

David Hernández

Website

Github

Linkedin

Understanding the Dynamics of Volatility, Returns, and Risk in Stock Investment: Correlation Between Return & Volatility



Gmail davidhernandex2002@gmail.com

INDEX

Introduction.....	2
Methodology and Technologies.....	3
Data Analysis and Results.....	4
Implications for Investment Strategy.....	5
Conclusions	6

1. Introduction

In the ever-evolving world of finance, the stock market remains a captivating arena that entices investors with its potential for substantial returns and wealth creation. However, alongside these enticing opportunities, the stock market introduces an element of uncertainty and risk, embodied by the concepts of volatility and risk. While often used interchangeably, it is essential to clarify the distinction between volatility and risk, as they play distinct roles in shaping investment outcomes.

1.1 Background

The concept of volatility in the financial markets has long been a subject of intense research and scrutiny. Financial scholars and practitioners alike have debated its significance in determining risk and the implications it holds for investment strategies. Volatility refers to the measure of price fluctuation or dispersion of returns over a specified period. It is a critical metric for assessing market dynamics and capturing the market's degree of uncertainty. Higher volatility implies greater price swings and potential for significant gains or losses.

On the other hand, risk encompasses a broader notion that encapsulates various factors affecting investment outcomes. While volatility represents one component of risk, it is not synonymous with risk. Risk encompasses multiple dimensions, including market risk, business risk, liquidity risk, and credit risk, among others. Unlike volatility, risk takes into account the probability of adverse events.

1.2 Objectives

The primary objective of this study is to establish a comprehensive understanding of the intricate relationship between volatility, returns, and risk in the context of stock investment. Through meticulous data analysis and robust statistical techniques, I aim to explore the following key aspects:

- Examine the historical volatility trends of 5000 American companies over a significant period.

- Investigate whether there exists a correlation between volatility and returns, and the extent of this relationship.
- Assess how risk, encompassing various dimensions, influences investment outcomes in different market scenarios.
- Provide valuable insights and recommendations to investors on leveraging volatility and managing risk to make informed investment decisions.

Methodology and Technologies

In this section, we outline the approach we undertook to analyze the relationship between volatility, returns, and risk in stock investment, as well as the technologies employed to facilitate the data collection, processing, and analysis.

Data Collection and Analysis Technologies

The methodologies applied in this study involved a combination of data collection and statistical analysis techniques. To gather the necessary data for the 5000 American companies, I employed a suite of technologies to ensure accuracy and comprehensiveness.

- **Python and pandas:** I utilized the **Python** programming language and the "**pandas**" library to streamline the collection and computation of data. This allowed us to efficiently manage, clean, and manipulate the datasets, ensuring the accuracy and reliability of our analysis.

- **MySQL:** For the creation and management of my **database**, I employed MySQL. This relational database management system provided a robust and structured environment to organize and store the data extracted from the API. The integration of MySQL ensured data integrity and facilitated complex queries for further investigation.

- **Power BI:** For graphical representation and **visualization** of the **analyzed data**, I utilized Power BI. This powerful business intelligence tool enabled us to create insightful and dynamic visualizations, aiding in the interpretation and communication of our findings.

Data Source

I sourced the data required for this study from the "EodHistoricalData" API, a reputable and comprehensive data provider. The API offers historical market data, including price, volume, and other relevant financial metrics for a wide range of companies. This dataset served as the foundation for our investigation into the volatility, returns, and risk associated with the selected companies.

Data Analysis and Results

In this section, is presented a comprehensive analysis of the collected data, shedding light on the intricate relationship between volatility and returns in the context of stock investment. The visual representation of the data offers valuable insights into patterns and trends that can significantly influence investment decisions.

The Data

The obtained data comprises a matrix with 6 columns and 745,000 rows, totaling over 4 million data points (Code information for obtaining this data is available on my GitHub.) .

ticker	date	month_std_deviation	month_profitability	correlation	p_value
HAL	2016-05	1,25100004673004	5,67000007629395	-0,27493399381	0,000853119
HALL	2016-05	0,361999988555908	-6,54300022125244	-0,29335099458	0,000359205
HALO	2016-05	0.602999985218048	-15.2410001754761	0.049013700336	0.559630990

Correlation

In the pursuit of comprehending the intricate interplay between volatility and returns, I employed the Pearson correlation coefficient—a statistical tool renowned for assessing the strength and nature of linear relationships between continuous variables. With values ranging from -1 to 1, the coefficient conveys the degree of correlation:

- A value close to 1 signifies a robust positive correlation, where both variables tend to rise together.
- A value close to -1 signifies a strong negative correlation, indicating that as one variable increases, the other tends to decrease.
- A value around 0 suggests a weak or negligible linear relationship between the variables.

However, while the coefficient furnishes insights into the relationship's nature, it alone does not establish the relationship's significance. This is where the concept of p-value becomes fundamental.

In the realm of correlation analysis, the p-value plays a vital role in gauging the robustness of the observed relationship against the backdrop of hypothesis testing. The null hypothesis (H_0) posits that no substantial correlation exists between the variables, while the alternative hypothesis (H_1) asserts the presence of a meaningful correlation.

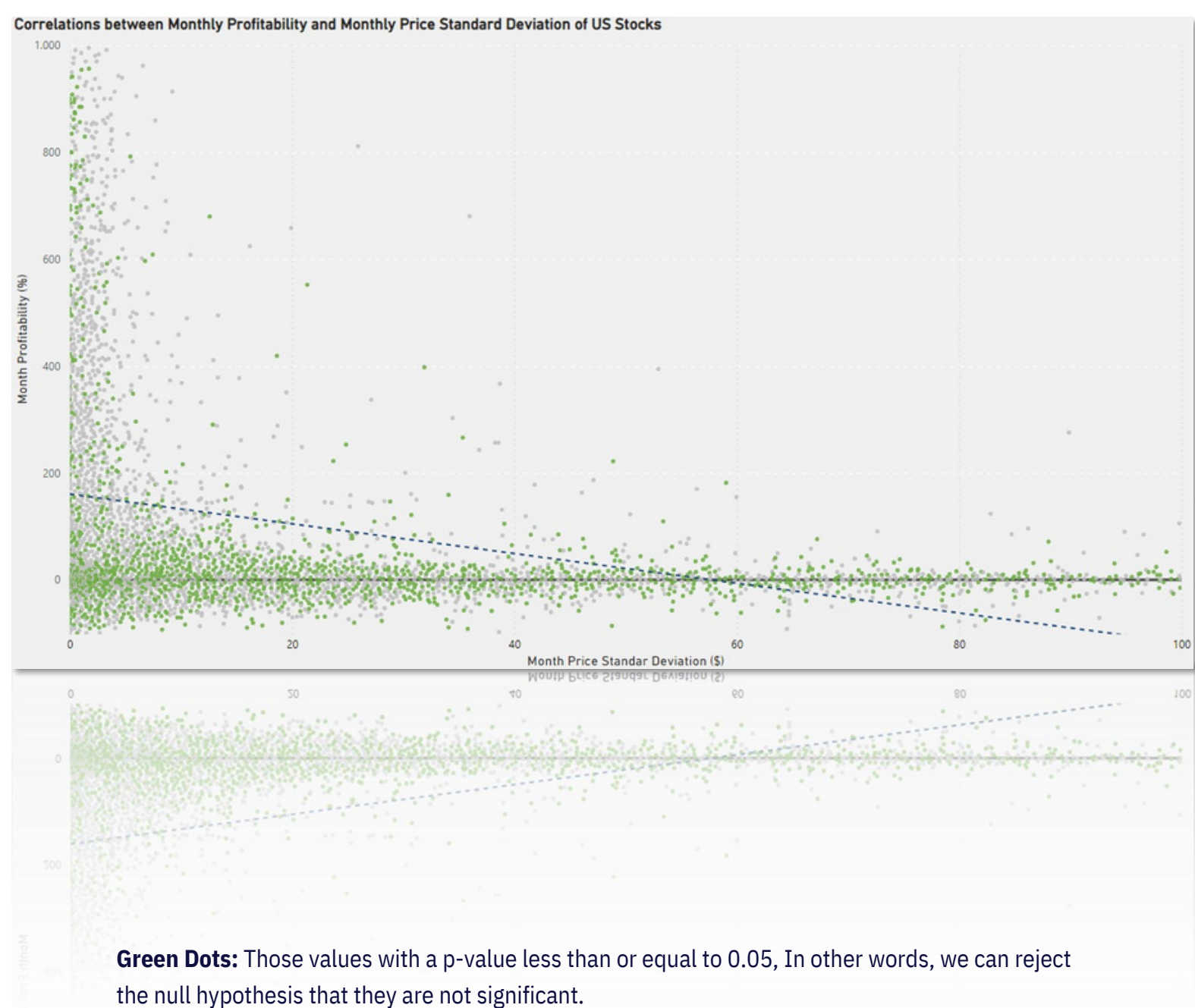
The p-value furnishes a measure of the likelihood of encountering the observed correlation coefficient (or one more extreme) under the assumption of the null hypothesis being valid. A small p-value, often considered less than 0.05, indicates that the observed correlation is statistically significant. In practical terms, this signifies the rejection of the null hypothesis in favor of the alternative hypothesis.

For my investigation, I computed the Pearson correlation coefficient to quantify the strength and direction of the relationship between volatility and returns. The accompanying p-value serves as a critical indicator of the statistical significance of the derived correlation. If the p-value demonstrates sufficient statistical significance (≤ 0.05), we can confidently reject the null hypothesis and conclude that a noteworthy correlation exists between the variables under study.

Volatility vs. Returns Analysis

One of the central aspects of my analysis involved plotting a scatter plot of monthly returns against the monthly standard deviation over the past decade. This visualization effectively captures the interplay between volatility and returns, allowing us to draw compelling conclusions.

***Figure 1: Correlations between Monthly Profitability and Monthly Price Standard Deviation of US Stocks (Data extracted from EODHD, data processing done by me.) ***



Trend Line: The trendline is drawn using both green and gray points.

Truncation: This graph is truncated to visualize the correlation between price and volatility, with a condition of month price standard deviation less than 100 and month profitability less than 1000 .

As observed in Figure 1, the scatter plot reveals a clear trend that underscores the relationship between volatility and returns. Companies with lower monthly standard deviations, indicating lower volatility, tend to exhibit higher monthly returns. Conversely, companies with higher volatility exhibit more varied returns, including both potential gains and losses.

Once the results from the table are observed, let's see what the numbers say:

The **average correlation**, considering those values where the **p_value is \leq zero**, is: **0.0227**.

The **average correlation**, considering all values, is **0.0086**.

A significant correlation cannot be established. Up until now, we were approaching the study by considering both variables as if one explained the other, or at least depended on it to some extent. However, the correlations indicate that they are largely unrelated.

Implications for Investment Strategy

In the realm of stock market investment, we can draw **two main conclusions**:

1. Higher volatility does not necessarily imply a lower or higher price.

2. Risk comprehension.

- Obviously, there are many definitions of risk. However, there is one that we may find interesting to analyze in this study. In this definition, risk is defined as volatility, and it tells us that there is a proportional relationship between volatility and expected returns.

The second assertion has already been refuted by the study's results. However, we need to make a series of logical assumptions to refute the first assertion "risk is defined as volatility". If we were to claim that risk is equal to volatility, based on the results of this study, we would also be asserting that the levels of risk when making an investment have no correlation with their potential returns. Obviously, making such a claim would be entirely incorrect.

To conclude, it is highly likely that volatility is one of the numerous variables that impact risk. However, it is entirely incorrect to assert that risk is a variable that can be defined as volatility. There are more logical variables for measuring investment risk than merely observing how the stock price fluctuates (default risk, debt ratio, business model, etc...).

With the results of this study, we can arrive at the conclusion that volatility and returns do not exhibit a significant correlation. Nevertheless, it's important to consider the limitations of this study.

The sample consisted of the daily prices of 5000 U.S. companies over the past 10 years. It's probable that by expanding the sample size and investigating other countries, the results could vary.