



Figure 1: The plot above shows the daily closing price time series of the 12 selected ETFs listed on the New York Exchange (NYSE). They all span more than ten years except DWT which spans over about two years. All the series end on October 25, 2019

The figures below show the trend in the predicted daily closing prices of the selected ETFs. I used an autoregressive model (i.e., AR(1)) to make the predictions where the model is been estimated by ordinary least square (OLS). I used 90% of each of the time series as the training sets and the remaining 10% for testing. The figures below plots actual time series (in blue) vs predicted series (in red). An important observation in almost all the plots was some sort of a spike or jump in the predicted prices immediately after the observe prices. This maybe due to an important feature in the datasets that my approach could not figured out. In the main work of the proposed project I intend to implement Long-Short Memory network model which may capture all the features of the data. Another observation in the plots below is that all the predicted prices tend to follow a downward trend except DWT and XOP. This is as result of poorly fit of the model to the data since the root mean square errors for all the series are far away from zero.

The main focus of the proposed capstone project is to predicted the time horizons of ETF misprings. Although I obtained the datasets for all the indices that the ETFs track, due to time constraint I could not proceed to the case of predicting the time horizons of the ETF mispricing. However, the the prediction of the time horizons will depend on the predicted closing prices.











