STRATEGY EVALUATION

1.0 Introduction

This empirical report aims to develop and assess the performance of trading strategies relative to a benchmark. To this end, we develop two strategies – a manual (rule-based) strategy and a machine learning strategy. These two strategies are then compared with each other and against the benchmark strategy. The benchmark strategy is a simple buy-and-hold strategy that purchases a stock and holds the position through the trading period. The buy-and-hold approach (i.e., benchmark) stems from one of the fundamental theories in asset pricing – The Efficient Market Hypothesis (EMH). In its weak form, the EMH posits that the price of a financial asset reflects all historical information, and that technical analysis has no merit. We argue that this may not always be the case and provide empirical evidence to the contrary in this report. We frame the hypothesis for this report around the weak form of the EMH as follows:

$\mathbf{H_0}$: Analysis of historical price patterns CANNOT yield higer performance relative to the benchmark strategy.
$\mathbf{H_A}$: Analysis of historical price patterns CAN yield higer performance relative to the benchmark strategy.

The analysis provides two pieces of evidence to reject the null hypothesis in favor of the alternative view that historical price analysis can yield relatively higher returns provided the investor knows what they are going for. In other words, skill is a differentiating factor. Our manual and machine learning strategies outperform the benchmark strategy in- and out-of-sample across several performance metrics. For all empirical analysis and experiments, we use the stock prices of J.P. Morgan with the ticker symbol "JPM."

2.0 Indicator Overview

We use the following technical indicators derived from stock prices as input variables in our manual strategy (rule-based) and machine learning (strategy learner). For ease of interpretation of the indicators, all indicators are standardized to a z-score. The approach is typically helpful in developing a rule-based strategy and machine learning.

- **Bollinger Bands (BB):** tracks the Simple Moving Average (SMA) to identify significant price excursions from the SMA for a given rolling lookback window. We derive a quantitative variable for BB as

$$\boldsymbol{BB_t = \frac{price_t - SMA_t}{2 \times standard\ deviation_t}}$$

- **Price-to-SMA ratio:** identifies how much the price of a stock moves relative to its long-run mean value. The SMA allocates equal weights to all prices within the rolling window and typically lags actual price movements. We derive the Price/SMA as

$$(\boldsymbol{Price/SMA})_t = \frac{\boldsymbol{price[t]}}{\boldsymbol{price[t-\ window:t].mean()}} - \mathbf{1}$$

- **Momentum:** measures the price changes at time t, relative to a specified lookback period. The formula is given as

$$\boldsymbol{Momentum_t} = \frac{\boldsymbol{price_t}}{\boldsymbol{price_{t-N}}} - \mathbf{1}$$

where N is the lookback period.

The optimal lookback and rolling window need to be determined to maximize our strategy's performance. To this end, we develop a simple SMA crossover trading strategy to evaluate which rolling window, when used, yields maximum performance relative to a simple buy-and-hold (benchmark). This brute force optimization often leads to overfitting the training dataset. We overcome overfitting by dividing the training dataset into four successive equal date ranges and training our crossover strategy on each tranche to identify the rolling window which beats the benchmark. Since different time series tranche may exhibit other behaviors, we average across the optimal rolling window in all four tranches to get the final optimal window size. Our simple crossover trading rule is as follows:

- BUY (= +1) when the shorter SMA is above, the longer SMA
- SELL (= -1) when the shorter SMA is below, the longer SMA.

For the shorter SMA, we try **`range(10, 31, 2)`** and **`range(30, 51, 2)`** for longer SMA range. The optimized rolling window turned to be 15 days. As such, all the SMA window and lookback used in our indicators was 15 days.

3.0 Manual Strategy

Our manual strategy is based on an intuitive interpretation of the chosen indicators to generate a trading signal or decision. The momentum variable indicates the strength of a positive or negative trend in anticipation that the trend will continue. For instance, when the momentum is strongly negative and persistent, it may indicate that selling pressures are driving the stock prices down and may signal a SELL opportunity. However, momentum indicator alone cannot yield a well informed trading decision. In our manual, we combined the indicators mentioned above to create trading signals as follows:

```
   holdings = 0
   new holdings = 0
if ((momentum < -1.0) or (Price/SMA > 1.0)) and BB > 2.0
           new holdings = -1000 # SELL
else if ((momentum > 1.0) or (Price/SMA < -1.0)) and BB < -2.0
           new holdings = 1000 # BUY
```

```
else:
           new holdings = holdings. # NOTHING
```

```
 At any point in the trading period, our total number of shares is
           shares = new holdings - holdings
```

The idea behind our trading rule is to identify points in time when there are pricing inefficiencies and exploit them to make a profit. To make a SELL or BUY decision, we evaluate market participants' interest in the stock by looking at its momentum and how much investors are currently pricing it relative to its long-run value. If investors are over-pricing the stock, its Price/SMA ratio will be greater than 1.0, signaling that it is over-bought. Similarly, price momentum relative to a previous period can indicate the trajectory of prices. The Bollinger band excursions over plus or minus two standard deviations are significant as they suggest that prices have over-shot their long-run mean and may revert. To confirm that we observe market inefficiencies, we construct harder rules/conditions that need to be met before making a trading decision. This ensures that we trade less to avoid unnecessary market impact and commissions. To this end, we only issue a SELL signal when we notice either momentum is less than -1 (signaling waning interest) or Price/SMA greater than 1 (indicating an over-bought) in addition to Bollinger band deviations greater than 2 (indicating mean reversion). We only enter a BUY position when the conditions are reversed and do NOTHING when none of the requirements are met. In each, consider how many shares we currently hold to determine how much we need to BUY/SELL.

Our strategy is practical, conservative, and robust over time as it is market regime invariant. In other words, it is not too sensitive to any specific state of market conditions. For instance, we notice that the in-sample performance around 2008-12 – 2009-10 is remarkably better than the benchmark (see Figure 1 below). That period (which was the peak of the Great Financial Crisis) is characterized by intensive market volatility and over-selling, and our indicators correctly capture the trading opportunities it presents. Table 1 reports some critical metrics on the in-sample performance of our manual strategy relative to the benchmark. It is also noteworthy to mention that our strategy makes only six trades during the entire two years of the in-sample period while still able to turn a profit after considering market impact and commissions on trades (see Figure 1 below).

Table 1: In-sample performance comparison: Benchmark vs. Manual strategy

Performance Metric	**Benchmark**	**Manual Strategy**
Sharpe ratio	0.1573	0.2817
Cumulative return	1.23%	7.97%

Final portfolio value[1]	$101,027.70	$107,965.35
Average daily return	0.017%	0.0258%
Standard deviation of returns	0.0170	0.0145
Number of trades	1	6

Figure 1: In-sample portfolio returns of Manual strategy vs. Benchmark from 2008-1-1 to 2009-12-31. Transaction cost applied: market impact = 0.005, commission = $9.95 per trade.

Our manual strategy was also compared to the benchmark out-of-sample. Table 2 and Figure 2 presents our findings. Benign market conditions with less volatility characterize the out-of-sample period. Our strategy continues to outperform the benchmark in this peaceful market environment, further emphasizing the strategy's robustness.

The cumulative returns of 7.97% vs. 16.04% (almost double) for the manual strategy for in-sample and out-of-sample may be explained by the different market regimes the stock prices cover (i.e., financial crisis vs. post-financial crisis). Intuitively, it is more difficult to turn a profit in trading when there is an ongoing systemic market risk observed during the financial crisis. On the other hand, the period after the crisis (out-of-sample period) was a recovery time for the entire economy with excellent growth potential. Hence, a good trading strategy could easily double its fortune, as observed with our manual strategy.

Table 2: Out-of-sample performance comparison: Benchmark vs. Manual strategy

Performance Metric	**Benchmark**	**Manual Strategy**
Sharpe ratio	-0.2569	0.6672
Cumulative return	-8.35%	16.04%

[1] Initial investment/startup cash is $100,000

Final portfolio value[2]	$91,445.70	$116,024.45
Average daily return	-0.014%	0.033%
Standard deviation of returns	0.0085	0.0077
Number of trades	1	7

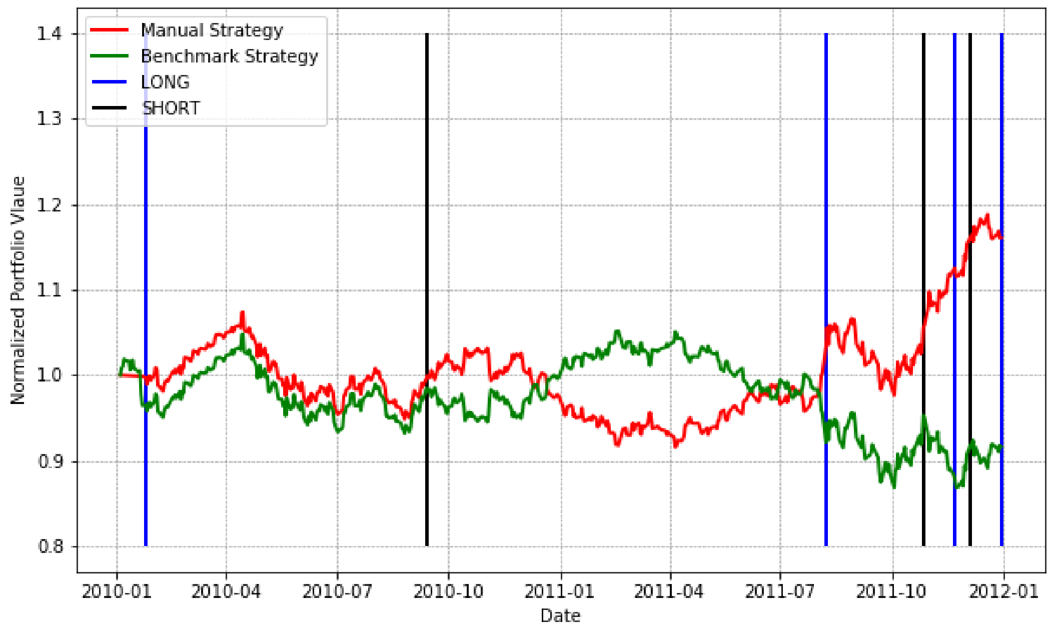

Figure 2: Out-of-sample portfolio returns of Manual strategy vs. Benchmark from 2010-1-1 to 2011-12-31. Transaction cost applied: market impact = 0.005, commission = $9.95 per trade.

4.0 Strategy Learner

We used the Random Forest learner as our choice of machine learning algorithm for making trading decisions. The trading decision was framed as a learning problem with discretized labels as an actionable trading signal. Our strategy learner accepts features in our three technical indicators and maps them onto a categorical target/label series. The categorical labels represent the trading decisions BUY, SELL, or do NOTHING. We construct the labels, features, and hyperparameter optimization as follows.

First, we generate features in the form of indicators from daily stock prices. Same indicators for the manual strategy were used here, namely – BB, Momentum, and Price/SMA ratio. For numerical stability and machine learning purposes, all indicators are standardized to a z-score. We initially start with a 15-day window used in the manual strategy and tune it for optimal performance. However, we vary the window for better in-sample performance. We also

---

[2] Initial investment/startup cash is $100,000

determine the optimal number of bags and leaf size to be used in the Random Forest learner in the training phase.

To this end, we split the training data into two – training (2008-1-1, 2009-6-30) and validation (2009-7-1, 2009-12-31) sets. Manual tuning and optimization of hyperparameters are done on the training set, and the model's performance is evaluated on the validation set. The sequential training on older data and testing on relatively recent data align with time series analysis. This sort of roll-forward validation prevents us from "peeking" into the future. Moreover, the approach also prevents us from using the test data in the training stage. The optimized rolling window was 10-days. The hyperparameters – leaf-size and number of bags -- for the Random Forest learner were at 5 and 15, respectively.

Second, we generate the categorical labels from stock returns by setting a predefined threshold above or below to make a trading decision. From the training dataset, we calculate a 10-days return as $\frac{prices[t+10]}{price[t]} - 1$. The BUY or SELL threshold is predefined as

SELL threshold = 0.02 + market impact
BUY threshold = 0.02 – market impact.

The interpretation is that we will be willing to make a trading decision if we observe return changes above or below 2% plus a market impact of 0.005. The threshold is based on an a priori two standard deviations in returns which is considered significant enough to gain attention. The label data then becomes easy to derive as

SELL = returns > SELL threshold
BUY = returns < BUY threshold
Y LABEL = SELL – BUY.

The Y LABEL (i.e., the categorical target series) then becomes (-1, 0, 1) for SELL, NOTHING, and BUY decisions, respectively.

The prediction of the labels informs our trading decision. The number of shares we buy/sell depends on the current shares we hold. If the Y LABEL = 1 (BUY), we buy 1000 shares, and if Y LABEL = -1 (SELL), we sell 1000 shares. If we are buying, new holding should be +1000 shares (LONG), and the number of shares we need to trade to reach that is (+1000 – current holdings). Similarly, for selling, we will need (-1000 – current holdings). The trading decision can be summarized in a simple logic as

```
Holdings = 0
If  Y LABEL  is in (1,-1):
        New holdings = 1000 if  Y LABEL == 1 else -1000
        Shares = New holdings - Holdings
```

5.0 Experiments

5.1 Experiment 1

In this experiment, we evaluate and compare the in-sample performance of the manual strategy, strategy learner, and the benchmark. Our initial hypothesis is that we expect the machine learning strategy (strategy learner) to outperform the manual strategy and the benchmark in-sample. The hope is that we can exploit more profitable trades due to the sophistication of the machine learning approach. The in-sample period is 2008-1-1 to 2009-12-31. We factor in market impact for metric performance evaluation at 0.005 per share price and $9.95 transaction cost per trade.

Table 3 below reports the performance metrics across all three strategies. The strategy learner outperforms the remaining two strategies with a wide margin. Notably, the strategy learner nearly doubles portfolio value in the in-sample period (see Figure 3 below). The reason is that the strategy learner makes more trades (131 in-sample).

Since the Random Forest algorithm used in the strategy learner randomly picks features (i.e., indicators) to split on, we cannot expect the same relative performance or outperformance every time, even for in-sample data.

Table 3: In-sample portfolio performance comparison: Benchmark vs. Manual Strategy vs. Strategy Learner

Performance Metric	**Benchmark**	**Manual Strategy**	**Strategy Learner**
Sharpe ratio	0.157	0.282	1.774
Cumulative return	1.23%	7.97%	79.94%
Final portfolio value[3]	$101,027.70	$107,965.35	$179,578.40
Average daily return	0.017%	0.026%	0.122%
Standard deviation of returns	0.017	0.015	0.011
Number of trades	1	6	131

[3] Initial investment/startup cash is $100,000

Figure 3: In-sample portfolio performance comparison of Benchmark, Manual Strategy and Strategy Learner from 2008-1-1 to 2009-12-31. Transaction cost applied: market impact = 0.005 and commission = $9.95 per trade.

5.1 Experiment 2

This experiment shows how market impact can affect portfolio performance and, subsequently, trading behavior. The market impact is how much the price of a stock moves against the trader when they buy or sell a stock. Typically, with a buy order, the market impact works such that by the time an order is settled, the price would be higher. The same but opposite occurs with a sell order. This additional cost is due to market slippage. Our hypothesis for this experiment is that we expect to see a reduction in portfolio values when the market impact is high. We explore this hypothesis with varying market impact (i.e., 2.5%, 1.25%, and 0.625%) while keeping commission cost at $0.00 and measuring key performance metrics using the strategy learner.

Table 4 and Figure 4 present our findings. As expected with an increase in market impact reduces portfolio value (see Figure 4). Notably, we also observe that the trading behavior in terms of the frequency of trades changes across different market impacts. For instance, we see that when market impact = 2.5% (see Table 4 below), 70 trades are made as opposed to 120 trades when market impact = 0.625%. The reason is that the strategy learner considers the market impact on determining the threshold by its construction. Therefore, the threshold increases with an increase in market impact. Subsequently, a higher threshold would require higher returns to be able to trigger a trade. Hence, the reduction in the number of trades at higher market impacts.

Table 4: In-sample assessment of market impact using Strategy Learner

Performance Metric	**Impact=0.025**	**Impact=0.0125**	**Impact=0.00625**
Sharpe ratio	-0.711	0.384	1.684
Cumulative return	-38.99%	14.26%	79.28%
Final portfolio value	$61,000.25	$113,718.12	$178,849.81
Average daily return	-0.081%	0.040%	0.122%
Standard deviation of returns	0.018	0.017	0.012
Number of trades	70	95	120

Figure 4: In-sample market impact assessment of Strategy Learner from 2008-1-1 to 2009-12-31. Transaction cost applied: market impact = [0.0025, 0.0125, 0.00625] and commission = $0.00 per trade.