Risk Management and Financial Engineering

Empirical Finance: Coursework 1

Imperial Global Asset Management: Investment Report

Talha Jamal - 01357497

February 8, 2024

Contents

1. Introduction

1.1 Abstract

This report presents the research carried out by Imperial Global Asset Management to build a globally diversified portfolio, using a systematic trading strategy backtested on over 100 years of data. It presents the assumptions, statistical tests, and economic intuition behind the systematic trading strategy used to build an optimized portfolio giving investors the chance to earn risk adjusted returns. The optimized portfolio consists of the US and UK stock indices combined with US short term bonds. The portfolio has a target volatility of 10% per annum taking into account the risk appetite of potential investors. This target volatility can be optimized further as per the investor's risk appetite.

1.2 Data

The report's research used historical time series data for US and UK stock indices, short term yields and long term yields. Data from 1792 to 2023 with monthly frequency was used. The values for short term and long Term yields were converted from annualized values into monthly values.

Figure 1.1 represents a graphical illustration of the US and UK market data used from 1792 to 2023.

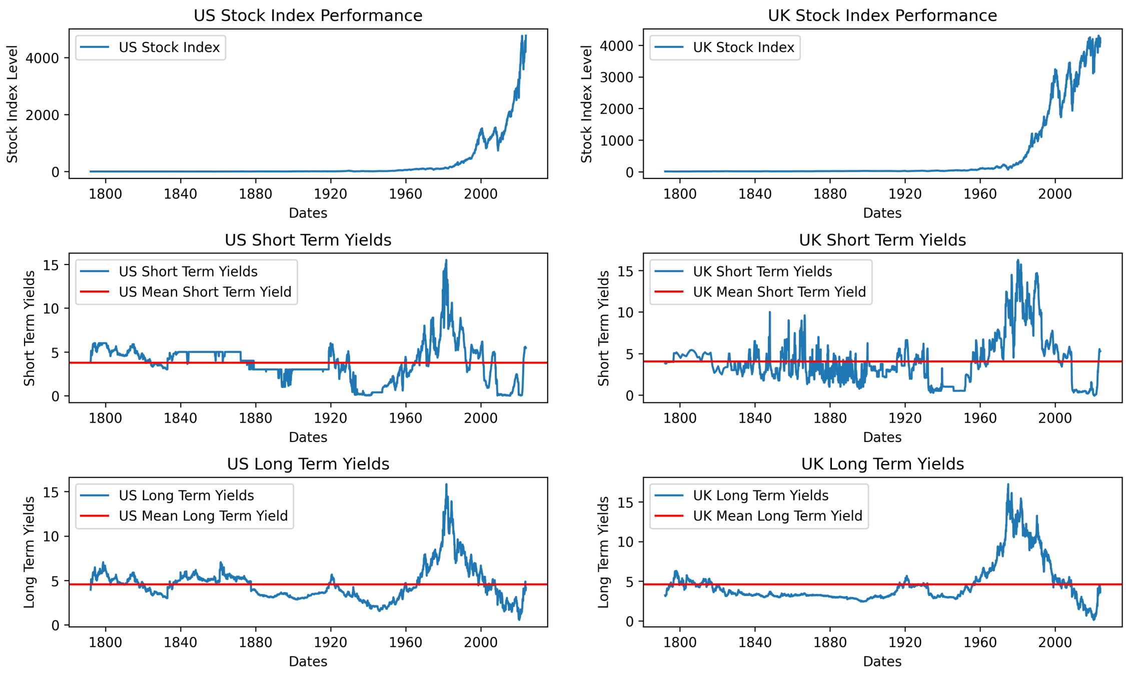

Figure 1.1: Market Data from 1792 to 2023

Both stock market indices follow a massive rise in their trajectory after 1970 & 1980. The UK stock market appears more volatile than the US stock market. Short term yields in both countries appear mean-reverting along a long run average, with short term yields appearing more volatile than long term yields. Long term yields on average appear higher than Short Term Yields owing to the liquidity premium and inherent risk in long term rates.

2. Time Series Analysis

Calculating returns for the US stock index and the change in yield for short and long term yields made each time series . Summary statistics on the three asset class available in our US universe were generated in order to analyse the statistical metrics of the asset classes themselves and their behaviour and relationship with respect to each other. ACF and PACF analysis was conducted to visually identify the relationship of each asset class with it's previously observed values. Finally, each stationary time series was fitted onto it's appropriate ARIMA model to identify the parameters of each time series model. Figure 2.1 below visualises all three time series data for the US and UK markets.

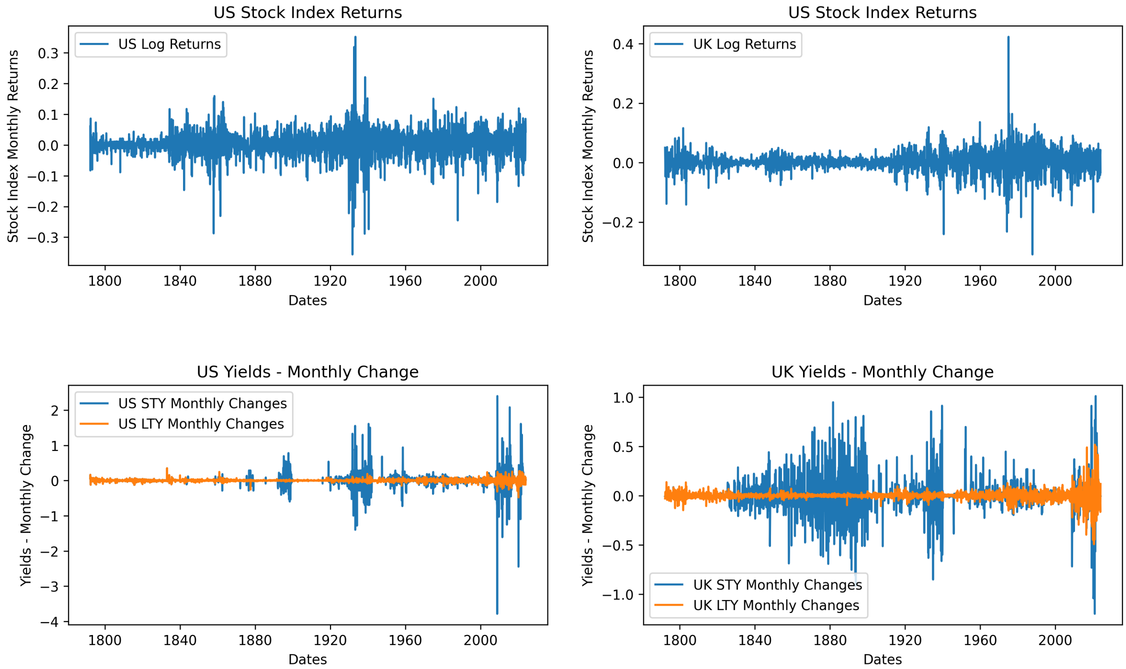

Figure 2.1: Stationary Time Series Data for all Asset Classes

2.1 Summary Statistics

Table ?? contains summary statistics for the three asset classes within the US Market universe. The stock index returns are centered around 0.27% per month with moderate volatility and slight potential for negative returns due to the negative skew. The high kurtosis indicates higher than normal probability of extreme values. The short term yield has a low average change but is extremely volatile. It's extremely high kurtosis points to extremely high chance of getting extreme changes in short term yields. The long term yield's change in values is more symmetrically distributed than the short term yield and and has moderate kurtosis compared to the stock index and short term yield. Figure 2.2 displays the distribution of the returns of each asset class.

Asset	Mean	Median	SD	Skewness	Kurtosis	Min	Max	ACF[1]	ACF[2]
Stock Index	0.0027	0.0025	0.043	-0.55	9.76	-0.356	0.352	0.105	-0.033
STY	0.00015	0.0	0.196	-1.69	85.72	-3.78.0	2.39	-0.04	-0.02
LTY	0.000025	0.0	0.043	-0.3056	16.97	-0.48	0.35	0.088	-0.022

Table 2.1: US Market Summary Statistics

The variance-covariance matrix in table 2.2 shows the behaviour of these three time series returns with each other. The high variance of the short term yield points to it's high volatile nature. The covariance reflects how the two asset classes move with respect to each other. The stock index is negatively correlated with the short term yield; as the short term yield rises, the stock market index falls. This is due to change in valuation of equities as the short term yield rises and investors demand a higher risk premium for investing in equities. The short term and long term yield are positively correlated - implying that they move together. However, the relationship is not that strong. The rest of the covariances reflect a positive but not very strong relationship.

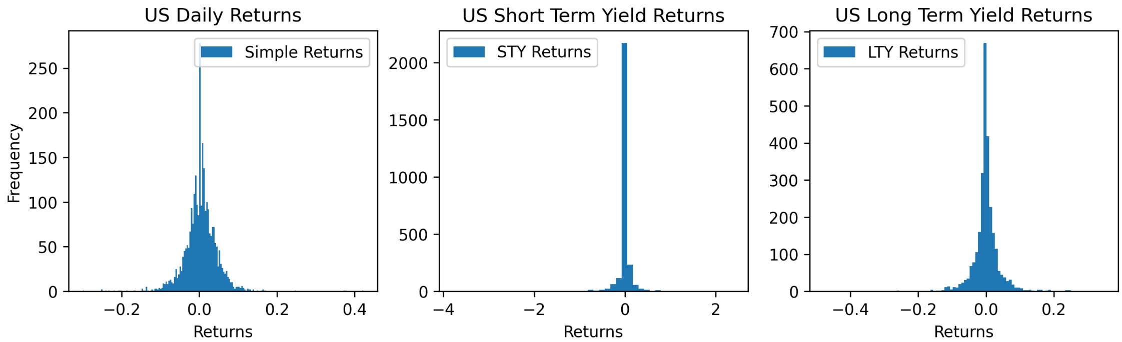

Figure 2.2: Distribution of Data for each Asset Class

	Index Returns	STY Change	LTY Change
Index Returns	0.001872	-0.000485	0.000073
STY Change	-0.000485	0.138073	0.000654
LTY Change	0.000073	0.000654	0.001866

Table 2.2: US Asset Class Variance Covariance Matrix

2.2 Time Series Specification and Parameters

Each stationary time series is visually inspected to identify it's autocorrelation and partial autocorrelation correlation with lagged data. The ACF and PACF at upto 35 lags is plotted in Figure 2.3. From visual evaluation, the US stock returns do not observe any significant correlation after the first lag. Since the return data is of monthly frequency, observing such low correlation with lagged data is expected. Data at higher frequency may have generated higher correlation at smaller lags compared to monthly data. The short term yield's returns observe a slightly significant negative correlation on the 3rd lag order - this may indicate that if the returns were above the mean 3 periods ago, they would now be below the mean, and vice versa. The long term yield's returns only observe a significant autocorrelation at the first lag order and nothing significant later on. In the plots of Figure 2.3, the correlation points higher than the blue band indicate significant correlations.

To assign the appropriate time series model to each data, we ensure that each time series is stationary via the the ADF test. Once stationarity was confirmed, we fitted each model onto ARIMA models with appropriate p and q lags and ran statistical tests to validate the parameters of our fitted models.

For the US stock index returns, only the first lag was significant in terms of it's correlation. When testing an ARMA(1,1) and a MA(1) Model, the MA(1) Model's parameter's coefficients were found to be statistically significant. The Ljung Box Test on the 1st lag had a test statistic of 0.002596 and a p value of 0.96. This indicated that within the residuals there is no autocorrelation at lag 1 and hence the model capture the data's autocorrelation structure well at this lag. The MA(1) model's parameters can be seen in Table 2.3.

For the short term yield returns, the plot showed negative correlation with the first 3 lags. By

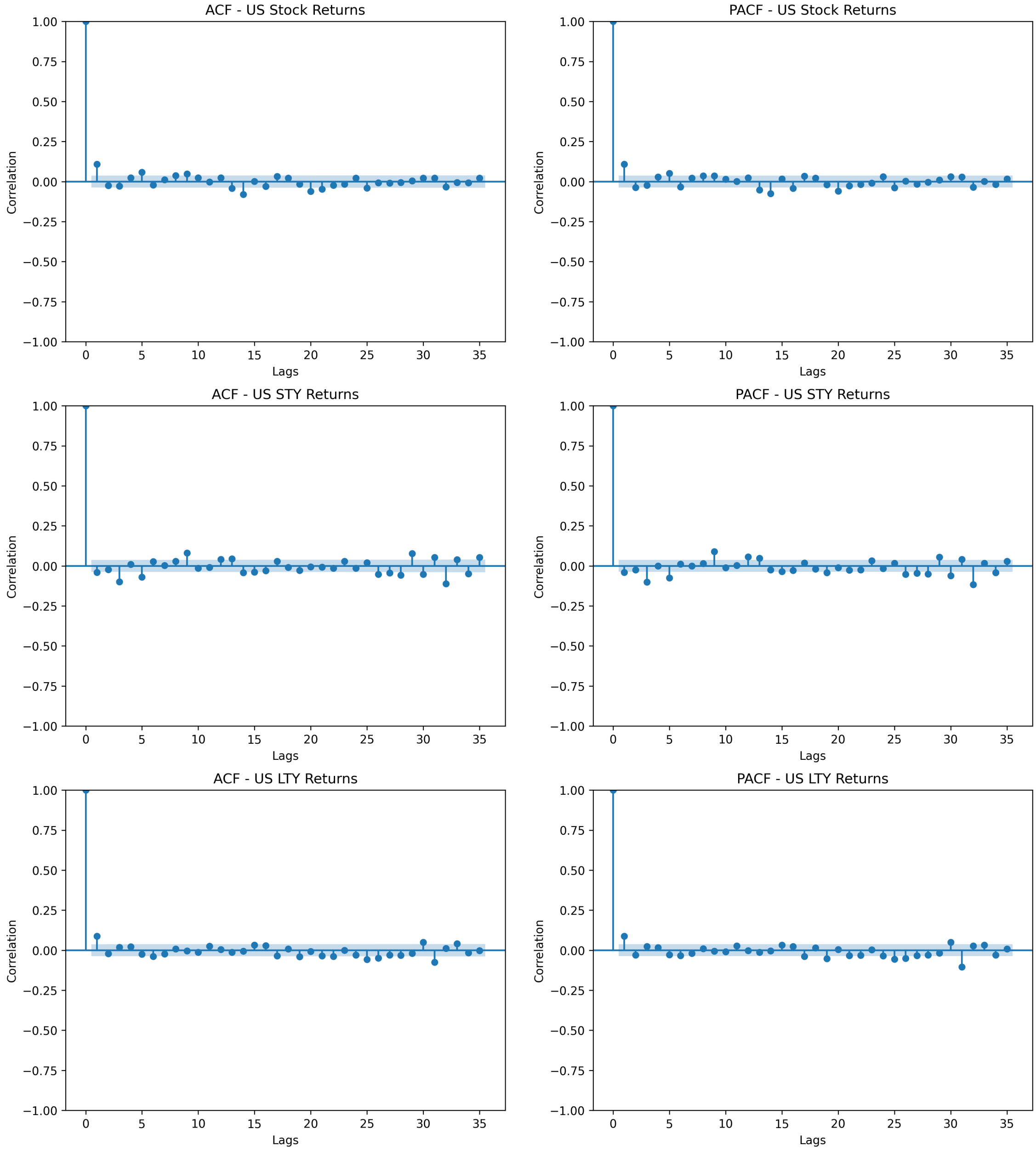

Figure 2.3: Autocorrelation and Partial Autocorrelation Analysis

	coefficient	std err	z	P>|z|	[0.025	0.975]
const	0.0037	0.001	4.087	0.000	0.002	0.005
ma.L1	0.1129	0.010	11.832	0.000	0.094	0.132
sigma2	0.0018	2.09e-05	88.133	0.000	0.002	0.002

Table 2.3: MA(1) Model for US Stock Index Returns

running a grid search algorithm to find parameters that were statistically significant, we found an AR(3) fit the data well compared to others. The Ljung Box Test till the 3rd lag had a test statistic of 0.032573 and a p value of 0.99. This indicated that within the residuals there was no autocorrelation at all lags up to the third one. Hence, the model appeared to capture the data's autocorrelation structure well. The AR(3) Model's parameters can be seen in Table 2.4.

	coefficient	std err	z	P > |z|	[0.025	0.975]
const	0.0037	0.001	4.197	0.000	0.002	0.005
ar.L1	0.1128	0.010	11.558	0.000	0.094	0.132
ar.L2	-0.0344	0.010	-3.297	0.001	-0.055	-0.014
ar.L3	-0.0223	0.010	-2.132	0.033	-0.043	-0.002
sigma2	0.0018	2.29e-05	80.157	0.000	0.002	0.002

Table 2.4: AR(3) Model for Short Term Yield Changes

For the Long Term Yield Returns, the plot showed positive correlation with only the first lag. We found an ARMA(1, 1) fit the data well compared to other models. The Ljung Box Test till the 1st lag had a test statistic of 0.02 and a p value of 0.90. This indicated that within the residuals there was no autocorrelation - they are white noise. Hence, the model appears to capture the data's autocorrelation structure well The ARMA(1, 1) Model's parameters can be seen in Table 2.5.

	coefficient	std err	z	P > |z|	[0.025	0.975]
const	2.287e-05	0.001	0.026	0.979	-0.002	0.002
ar.L1	-0.3087	0.079	-3.915	0.000	-0.463	-0.154
ma.L1	0.3980	0.078	5.123	0.000	0.246	0.550
sigma2	0.0018	1.7e-05	107.606	0.000	0.002	0.002

Table 2.5: ARMA(1, 1) Model for Long Term Yield Changes

3. Forecasting US Stock Returns

The one period ahead US stock returns were predicted based on a Term Spread Factor. We create this factor by subtracting the US long term yield by the US Short Term Yield and regress the stock returns on the one period lagged term spread values. A rolling and expanding window of 10 years was used in all regression models. The competing model (3.2) and the benchmark model(3.1) were statistically analysed to evaluate whether the competing model outperformed the market.

1. Benchmark model

$$y_t = \alpha + \varepsilon_t \tag{3.1}$$

2. Competing model

$$y_t = \alpha + \beta x_{t-1} + \varepsilon_t \tag{3.2}$$

3.1 Expanding Window Regression

The table 3.1 summarises the out of sample performance of the Expanding Window Regression with a starting window of 10 years.

MSE Benchmark Model	MSE Competing Model	Out of Sample R Squared	Delta RMSE
0.001937	0.001945	-0.004367	-9.5993e-05

Table 3.1: Expanding Window Regression Out of Sample Performance

3.1.1 Equal Predictive Ability of Benchmark & Competing Model

The model's out of sample efficacy was evaluated with the Mean Square Error between their out of sample prediction and actual observed value. The competing model has a slightly higher Mean Square Error than the benchmark model, indicating that the benchmark model performed very slightly better than the competing model. The Clarke and West Test statistic tested for the hypothesis that the MSE of the Benchmark is equal to the MSE of the Competing Model. With a CW Statistic value of -0.796, the null hypothesis is not rejected and the MSE of the Benchmark and Competing Model can be considered not statistically different to each other.

3.1.2 Out of Sample Performance

The Competing Model's out of sample R Squared was negative and extremely low - indicating this model does not perform well on out of sample data. The negative and extremely small delta RMSE between the two models suggests that both models perform very similar to each other on out of sample data - with the competing model performing very slightly worse than the benchmark model. These results suggest that choosing a model to forecast one period forward returns might depend on other factors such as simplicity of the model and calculations. These results suggest going with the benchmark model - the long run mean of the sample - may be better at forecasting returns than the competing model.

3.1.3 Market Timing Capability

The Market Timing Capability of the Competing Model was evaluated by calculating the ratio for the number of times the direction of the returns was correctly predicted by the model to total predictions. This ratio's Z Score was then used to evaluate its statistical significance. The

Competing Model had a ratio of 0.51539 of correctly predicting the direction of the returns and a Z Score of 1.5894. At a 5% significance level the null hypothesis that states the ratio for this model correctly forecasting the direction of returns equals 0.5 cannot be rejected. This result implies that the Expanding Window Competing Model has no market timing capability.

3.2 Rolling Window Regression

The table 3.1 summarises the out of sample performance of the Rolling Window Regression with a rolling window of 10 years.

MSE Benchmark Model	MSE Competing Model	Out of Sample R Squared	Delta RMSE
0.0.002066	0.001948	-0.06061	-0.001318

Table 3.2: Rolling Window Regression Out of Sample Performance

3.2.1 Equal Predictive Ability of Benchmark & Competing Model

The out of sample efficacy of both models was evaluated with the Mean Square Error of their out of sample prediction and actual observed value. The competing model has a slightly lower Mean Square Error than the benchmark model, indicating that the competing model performed very slightly better than the benchmark model. The Clarke and West Test statistic tested for the hypothesis that the MSE of the Benchmark is equal to the MSE of the Competing Model. With a CW Statistic value of -0.10913, the null hypothesis is not rejected and the MSE of the Benchmark and Competing Model can be considered not statistically different to each other.

3.2.2 Out of Sample Performance

The Competing Model's out of sample R Squared was negative and extremely low - indicating this model does not perform well on out of sample data. The negative and extremely small delta RMSE between the two models suggests that both models perform very similar to each other on out of sample data. These results suggest going with the benchmark model - the long run mean of the sample - may be better at forecasting returns than the competing model due to simplicity of the model.

3.2.3 Market Timing Capability

The Market Timing Capability of the Competing Model was evaluated by calculating the proportion of correct return direction predictions. With the Competing Model's prediction rate of 0.53378 and Z Score of 3.495, the model significantly outperformed random chance at the 5% significance level considering the Z Score was much higher than the critical value of 1.65. This result suggests that the Rolling Window Competing Model can time the market and using a rolling window of 10 years of data may explain more variation in future returns than an expanding window.

3.3 Analysis

Overall, the rolling and expanding window regression models do not perform well on out of sample data. There might not be enough correlation between monthly values to forecast future returns. Further analysis on higher frequency data such as daily or intraday values may be useful to see whether these models can forecast the 1 day forward returns better with higher frequency data.

4. Optimization for Target Volatility

Continuing this report's quantitative research, the benchmark model 3.1 and competing model 3.2 were used to forecast returns for the UK Stock Index and the GBP/USD currency pair. The goal was to build an optimized globally diverse portfolio consisting of the US stock index, the UK stock index, and the US short term yield Bonds with a target volatility of 10%.

The data for this quantitative analysis was a subset of the original dataset. To accurately forecast 1 month future returns in 2024, data from 1792 till 1940 is not relevant. Using data from 1940 to 2023 gives us 996 data points. By assuming the returns are normally distributed and using the standard deviation of returns calculated in Table 2.1, we can see from the equation below that using 996 data points results in a 0.2% Margin of Error in the point estimate of a return. This size of historical data gives us enough data points to test our model out of sample and is still relevant in terms of the market structure of today.

$$ME = z \cdot \frac{\sigma}{\sqrt{n}} = 1.96 \cdot \frac{0.043}{\sqrt{996}} = 0.2\%$$

4.1 Forecasting FX Returns of GBP/USD

A rolling and expanding window regression for the FX Returns of the GBP/USD currency pair. The ultimate goal is to use the forecast of FX returns to convert the forecast of the UK Stock Index into dollar terms for our optimized portfolio. An Interest Rate Differential Factor was created by subtracting the US short term yield from the UK short term yield. The one period lagged interest rate differential values were then regressed on the FX Returns.

1. Benchmark model

$$e_t = \alpha + \varepsilon_t \tag{4.1}$$

2. Competing model

$$e_t = \alpha + \beta z_{t-1} + \varepsilon_t \tag{4.2}$$

4.2 Expanding Window Models - Out of Sample Performance

Model	MSE Benchmark	MSE Competing Model	OOS R Squared
US Stock Returns	0.001816	0.001798	0.01038
UK Stock Returns	0.002492	0.002454	0.01507
FX Returns	0.0005942	0.0005949	-0.00117

Table 4.1: Expanding Window FX Regression Out of Sample Performance

4.3 Rolling Window Models - Out of Sample Performance

Model	MSE Benchmark	MSE Competing Model	OOS R Squared
US Stock Returns	0.00188	0.001804	0.04011
UK Stock Returns	0.002583	0.002486	0.03765
FX Returns	0.0005978	0.0006068	-0.01504

Table 4.2: Rolling Window FX Regression Out of Sample Performance

4.4 Portfolio Optimization

Monthly Returns data from 1940 to 1950 was used as the initial 10 year window for the rolling and expanding window. With each benchmark and competing model we built a portfolio optimized for a target volatility of 10%. The Portfolio was optimized by minimizing the difference between the volatility of the portfolio and the target volatility by finding and choosing appropriate weights. This Python code for this function is present in the appendix ‘A.

The return matrix consisted of three assets: US Index Returns, UK Index Returns, and the Risk Free Rate. At each trading day, the return matrix consisted of the actual return data up until that point in time. For example, after training the regression model with 10 years of data, the return matrix would include returns from t up until t - 120.

$$\begin{bmatrix} US_{t-120} & UK_{t-120} & RFR_{t-120} \\ US_{t-119} & UK_{t-119} & RFR_{t-119} \\ \vdots & \vdots & \vdots \\ US_t & UK_t & RFR_t \end{bmatrix}$$

Each model was used at time t to forecast the US and UK index return for t+1. This forecast return was then added to our return matrix in order to build a return matrix for t+1. The Risk Free Rate for t+1 was assumed to be the Risk Free Rate at time t. The return matrix for time t+1 would therefore look like this (the last row indicates the forecast return for t+1 at t):

$$\begin{bmatrix} US_{t-120} & UK_{t-120} & RFR_{t-120} \\ US_{t-119} & UK_{t-119} & RFR_{t-119} \\ \vdots & \vdots & \vdots \\ US_t & UK_t & RFR_t \\ US_{t+1} & UK_{t+1} & RFR_t \end{bmatrix}$$

For the expanding window model all historical returns up until time t were used. A new row for the forecast returns was added at the end of this matrix. For the Rolling Window Models all historical returns from t back till t-120 were used. A new row of forecast returns was then added at each iteration. This matrix was used to calculate the covariance matrix at each point in time before rebalancing the portfolio every month. The portfolio variance σ_p^2 was calculated via the formula:

$$\sigma_p^2 = \mathbf{w}'\mathbf{\Sigma}\mathbf{w}$$

where:

- $\mathbf{w}$ is the matrix of portfolio weights
- $\mathbf{w}'$ is the transpose of the matrix of portfolio weights
- $\mathbf{\Sigma}$ is the covariance matrix of the asset returns

The table 4.3 has detailed information on the performance of these strategies over time:

Strategy	Avg Ret	Avg Vol	SR	SO	Performance Fee
Expanding BEN	5.78%	10.13%	0.61	0.23	
Expanding MOD	5.76%	10.14%	0.61	0.23	220 BPS
Rolling BEN	5.53%	10.37%	0.52	0.47	
Rolling MOD	5.53%	10.37%	0.52	0.47	843 BPS

Table 4.3: Performance Metrics of each Rebalancing Strategy

The Expanding Regressions overall outperformed the Rolling Regression in terms of higher average annual return, lower average annualized Volatility, and higher sharpe ratio. The Performance

Fee in table 4.3 represents the fee a trader would be willing to pay per annum to switch from the Benchmark Strategy to the Competing Model Strategy. Given that switching from a Benchmark Strategy to Competing Model strategy would not improve the Portfolio Performance, we would not be willing to pay the extra performance fee for the Competing Models. The Figure 4.1 below shows how the Expanding Benchmark Model beats the Rolling Benchmark Model over time:

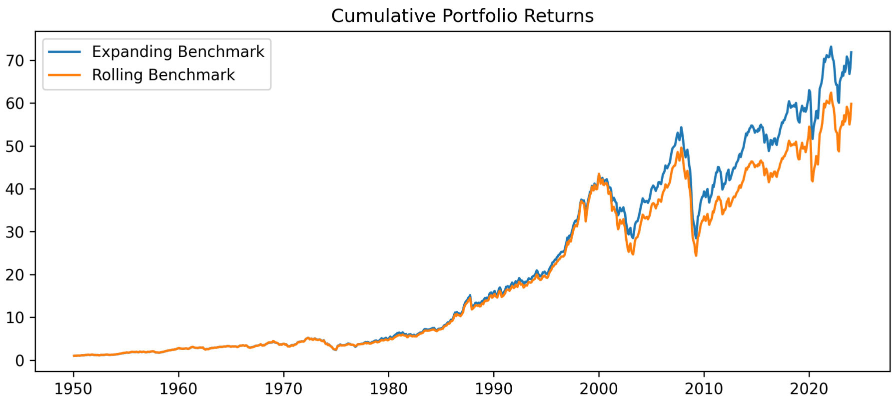

Figure 4.1: Cumulative Portfolio Returns

The Expanding Window Models overall perform better than the Rolling Window Models. Ex post volatility (realized volatility at t+1 after rebalancing the portfolio) was better targeted with the expanding window models. This can be seen in the figures below which present the Cumulative Return, Porttfolio Weights, the 1 Year Rolling Sharpe, and Annualized Volatility over time for each of these models.

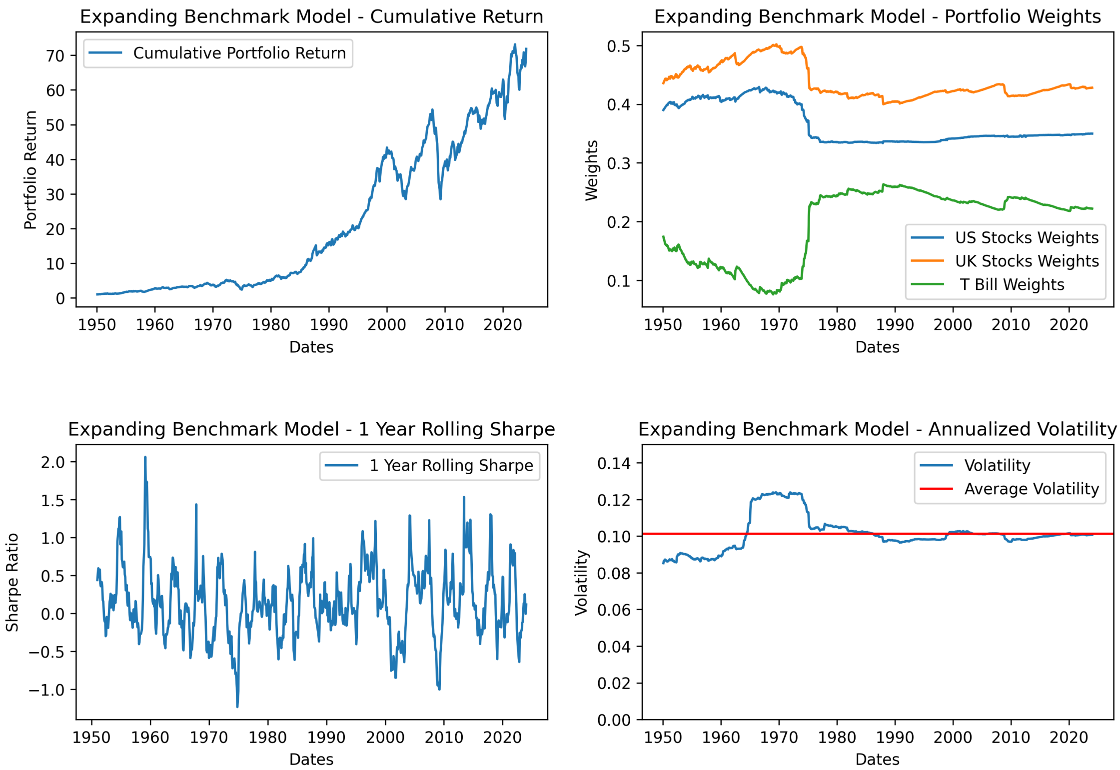

Figure 4.2: Expanding Benchmark Model

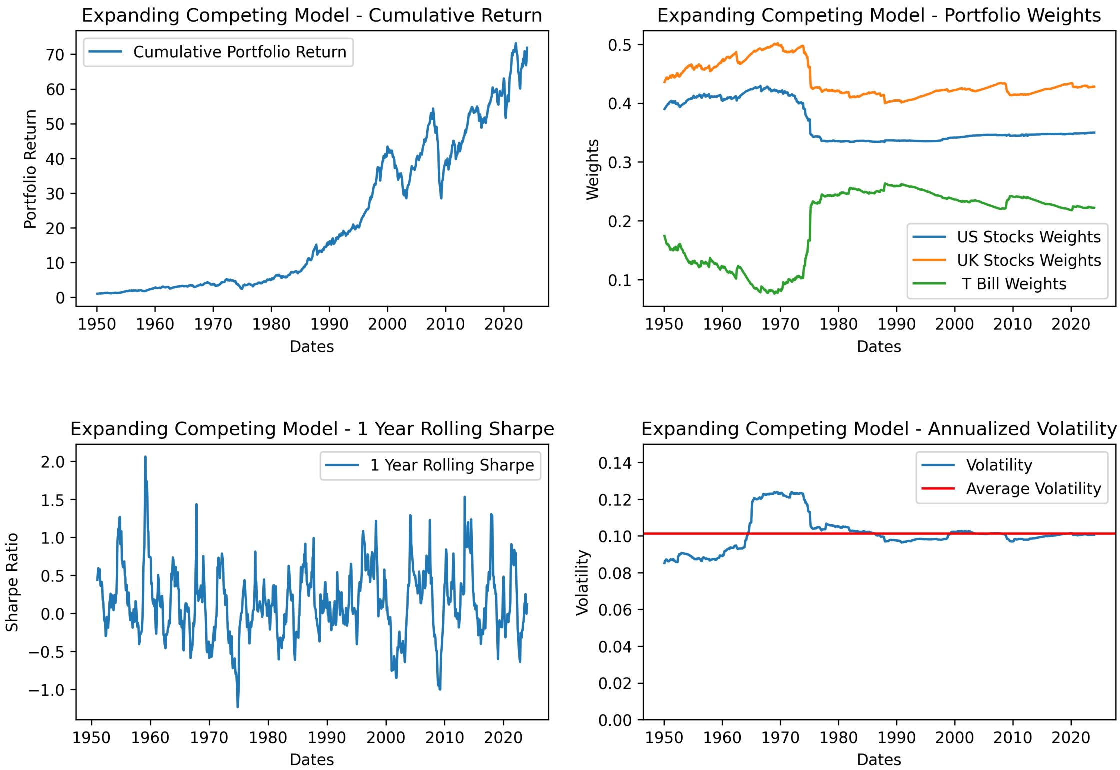

Figure 4.3: Expanding Competing Model

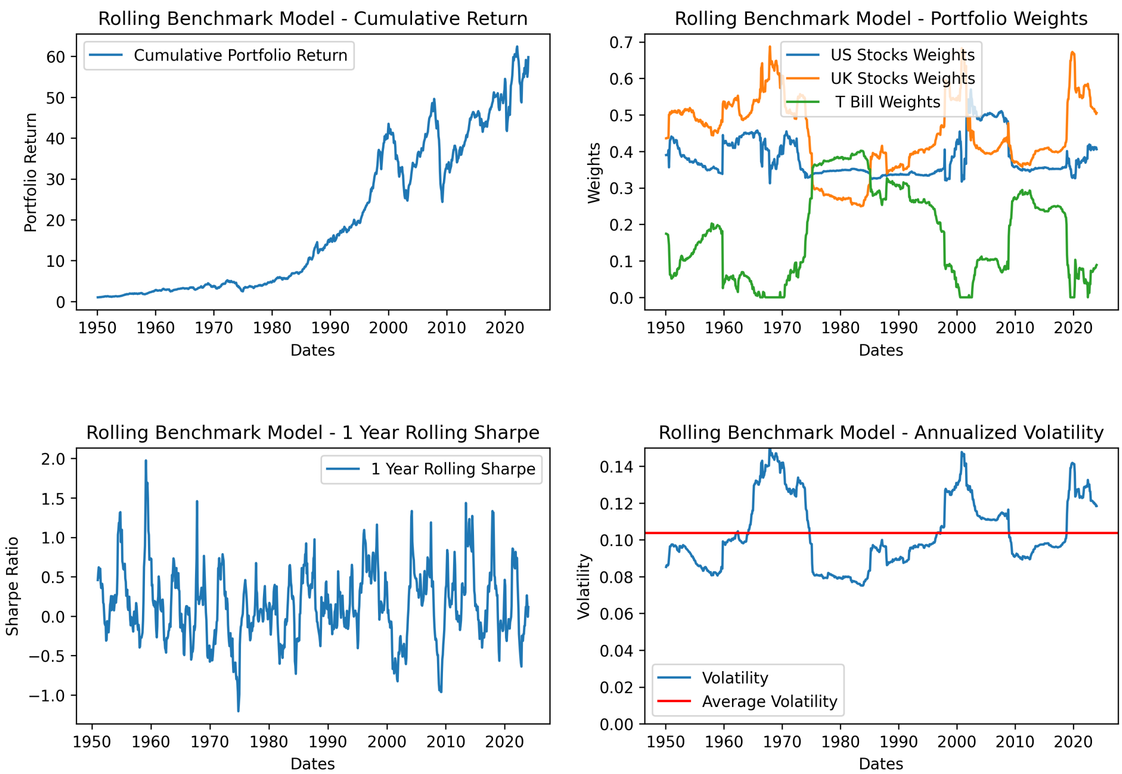

Figure 4.4: Rolling Benchmark Model

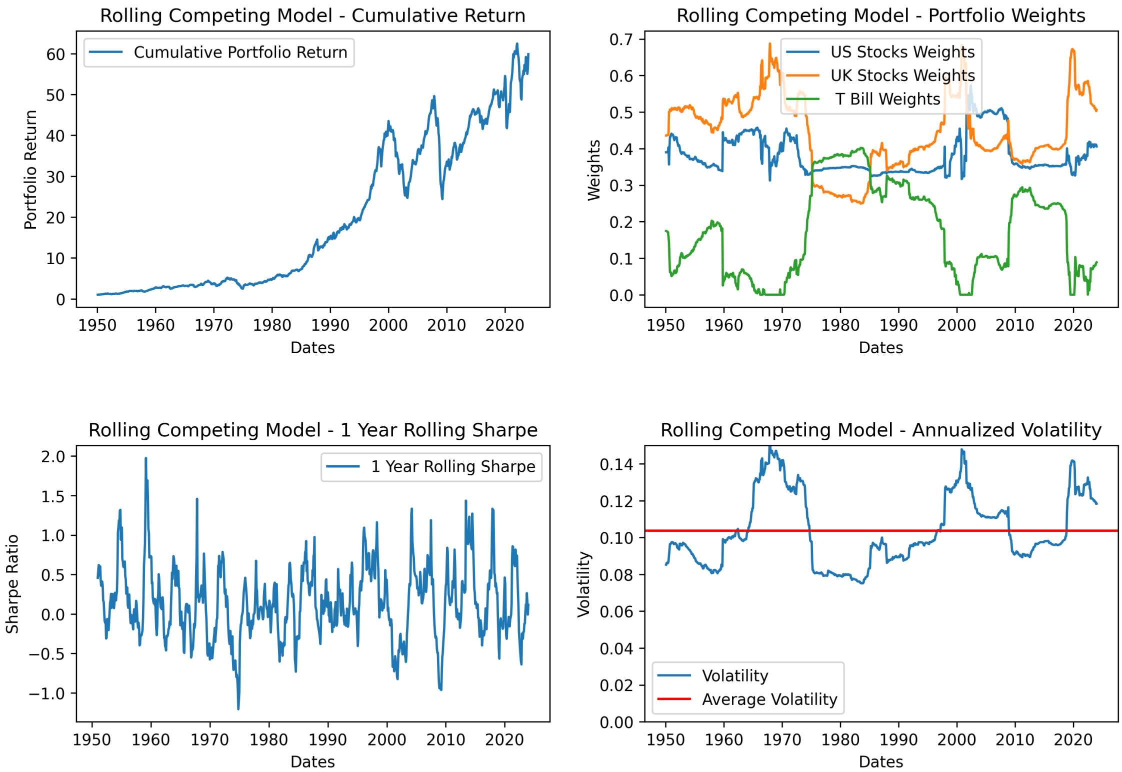

Figure 4.5: Rolling Competing Model

4.5 Conclusion

The Quantitative Research in this report produced 4 simple models to forecast future returns and and build optimized portfolios targeting a volatility suitable to investor's preferences. The simple benchmark models outperformed the competing models overall and would have provided an investor a 70x return on their investment between 1950 and 2023. Further analysis with higher frequency data may lead to identification of better models that are able to forecast future returns better than models using monthly data and can optimized the portfolio to the target volatility with a much lower standard error.

A. Appendix

The below function was used to optimize the portfolios in this research report. The full code for this project can be found at Talha Jamal's GitHub Profile: Empirical Finance Project

```
def target_volatility_optimization(cov_matrix, target_annual_vol):
    n_assets = len(cov_matrix)
    target_monthly_vol = target_annual_vol / np.sqrt(12)
    initial_guess = np.ones(n_assets) / n_assets

    # Objective function: Minimize the absolute difference between the target and actual portfolio volatility
    def objective_function(weights):
        current_volatility = portfolio_volatility(weights, cov_matrix)
        return abs(current_volatility - target_monthly_vol)

    # Constraint: Weights sum to 1
    constraints = ({'type': 'eq', 'fun': lambda weights: np.sum(weights) - 1})
    # Bounds for each weight to be between 0 and 1
    bounds = tuple((0, 1) for asset in range(n_assets))

    # Optimization
    result = minimize(objective_function, initial_guess, method='SLSQP', bounds=bounds, constraints=constraints)

    if result.success:
        return result.x  # Optimized weights
    else:
        raise ValueError("Optimization failed")
```