# Evaluating the Risk-Return Trade-Off within the S&P500 with reference to Sector-Specific dynamics

**Executive Summary**

This project investigates the risk-return trade-off within the S&P 500, focusing on how risk, as measured by volatility, influences stock returns. The project goes on to further investigate how sector-specific risk impacts stock returns.  The study tests two key hypotheses derived from the Capital Asset Pricing Model (CAPM) and sector-specific dynamics:

1. **Stocks with more volatility than the market deliver average daily returns that are at least 20% higher than stocks with equal/lower volatility than the market, consistent with the Capital Asset Pricing Model.**

Supported by the data. High-beta stocks showed significantly higher average daily returns compared to low-beta stocks. However, beta alone is not a robust predictor of returns – inconsistent with the CAPM.

1. **Sector is the most important feature in determining average daily returns due to its influence on risk profiles.**

Not supported. Sector was not the most significant factor in predicting returns. Instead, volatility, dividend yield, and beta emerged as stronger predictors, with sector playing a secondary role.

The analysis revealed that while CAPM’s prediction of a positive relationship between beta and returns holds to some extent, the model’s linear assumptions are insufficient to fully explain stock returns. There is no overarching factor -whether risk or sector- to entirely model returns. Stock returns can be thought of as broader trends and multi-factor models, as the risk-return relationship is influenced by a variety of factors.

**Introduction and Motivation**

The risk-return trade-off is a fundamental principle of investment, that states potential returns increase with an increase in risk (Chen, J., 2024). This project investigates this relationship within stocks that are components of the S&P 500, focusing on how sector-specific risk metrics such as volatility influence returns.

The Capital Asset Pricing Model (CAPM), developed by William Sharpe and John Lintner in the 1960s, formalises the risk-return trade-off by establishing a **linear** relationship between the risk and expected return of an asset. Sharpe presented *A theory of market equilibrium under conditions of risk* where he argued that a rational investor can only obtain a higher expected rate of return on their investments by incurring additional risk (Sharpe, W., 1964) as represented in Figure 1, where the capital market line illustrates the trade-off between risk and return. Beyond the interest rate (risk free rate of return), investors incur additional risk for higher expected rates of return.

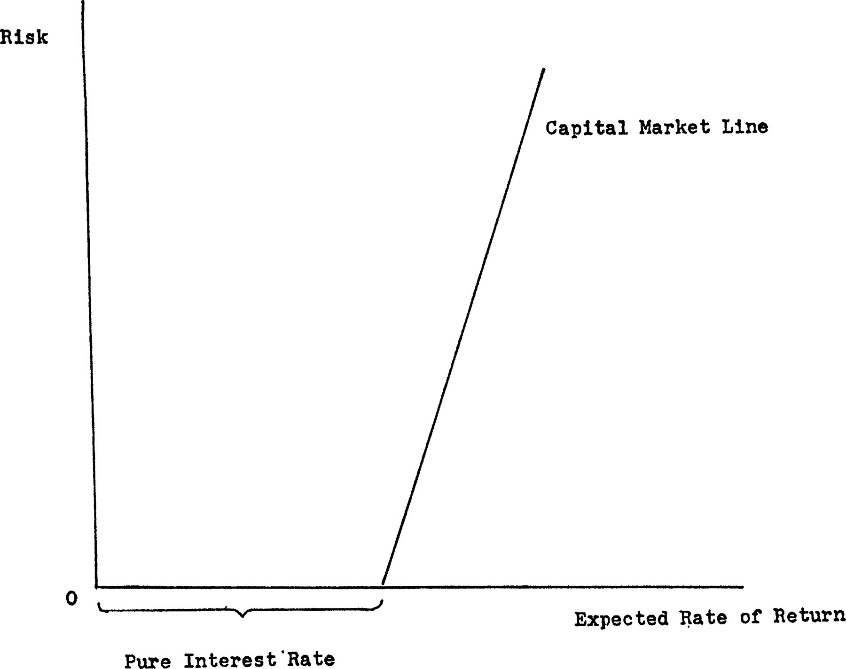


Figure : A representation of William's Sharpe's view of the capital market from **A theory of market equilibrium under conditions of risk** (Sharpe, W., 1964)

The CAPM formula is expressed as:

(2.3% for this study, based on 5-year Treasury yields from 2018)

(volatility relative to the market)

(expected return of the market minus the risk-free rate)

**Beta** (β) measures an asset’s volatility relative to a benchmark such as the S&P500 representative of behaviour of the general market. A beta of greater than 1 indicates higher volatility than the market while a beta less than 1 suggests lower. *Investors often target high beta, taking on more risk in hopes to gain from the volatility.*

This project tests two key hypotheses derived from CAPM and sector-specific risk-return dynamics:

1. **"Stocks with a beta higher than 1.0 deliver average daily returns that are at least 20% higher than stocks with a beta lower than 1.0, consistent with the Capital Asset Pricing Model prediction."**
2. **“As Sector is a key determiner of risk profile, it acts as the most important feature in determining average daily returns”**

These hypotheses are rooted in financial theory, and the principles of the risk-return trade-off. The first hypothesis addresses the theory that greater systematic risk should offer higher returns as compensation. The second hypothesis builds on the observations of similar sector dynamics, where companies often exhibit similar risk-return profiles due to factors such as regulatory environments and technological advancements.

For instance, the Technology sector, characterized by rapid innovation and high growth potential, often experiences wide dispersion in returns, presenting both opportunities and risks for investors (Cutter, P., 2024). Similarly, sectors like Energy and Consumer Discretionary exhibit distinct risk profiles driven by commodity price fluctuations and consumer demand cycles, respectively.

By analysing data from the components of the S&P 500 for 5 years (2020-2024) this study employs various statistical and machine learning techniques to examine the validity of these hypotheses and provide a data driven perspective on the relationship between risk and return.

**DATA AND METHODS**

The dataset consists of individual records for companies listed in the S&P 500, an index fund of leading publicly traded companies in the US. The dataset stock information, risk metrics, return measures, and financial stock metrics. The analysis employed statistical and machine learning techniques, including Pearson’s correlation, linear regression, ANOVA, and random forest feature importance analysis.

A screenshot of a computer

Description automatically generated

**Analysis of CAPM Theorem, modelling returns as a simple linear equation**

A screen shot of a graph

Description automatically generated

Figure : Scatter Plot of beta plotted against average daily returns, colour-coded by sector

**Beta vs Average Daily Return**

**Linear Correlation Analysis of Beta and Average Daily Return**

Pearson’s Product Moment Correlation Coefficient was used to measure the strength of the linear relationship between beta and average daily return. This analysis aimed to quantify the strength of the presumed positive relationship, supporting evidence for the CAPM’s prediction that higher beta stocks yield higher returns.

The PMCC analysis returned a coefficient of 0.401 and a p value of 2.306 x 10-19 at the 5% one tailed significance level, which is statistically significant. This test indicates a moderate positive linear correlation, suggesting that to an extent, stocks with higher beta tend to have higher average daily returns.

**Linear Regression of Beta and Average Daily Return (AIC -5949.)**

To further explore this linear relationship between beta and average daily return, a Ordinary Least Squares simple linear regression model was employed. Since CAPM suggests that returns are driven by beta, beta is considered our independent variable (x) and average daily returns the dependent variable (y).

The slope of the regression line suggests that for every 1% increase in beta, the average daily return increases by 0.007% per day. The p value of 0.000 suggests that this value is statistically significant. The intercept is not statistically significant (p = 0.707), and therefore when beta is 0, the expected return is not meaningfully different from 0, which aligns with financial theory - cash is considered to have a beta of 0.

The simple regression line becomes**:** Average daily return = 0.0007 x Beta

A graph with blue dots and red line

Description automatically generated

Figure : Simple Linear Regression Model (OLS) for Beta against Average Daily Return with its 95% confidence interval

The R-squared statistics of 16.8% suggested that beta alone does not strongly predict returns, and additional factors could play a role. Given the moderate effect size of beta as a predictor of average daily returns, **it does not appear that systematic risk is linearly correlated to higher returns in the way proposed in the CAPM theorem**. The CAPM formula has beta as a variable while all other factors remain as constants - if this were the case, we would expect a larger effect size.

**Addressing Hypothesis 1**

A diagram of a group of blue rectangular objects

Description automatically generated with medium confidence

Figure : Box Plot demonstrating the differences in return for low beta (<= 1) and high beta (>1) companies

**Analysis of Variance (ANOVA) and T-Statistic Testing**

An ANOVA was conducted to determine if there was a significant difference between the two groups. The F statistic, 50.143 was significant with a p value of 5.23e-12. A T test was used to get the absolute mean difference, and within the 95% confidence level, the lower bound still represented a 43.06% difference between the two. As this exceeds the 20% increase in the hypothesis, the hypothesis is supported by the data for the most part, however this may not be entirely consistent with the CAPM model as beta alone does not reflect a strong linear relationship with returns.

**Addressing Hypothesis 2**

**Assumption: Sector as a key determinant of systematic risk**

**A comparison of a graph

Description automatically generated with medium confidence**

Figure : Box plots to show the distribution of volatility and beta by Sector

The most volatile sectors (beta), by mean are Information Technology (1.211372), Consumer Discretionary (1.166521) and Energy (1.127928). *The contrast in risk profiles across sectors suggest the possibility that sector could be a significant indicator of risk and therefore of returns.*

**Multiple Linear Regression (AIC -5976.)**

Following the results of the linear regression, multiple regression was employed to include sector dummies as predictors. C(Sector) was added as another standalone variable, then as an interaction term **Beta \* C(Sector)** on the assumption that the effect of beta on average daily return may differ by sector. This increased the R2 statistic to 27.7% and the lower AIC value compared to the Simple Linear Regression model indicated a better model fit.

**Random Forest Feature Importance Analysis**

Feature Importance Analysis was employed to assess the factors influencing return. The R2 statistic indicated the model accounted for 43.15% of the variation in average daily returns. The key predictors were found to be Volatility (0.264), Dividend Yield (0.199) and Beta (0.164). Notably, financial indicators, including risk, have stronger predictive capability than Sector. Based on this dataset, Sector does not emerge as the most significant factor in predicting returns leading to conclude that the second hypothesis is unlikely.

A graph with blue dots

Description automatically generated

Figure : Random Forest Predictions for Average Daily Returns plotted against actuals, showing a mostly trend with some outliers

**DISCUSSION AND CONCLUSIONS**

The analysis conducted support the first hypothesis for the most part, which posited that stocks with a beta higher than 1.0 would deliver average daily returns at least 20% higher than those with a beta lower than 1.0, in line with the Capital Asset Pricing Model (CAPM). While the findings from the one-way ANOVA and t-tests indicate a significant difference in the performance between stocks with high and low betas (of at least 43.06% when compared to the 20% threshold originally hypothesized). However, prior analysis of the linear relationship between beta and returns including the moderate R-squared value in the linear regression (0.168) suggests that beta alone is not a sufficient predictor of average daily returns. There is an association of high beta and high returns, but not modelled by the CAPM model.

Broader literature (Fama, E.F. and French, K.R., 2004) claim that the CAPM model as developed by Sharpe was never an empirical success. They analysed share returns on the American and New York Stock Exchanges and found that the linear relationship was not present when comparing beta and return over longer periods. Some of the assumptions of the model are considered unrealistic, such as assuming the risk-free rate will remain the same over a holding period. (Kenton, W., 2024). However, this model paved the way for further research into other variables such as size, price ratios and momentum that add to the explanation of average returns explained by beta.

As seen in the analysis, it appears there is no overarching factor -whether risk or sector- can entirely model returns. Stock returns can be thought of as broader trends and multi-factor models, as the risk-return relationship is influenced by a variety of factors. The relationship between the two is not necessarily unidirectionally causal, as higher returns can cause more volatility, if it triggers more investor activity. There are other third-party factors such as the macroeconomic environment and investor sentiment that can affect both price and volatility.

*Limitations of the analysis*

* The model makes use of null hypothesis testing, typically building on the assumption that the data is distributed normally. The data here may not be representative of the true population, so any trend observed may not be considered empirical evidence for or against a hypothesis.
* Temporal trends were not accounted for, as the model did not use time-series analysis
* The analysis assumes normal distribution, which may not hold true for all financial data.

In conclusion, while the CAPM provides a foundational framework for understanding the risk-return trade-off, stock returns are influenced by a complex interplay of factors beyond systematic risk. Future research could explore multi-factor models and temporal dynamics to better capture the nuances of this relationship.

**References**

Chen, J. (2024) *Risk-return tradeoff: How the Investment Principle Works*, *Investopedia*. Available at: https://www.investopedia.com/terms/r/riskreturntradeoff.asp (Accessed: 09 December 2024).

Cutter, P. (2024) *Navigating Tech Volatility*, *The merits of using long and short strategies to navigate volatility in the technology sector.* Available at: https://www.schroders.com/en-us/us/intermediary/insights/navigating-tech-volatility/?utm\_source=chatgpt.com (Accessed: 04 January 2025).

Fama, E.F. and French, K.R. (2004) ‘The Capital Asset Pricing Model: Theory and evidence’, *Journal of Economic Perspectives*, 18(3), pp. 25–46. doi:10.1257/0895330042162430.

Freeman, D., Pisani, R. and Purves, R. (2014) ‘Chapter 8: Correlation’, in *Statistics*. 4th edn. New York City, New York, United States: W. W, Norton & Company, Inc., pp. 119–141.

Kenton, W. (2024) *Capital Asset Pricing Model (CAPM): Definition, formula, and assumptions*, *Investopedia*. Available at: https://www.investopedia.com/terms/c/capm.asp (Accessed: 03 January 2025).

Sharpe, W.F. (1964) ‘Capital asset prices: A theory of market equilibrium under conditions of risk\*’, *The Journal of Finance*, 19(3), pp. 425–442. doi:10.1111/j.1540-6261.1964.tb02865.x.

**Use of Artificial Intelligence**

AI was used to brainstorm succinct and suitable titles for this project

A screenshot of a black and white screen

Description automatically generated

AI was used for repetitive and information purposes, such as adding basic descriptions for some of the financial features not directly related to the investigation.

**A screenshot of a computer

Description automatically generated**

AI was used to help to format diagrams side by side, two in one for the histograms I had programmed independently.

A screenshot of a computer

Description automatically generated