QQQ vs SMH: Has Software or Hardware Benefited More From AI?

Aidan Wardak

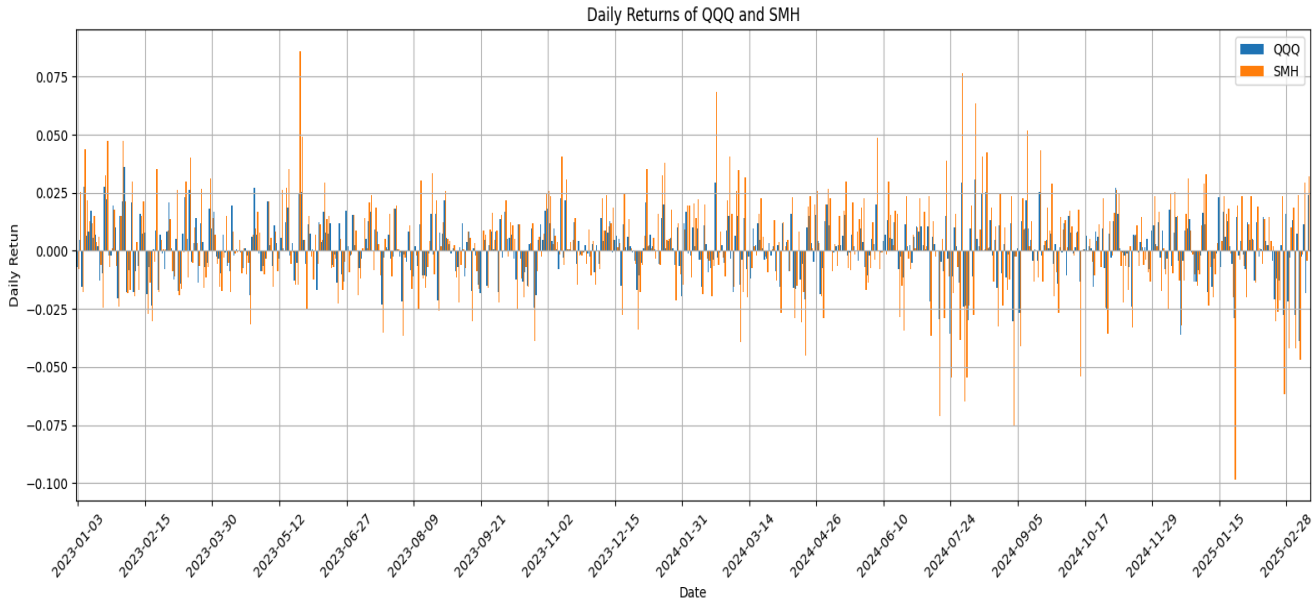
Introduction

The two ETFs being analyzed are Invesco QQQ Trust, ticker ‘QQQ’, and VanEck Semiconductor ETF, ticker ‘SMH’. I am interested in finding which ETF has performed better daily after the release of ChatGPT and subsequent AI technologies, so that we can make suggestions as to which stocks have benefitted more, computer software oriented companies (QQQ) or computer hardware oriented companies (SMH). Originally, I planned to perform a two-sample t test to compare the mean daily returns of each ETF, but when checking the conditions for statistical inference, I found that although the ETFs necessarily focus on different subsectors of technology, it would be hard to say their daily returns are independent of each other. This was further suggested when the Pearson correlation between the daily returns for QQQ and SMH was found to be 0.8666. This led me to believe that a paired t test for the difference of mean daily returns between QQQ and SMH would be better suited to compare daily performance of the ETFs as the observations of each sample are not independent of each other. In addition, the selection of the paired t test would reduce noise in the data from market shocks which could inflate variability between the daily returns. The two-sample t test would not be able to reduce this noise, leading to an increased possibility of error.

Data Collection and Preparation

The price data necessary for the paired t test was downloaded from Yahoo Finance using the yfinance library. The historical price data was collected from December 30, 2022 to March 25, 2025. These start and end dates were chosen to capture when financial markets began to process the release of new AI technologies, and specifically, ChatGPT which was released on November 30, 2022. I lagged the start date by a month to capture when sentiment towards AI picked up and set the end date to March 25, 2025 to reduce the noise in the data caused by recent tariffs. Closing prices were then downloaded for each ETF and the daily percent change of the ETFs (only on trading days) was pulled from the data and this represents daily returns. The final

dataset included 551 total differences (paired observations) of daily returns between QQQ and SMH. The individual data sets before being subtracted from each other are:



Statistical Test

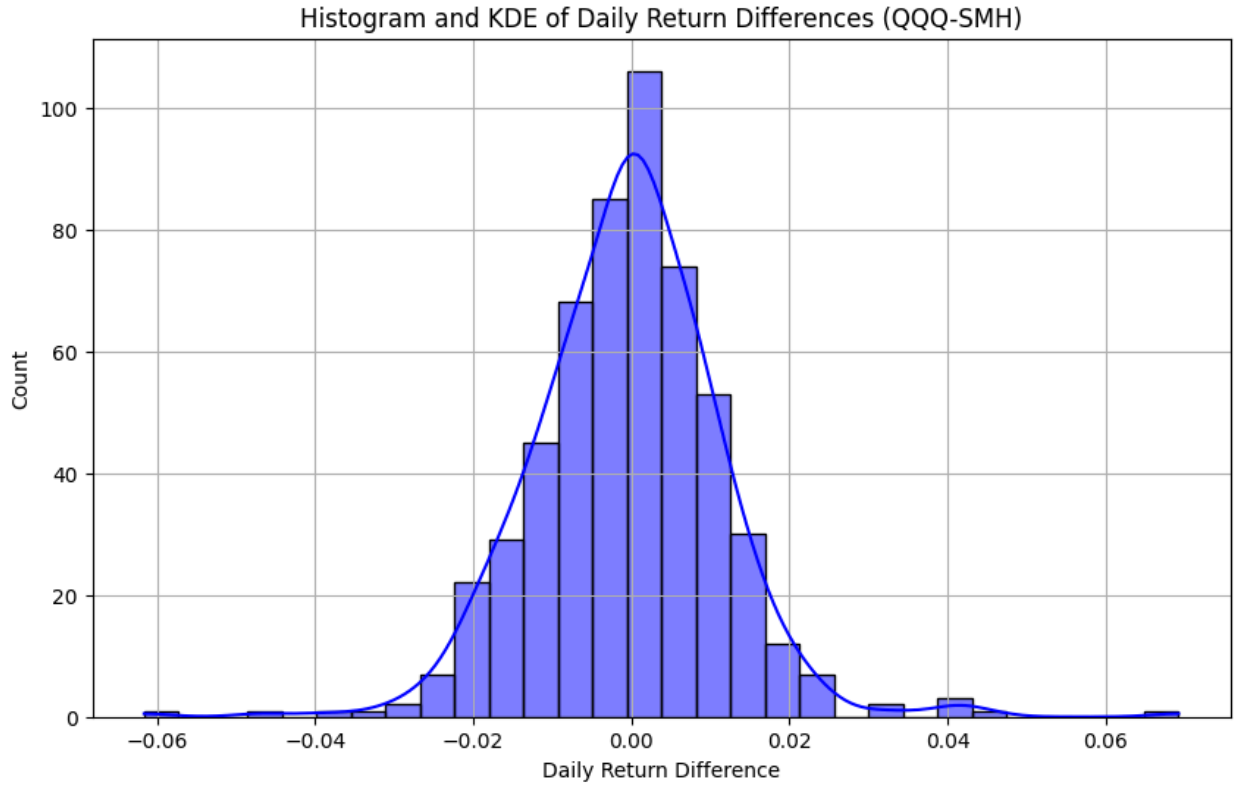
Parameter

We want to determine if QQQ has different mean daily returns than SMH in the post-ChatGPT era at the 5% significance level. Let mean(d) be the mean difference of the daily returns between QQQ and SMH (QQQ-SMH). The null hypothesis is H0: mean(d) = 0 and the alternative hypothesis is Ha: mean(d) /= 0.

Conditions

a) Simple Random Sampling: the sample includes all non-NA data values from QQQ and SMH. While random sampling was not used in the paired t test, it is reasonable to consider the sample of all differences in daily returns between QQQ and SMH in the period to be pseudo-random as they result from the stochastic nature of financial markets themselves.
b) Independence: A lag-1 autocorrelation analysis was performed for QQQ and SMH and I found it to be -0.0045 and -0.0423, respectively. As these values are very close to zero, it is reasonable to assume independence for the purpose of the paired t test.

c) Normality: Given the sample size is n = 551, the central limit theorem can be assumed as n > 30, leading us to conclude the data is approximately normal. A histogram is shown to confirm this:



Calculations

I imported stats from scipy and performed a paired t test. This produced a t statistic of -1.0411 and a p-value of 0.2983.

Conclusion

Because the p-value of 0.2983 is greater than the level of significance, 0.05, we fail to reject the null hypothesis that mean(d) = 0. There is not convincing statistical evidence that the mean difference in daily returns between QQQ and SMH is different from zero in the post-ChatGPT period.

Discussion

The statistical test led to the conclusion that there is not convincing statistical evidence that the mean difference in daily returns between QQQ and SMH is different from zero in the post-ChatGPT period. This result suggests that QQQ did not perform better than SMH on a daily basis and hence that broad technology did not perform better than semiconductor companies and vice versa. From a trading perspective, there is not statistically significant evidence that traders would have a higher return with QQQ than SMH and vice versa. Both of the ETF's maintain a strong, positive, linear association with a pearson correlation of 0.8666, further suggesting that trading QQQ or SMH would not result in significantly different daily returns. The expansive market sentiment towards AI seems to have impacted QQQ and SMH very similarly, showing how broad technology and hardware companies have shared gains from the trend. Yet, there are several limitations of this statistical test that should be noted. The test considered only daily returns, potentially missing weekly or intra-day changes in price that affect the tradability of the ETFs. In addition, it did not account for risk and volatility, further limiting the trading implications of this test (see appendix A for a better suited paired t test). Furthermore, the performance of semiconductor companies and broad technology company stocks may diverge in the future due to unforeseen changes in markets. Future research could consider Sharpe ratios of both ETFs to evaluate risk-adjusted returns or volatility tests to compare volatility between ETFs.

Conclusion

This analysis has shown QQQ is not different to SMH in terms of daily returns in the post-ChatGPT period and suggests that broader technology companies and semiconductor companies have both benefited from the trend towards AI. Future research could be done to provide more statistical evidence of this claim or be used to further evaluate QQQ versus SMH as trading options.

Appendix A.1

In this section, a paired t test with beta-adjusted returns will be employed for increased insight into whether trading QQQ or SMH is a better option for reducing risk and increasing profit.

Parameter

We want to determine whether QQQ or SMH has different mean beta-adjusted daily returns in the post-ChatGPT period (December 20, 2022 - March 25, 2025) at the 5% significance level. Let mean(d) be the mean difference of the beta-adjusted daily returns between QQQ and SMH (QQQ-SMH) in the post-ChatGPT period. The null hypothesis is H0: mean(d) = 0 and the alternative hypothesis is Ha: mean(d) /= 0.

Conditions

All conditions are fulfilled for the reasons given previously.

Calculations

I imported stats from scipy and performed a paired t test. This produced a t statistic of -0.0153 and a p-value of 0.9878.

Conclusion

Because the p-value, 0.9878, is greater than the level of significance, 0.05, we fail to reject the null hypothesis that mean(d) = 0. There is not convincing statistical evidence that the mean difference between beta-adjusted returns for QQQ and SMH in the post-ChatGPT period is different from zero.

Appendix A.2

In this section, the findings of this statistical test will be discussed briefly. The statistical test led to the conclusion that there is not convincing statistical evidence that the mean difference between beta-adjusted returns for QQQ and SMH in the post-ChatGPT period is different from zero. This result indicates that after controlling for beta, the daily performance of QQQ and SMH was not statistically different over the studied time period. This data combined with the results of the previous paired t test suggest that trading QQQ or SMH during the post-ChatGPT period does not confer benefits either way on a daily basis despite differences in volatility.

Return Type	t-statistic	p-value	Conclusion
Raw daily returns	-1.0411	0.2983	Fail to reject null hypothesis

Beta-adjusted daily returns	-0.0153	0.9878	Fail to reject null hypothesis

Appendix B

The full code used in this paper can be found at:

- https://github.com/wardakaidan/qqq_vs_smh_t-test