price valuation</u>

price ($/share): 118.62

Furthermore, we also provide PEPSICO's *current valuation* on Dec. 10th, 2025

For Equity valuation

Equity value (millions) = 203,790 millions

For Per chare price valuation

Per share price = 149.04 $

Comparing our model to a simple dcf, our model produces significantly lower valuation on both mean and median equity and per share price estimate. Our model's 90% interval, however, captures the simple dcf's prediction easily.

Nevertheless, it is worth noting that PEPSICO is currently trading significantly above both traditional model's and our model's valuation, and even our model's 90% confidence interval fails to capture it at its upper bound. potentially suggesting that PEPSICO is trading at a significantly higher price than it should be, further indicating a possible short for traders.

Conclusion and Next Steps

In conclusion, valuation by our model and a simple dcf model both suggests that PEPSICO is trading at a significantly higher price than it should be as the average per share price estimated by models is 98.9 while the stock is currently traded at 149, further indicating a possible short for traders.

Nonetheless, our model's drastically lower valuation compared to that of a simple dcf raises questions. A possible reason for this is that we are using the company's data for the past 75 years, very possibly capturing the company changing regime and improving its business model in this long period of time. For example, PEPSICO probably has increased and stabilized its Ebit margin, ROIC in recent years and it is extremely unlikely that the company could return to its situation decades ago, meaning that data from a long period of time ago could be irrelevant in our prediction model and pollute our results, which is most likely taking form in an unreasonably lower valuation, that we actually observe. As a result, I think one major change our model must implement is capturing a shorter period of time for the monte carlo layer, avoiding data pollution by, for example, selecting only the most recent 10-15 years.