**Exchange Traded Fund 1 Year Return and Asset Classes**

**Introduction**

An Exchange Traded Fund (ETF) is a bundle of securities that are traded together like a stock and typically represent an underlying asset or market sector. For example, the iShares Core S&P 500 ETF (IVV) tracks the S&P 500 stock index, an index that monitors the performance of 500 of the largest market capitalization companies in the United States. The IVV can be bought and sold daily, like any other liquid stock. There is an underlying assumption that by investing in a bundle of securities instead of individual securities in a market, the investment is more risk adverse because of lower volatility and more resistance to systematic factors that can negatively affect individual assets. There are many different types of ETFs on the market and an abundance of ETF providers. The provider of ETFs are market makers, entities that buy and sell securities when there is not another party to buy and sell that security. Thus, the barrier to entry to become a provider of ETFs is extremely high. The largest players in the market are BlackRock iShares, Vanguard ETFs, and State Street iShares. These companies provide many different types of ETFS, representing unique underlying assets and different investment strategies. Each underlying asset moves with patterns unique to the market or index that the ETF tracks. A few of the different types of asset classes are Commodity Stocks (Gold, Silver, and Crude Oil), Equity Stocks (S&P 500, Technology Sector, Agriculture) and Fixed Income Bonds (U.S. Treasury Notes, Small and Large Market Capitalization Corporate Bonds, and Foreign Debt). Moreover, there are basically two avenues for the market makers to monitor these ETFs. By taking a passive strategy, a market maker will buy and sell securities to track another index. For example, the iShares Core S&P 500 ETF attempts to buy and sell securities in the list of 500 of the largest companies in the U.S. proportionally to match how the S&P index is performing. Another strategy is active investing, where a market maker will buy and sell securities to beat an index, and thus seeking greater returns. Therefore, the research question of interest to any intelligent investor is whether different Asset Classes and Investment Strategies correspond to greater returns. By investigating this question, any potential investor will have a more transparent view of the overall market for ETFs and be able to make more informed decisions when faced with choosing an Asset Class or an Investment Strategy.

**Data Summary**

This study is entirely observational because the data was retrieved by iShares ETF report, and thus, was not experimentally yielded. BlackRock iShares provides financial advisors and investors with data on the performance of an ETF and corresponding metrics that characterize the equity. Net Asset Value (NAV) is the summation of the value of each of the underlying assets that are bundled into the ETF. For example, if an ETF was only made up of Apple (APPL) trading at $200, Amazon (AMZN) trading at $1900, and Microsoft (MSFT) trading at $150, then the NAV for this ETF would be $2250. This metric is interesting for investors because there is an assumption that large market capitalization companies have slower growth over time, so companies with higher NAV usually correspond to companies with higher market capitalization. Another metric used to characterize ETFs is the Expense Ratio. The Expense Ratio is the proportion of the amount charged by market makers to manage the fund versus the total value of the fund. ETFs with higher Expense Ratios are usually traded less frequently than ETFs with lower Expense Ratios. Moreover, unlike the NAV, higher Expense Ratios correspond to ETFs with lower market capitalization companies. Thus, for this study, 4 different variables were analyzed. The Asset Class is a Random Effect Categorical Variable because the three different asset classes being analyzed represent a small sample of all of the possible Asset Classes for ETFs. The Investment Style is a Fixed Effect Categorical Variable because there are only two types of Investment Styles, active and passive. The two covariates being analyzed are the NAV and Expense Ratio for individual ETFs. The treatment design for this observational study is a crossed factorial design because every Asset Class has both types of Investment Styles.

**Hypotheses**

For testing the main effect of the Asset Class on the 1 Year return for an ETF, the null hypothesis is that each Asset Class has the same 1 Year returns.

The alternative hypothesis is that each Asset Class has a different mean 1 Year returns.

For testing the main effect of the Investment Style on the 1 Year return for an ETF, the null hypothesis is that both Investment Styles have the same 1 Year returns.

The alternative hypothesis is that different Investment Styles have different mean 1 Year returns.

Lastly, there may be an interaction term between the Asset Class and the Investment Style. Thus, the hypotheses for testing the interaction term is:

**Exploratory Data Analysis**

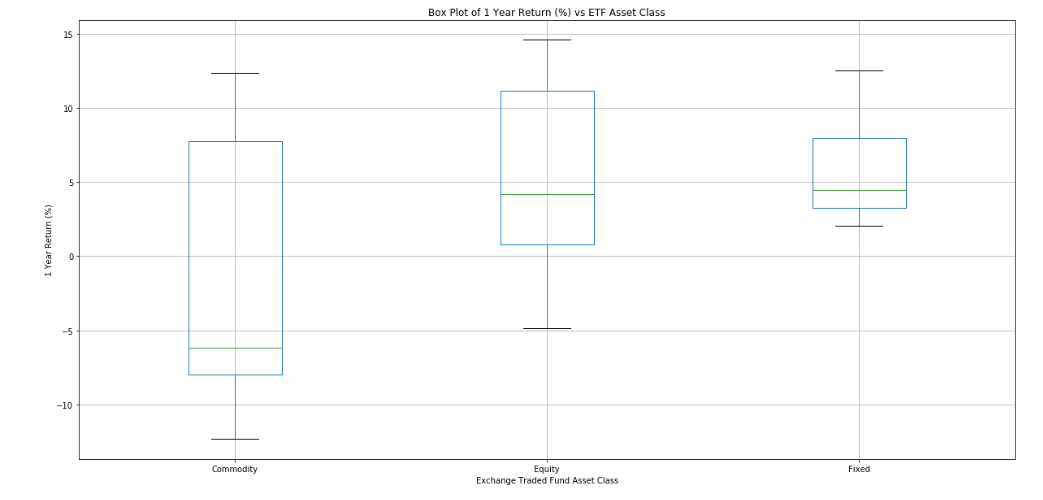
****

Figure 1: Box Plot of 1 Year Return versus the ETF Asset Class

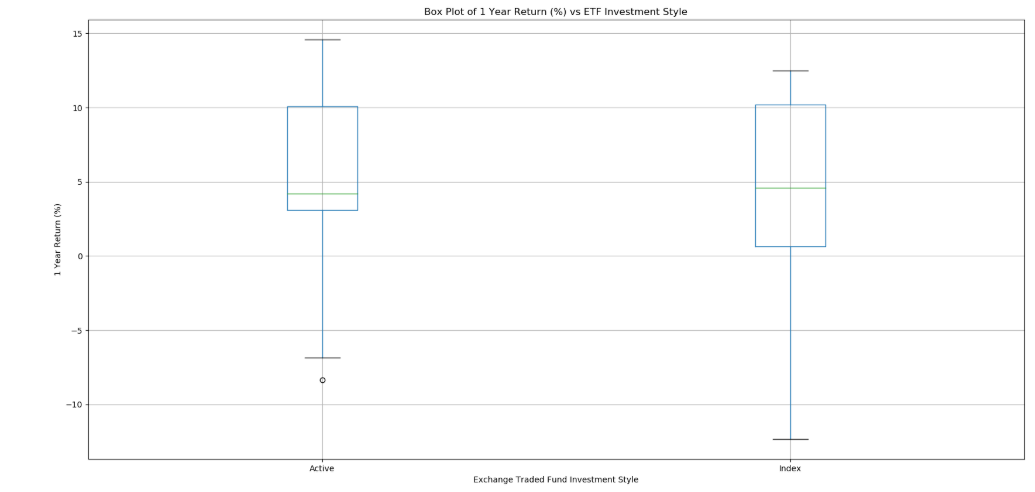


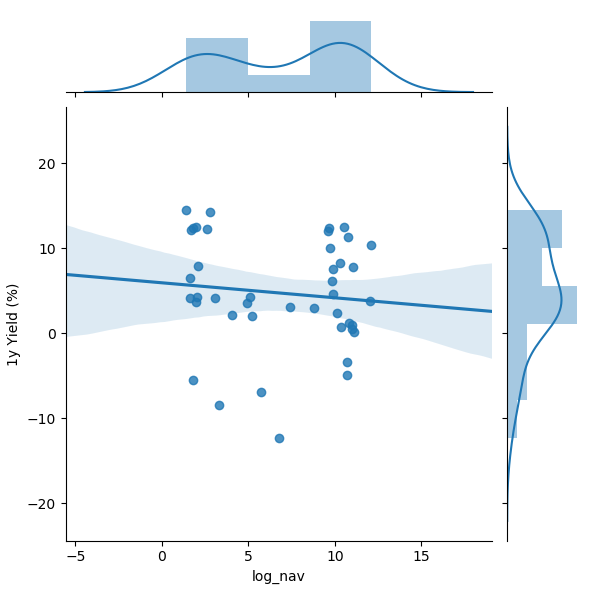
Figure 2: Box Plot of 1 Year Return vs ETF Investment Style

Figure 3: Joint Distribution Plot of 1 Year ETF return regressed on the Log Transformed NAV

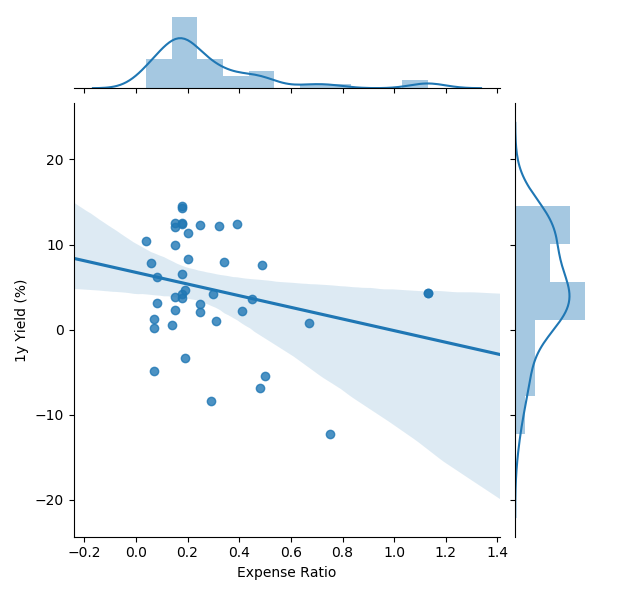


Figure 4: Joint Distribution Plot of 1 year ETF return regressed on the Expense Ratio

Initial analysis of the main factors for the ETF return analysis raises concern over the linear relationship between the covariates and the response variable. From Figure 3, it is clear that there only exists a linear relationship if the NAV is log transformed. It is common to log transform monetary data because it is always positive definite and usually follows a right skewed distribution. Looking at the distribution plot, a log transform makes the data appear more normal, yet remains bimodal. However, after transforming NAV, there seems to be an inverse linear relationship. Similarly, the Expense Ratios also have a negative linear relationship with the Yield. Moreover, the box plots from Figure 1 shows a large spread in data for the variables split along the Asset Class factor. Similarly, the values in Investment Style seem equally spread. This raises a concern over the ability of the factors to predict the mean of the Yields. However, after introducing a covariate, there is reason to believe that there is more of a linear relationship in the data than appears from the EDA.

**Analysis of Variance**