PROJECT 8: STRATEGY EVALUATION

Christopher Jimenez
cjimenez36@gatech.edu

1 INDICATORS

1.1 Price/SMA

A simple moving average (SMA) is the arithmetic mean of a time series for a moving window of a specified size. Comparing price to its SMA requires 2 vectors, so we can condense this information into a single vector by dividing the price by the SMA to create a ratio. If the ratio is > 1 and increasing, then we have a buy signal. If the ratio is < 1 and decreasing, then we have a sell signal. For both Manual and Strategy learners, the lookback window was optimized to maximize performance. A range of 10, 20, 30, 40, and 50 weeks were attempted and 20 weeks produced the best results. The ratio used to determine buy/sell signals was optimized against the in-sample period as well via repeated trials.

1.2 Bollinger Band® Percentage[1]

Bollinger Bands® are a measure of volatility, often used to identify potential shifts in volatility and take advantage of them as trading opportunities. Bollinger Bands® are calculated by taking the SMA (see 1.1) and adding 2 standard deviations to create an upper bound or subtracting two standard deviations to create a lower bound. To compress the information in the Bollinger Bands®, we can use Bollinger Bands® Percentage, which is the difference of the price and lower band, divided by the difference of the upper band and lower band.

This percentage will tell us how far or close the price is from the lower band, which we can use to determine whether to buy or sell. If %B is < 0, then we may want to buy because we expect the price to go up or sell if %B is > 1. For both Manual and Strategy learners, the lookback window was optimized to maximize performance. A range of 10, 20, 30, 40, and 50 weeks were attempted and 20 weeks produced the best results. The ratio used to determine buy/sell signals was optimized against the in-sample period as well via repeated trials.

1.3 Golden Cross Ratio[2]

A Golden Cross uses two SMAs with different size windows, one short-term, one medium-to-long-term and uses points where they cross as a signal to buy or sell. When the short-term SMA crosses the long-term SMA in an upward trajectory, it is a Golden Cross, signaling a buy as a bullish trend is expected, but when it crosses in a downward trajectory, it is used as a sell signal because a bearish trend is expected. To condense the information of the SMAs into a single vector, we can create a ratio by simply dividing the short-term SMA by the long-term SMA. When the ratio exceeds 1.0, then we have a potential buy signal, and when it is less than 1 we have a potential sell signal. For both Manual and Strategy learners, the lookback window was optimized to maximize performance. A range of 10, 20, 30, 40, and 50 weeks were attempted and 20 weeks produced the best results for the short-term. The long-term SMA was a fixed ratio (4x the short-term) to mimic the proportions of the commonly used SMA(50) & SMA(200) windows. The ratio used to determine buy/sell signals was optimized against the in-sample period as well via repeated trials.

2 MANUAL STRATEGY

All indicators selected rely on SMA in some way, so I opted not to make the indicators rely on one another to provide a sell or buy signal. If any one indicator indicated a buy or sell, then that action would be taken. For example, if Bollinger Band® Percentage (BBP) dropped below 0 indicating it is oversold, then a buy was executed (or a sell signal if greater than 1 because it is overbought), regardless of whether the other two indicators supported the action or not. However, both Price/SMA and Golden Cross Ratio (GCR) rely on lagged versions of themselves to provide additional information. For example Price/SMA returns a buy signal if it is equal to or greater than .85 (indicating bullish price movement) AND the previous period was below .85 (and the inverse for a sell signal). By checking the previous period, we can determine the trajectory of the Price/SMA line to predict whether it will intersect and provide a strong buy/sell signal. This allows us to maximize potential returns assuming the trajectory of the ratio continues because we will know to provide an early buy/sell signal to capture as much of the price movement as possible. Finally, GCR follows similar logic to Price/SMA while accounting for longer term trends. GCR provides a buy/sell based on whether an intersection just occurred by looking at the current value of the indicator and that of the previous period. For example, if the ratio is above or equal to 1 (bullish trend), a buy signal is only generated if the previous period was below 1 (and the inverse is true for a sell signal caused by a bearish trend).

Overall, this strategy was simple but effective, as it was able to achieve positive cumulative returns for both in and out of sample tests. This strategy leverages the ability of Bollinger Bands® to factor in volatility, Price/SMA's ability to react to short-term trends in price movement, and Golden Cross Ratio's ability to account for longer-term trends. In Figure 1 below, we can see how the performance of the Manual Strategy compares to the benchmark in the in-sample period, where blue lines represent long entry points and black indicates short entry points. It is evident that the Manual Strategy outperformed the benchmark by a significant margin, with a significant number of trades. The frequency in trades is likely the result of having three separate indicators working together, yet independently, in the sense that any one of them can provide a buy or sell signal, rather than combining them into a single, more potentially conservative metric. The two lines rarely intersect, with almost linearly increasing gains for the Manual Strategy.

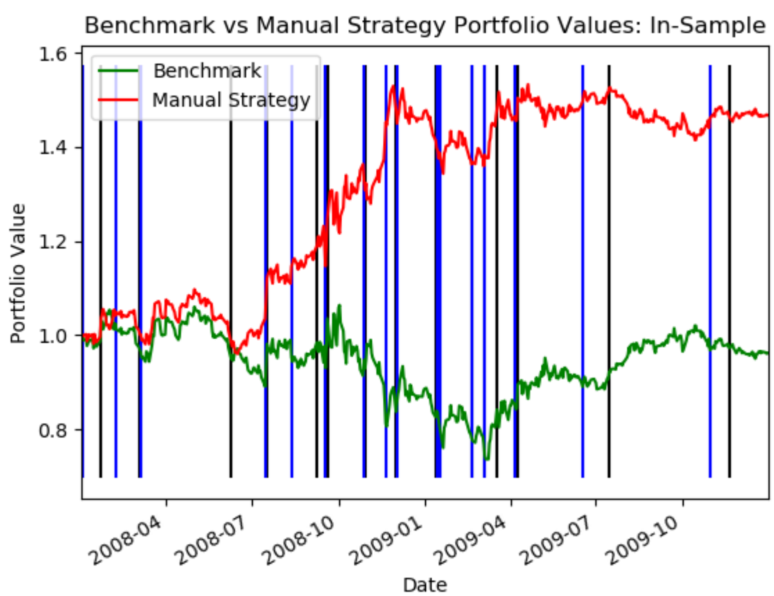

Figure 1—In Sample: Manual Strategy vs Benchmark

In Figure 2, we can see how the Benchmark and Manual Strategy compare across both the in and out of sample periods. Both methods had a significant drop off in performance, although the Manual Strategy was able to generate positive cumulative returns in both scenarios whereas the benchmark was negative in both cases.

Strategy	In Sample Cum Returns	Out of Sample Cum Returns
Manual	+46.62%	+6.71%
Benchmark	-3.79%	-13.37%

Figure 2— Manual Strategy vs Benchmark Performance Comparison

In Figure 3 below, we can see how the performance of the Manual Strategy compares to the benchmark in the in-sample period in detail, where blue lines represent long entry points and black indicates short entry points. The Manual Strategy outperformed the benchmark and achieved positive returns, but by a significantly smaller margin than the in-sample period. Interestingly, there were a couple of periods in which the Benchmark was outperforming the Manual Strategy. Much like the in-sample period, we see a notable amount of trades happening in the out of sample period.

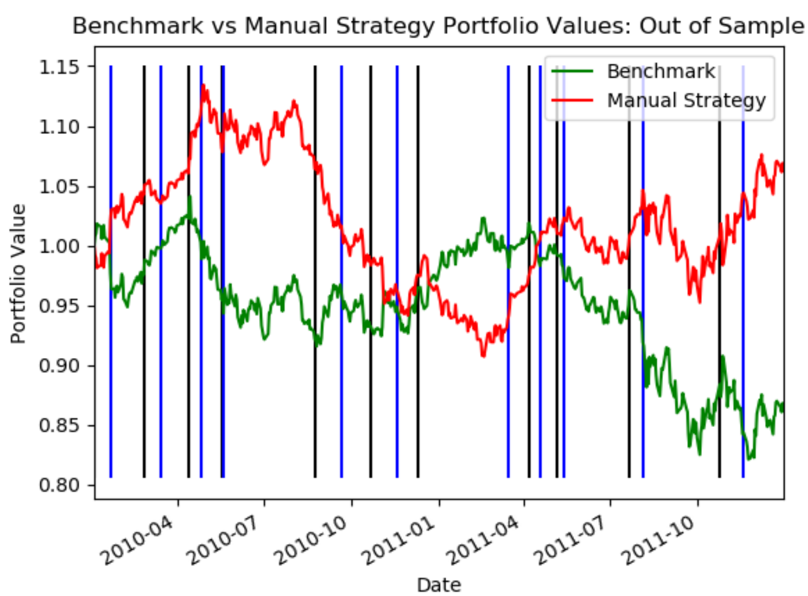

Figure 3—Out of Sample: Manual Strategy vs Benchmark

In Figure 4 below, we can see how the Manual Strategy compared to the benchmark across several metrics. As noted previously, the Manual Strategy saw a significant drop in performance in the out of sample period, which is likely due to overfitting. The Manual Strategy was optimized for the in-sample period, which would not necessarily be indicative of trends in the out of sample period, which are subject to some amount of randomness and externalities. Additionally, the training period and testing periods were equal length. Often, the train/test split skews towards a larger training set to reduce bias. More generally, it is difficult to forecast the behavior of a time series over the course of multiple years on new data. The poor performance of the benchmark can be attributed to randomness since there was no true strategy behind the choice to go long for the duration of both periods. The Manual Strategy consistently outperformed the Benchmark in returns (cumulative and daily mean), but interestingly shared similar levels of volatility in the out of sample period. Although Benchmark had lower volatility overall, likely due in part to only making one trade and not attempting to time the market, it had overall negative gains, so the Manual Strategy is still superior.

	Cumulative Returns		Mean Daily Returns		STDEV Daily Returns	
Strategy	In Sample	Out of Sample	In Sample	Out of Sample	In Sample	Out of Sample
Manual	+46.62%	+6.71%	+30.42%	+1.75%	19.35%	5.18%
Benchmark	-3.79%	-13.37%	-5.71%	-5.08%	6.67%	4.95%

Figure 4— Manual Strategy vs Benchmark Detailed Performance Comparison

3 STRATEGY LEARNER

For my Strategy Learner, I opted to use a Random Forest Classifier (Random Tree Learner + Bagging). To frame this trading problem as a learning problem, I leveraged the same indicators as my Manual Strategy, including using their lagged versions. To turn this into a classification problem, I first had to discretize the possible output because returning a real number would not directly translate to particular trading actions. My labels were {-1, 0, 1}. -1 denotes a sell (short position), 0 denotes "do nothing", and 1 denotes a buy (long position). Next I had to determine where to apply each label to my training data. The criteria I used to classify my training data was the daily returns of n+5 days into the future. The idea behind this criteria is that my indicators should have enough predictive power to approximate the returns 5 days into the future and enable me to make trading decisions based on those expected gains and my current holdings. Through repeated trials using the in-sample data, I landed on if n+5 day returns > .019 + market impact then it would be a buy, n+5 day returns < -.019 - market impact then it would be a sell, and any other situation would be to do nothing. Once I obtained the appropriate labels, I could pass them through to my market simulator to evaluate their performance and tune for optimization.

In addition to deciding on criteria for the labels, the learner itself had additional parameters that needed tuning. The lookback period remained at 20 weeks for SMA calculations for consistency in indicator calculations with the Manual Strategy. The Random Tree learner that was passed to the Bagging learner required testing different leaf sizes, which essentially determines when the tree should stop branching, to prevent overfitting. Smaller leaf sizes lead to deeper trees which have a tendency to overfit. After several trials using different values, I ultimately landed on a leaf size of 5 because increasing it reduced performance significantly and going lower could lead to severe overfitting. Another key parameter was the number of bags for the Bag learner, which essentially determines how many separate Random Trees to create and aggregate to produce an ensemble model. I tested several different values and generally, increasing the number of bags from 20 did not provide a significant performance boost and reducing the number of bags was detrimental to performance. Since there was no increase to performance, I opted to go with the default 20 because I did not want to increase run time by increasing the number of bags to "be safe". Oftentimes, the larger the number of bags, the less overfitting that will occur because you capture more possible scenarios, but there are diminishing returns to increasing the number of bags.

4 EXPERIMENT 1

For this experiment, I compared the Benchmark , Manual Strategy, and Strategy Learner. My expectation going into this experiment is that the Benchmark would have the worst performance since it did not utilize any logic to take advantage of trends in the stock price, that the Manual Strategy would outperform Benchmark because it was optimized using tried and true trading indicators, and that the Strategy Learner would perform the best because it would be able to pick up and learn form patterns that would be difficult or impossible to do manually. See setup information for the experiment below:

- Stock Symbol used: JPM
- Possible positions: -1000, 0, or 1000

- Restrictions on leverage: None
- Sample period: January 1, 2008 to December 31, 2009
- Starting cash: $100,000
- Possible actions: buy or sell only
- Commission: $9.95
- Market Impact: .005
- SMA Window: 20
- # of bags: 20
- Leaf size: 5

Each trading strategy was executed using the above settings on the in-sample period without testing on out of sample data. Final predictions were simulated using a market simulator to determine daily returns, which were then normalized and plotted.

As seen in Figure 5 below, my initial expectations of performance were correct. Benchmark performed the worst of the three, likely due to being solely reliant on randomness/luck for returns and not utilizing available information to make informed trading decisions. The Manual Strategy performed better because it leveraged indicators to inform trading decisions. The Strategy Learner performed the best by a wide margin because of its ability to automatically detect patterns and exploit them. This level of performance on in-sample data is to be expected because this is the data we use to optimize our model. Without testing on out of sample data to inform the final model selected, we run the risk of overfitting because we do not have a view into how well the model generalizes outside of this data set. For example, with the Strategy Learner, if we decreased the leaf size to 1, we could likely boost performance of the model on in-sample data and likely worsen out of sample performance because the criteria for the classification has become too specific. In this sort of situation, I would almost always expect similar results.

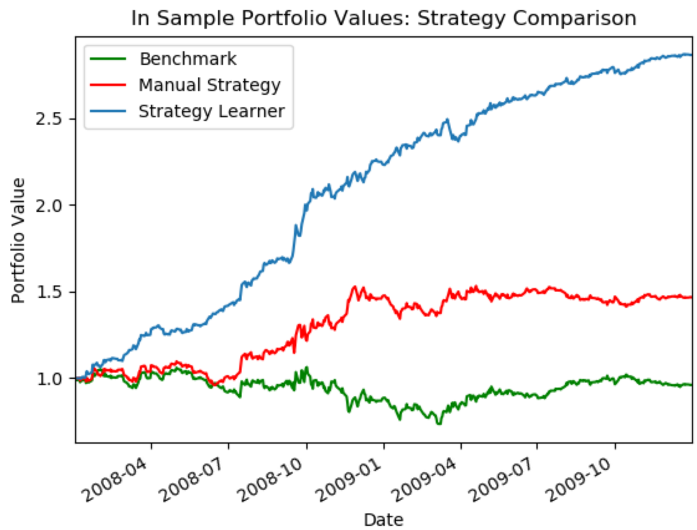

Figure 5— Manual Strategy vs Benchmark vs Strategy Learner

5 EXPERIMENT 2

For experiment 2, we are investigating the influence that market impact has on trading performance for the Strategy Learner. In this experiment, we run the Strategy Learner on the same in-sample period, with zero commission and varying market impact rates. To evaluate how trading behavior changes according to market impact, I tested .000, .005, and .010 impact levels to observe how overall portfolio returns changed, as well as the volatility of those returns. Please see below for experiment set up details.

- Stock Symbol used: JPM
- Possible positions: -1000, 0, or 1000
- Restrictions on leverage: None
- Sample period: January 1, 2008 to December 31, 2009
- Starting cash: $100,000
- Possible actions: buy or sell only
- Commission: $0
- Market Impact: .000, .005, and .010
- SMA Window: 20
- # of bags: 20
- Leaf size: 5

I would expect overall returns to decrease as impact increases because trades with smaller returns would become no longer worthwhile, limiting overall potential. On the other hand, I would expect volatility (standard deviation) of daily returns to decrease as the number of trades would likely decrease, resulting in fewer swings or shifts in daily returns.

In Figure 6 below, we can see how the portfolio value changed across different levels of market impact. The results returned somewhat unexpected behavior. The highest returns were realized at .005 impact, presumably because some of the trades with lower expected returns were not executed and those trades happened to go against expectations, which resulted in gains via reduced trades. However, as expected, the lowest performing model was the .010 level, presumably due to the reduced potential gains from trades that were deemed not worthwhile.

In Figure 7 below, we can see that volatility did not show any meaningful pattern as the lines between the three values of market impact intersected quite frequently, with no front runner in terms of overall volatility across the period. While .005 had the most extreme values, it was not consistently the most volatile, suggesting that the spikes could be considered outliers.

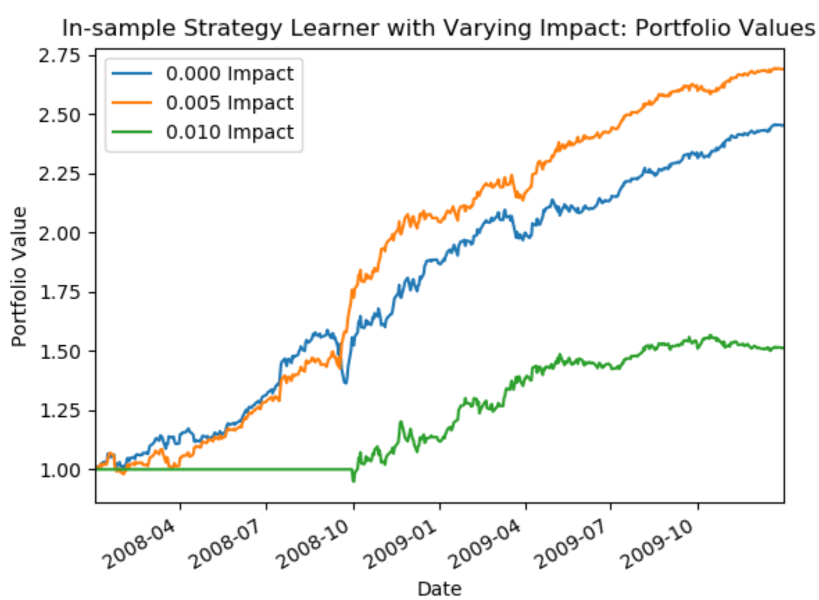

Figure 6— Strategy Learner returns at differing impact levels

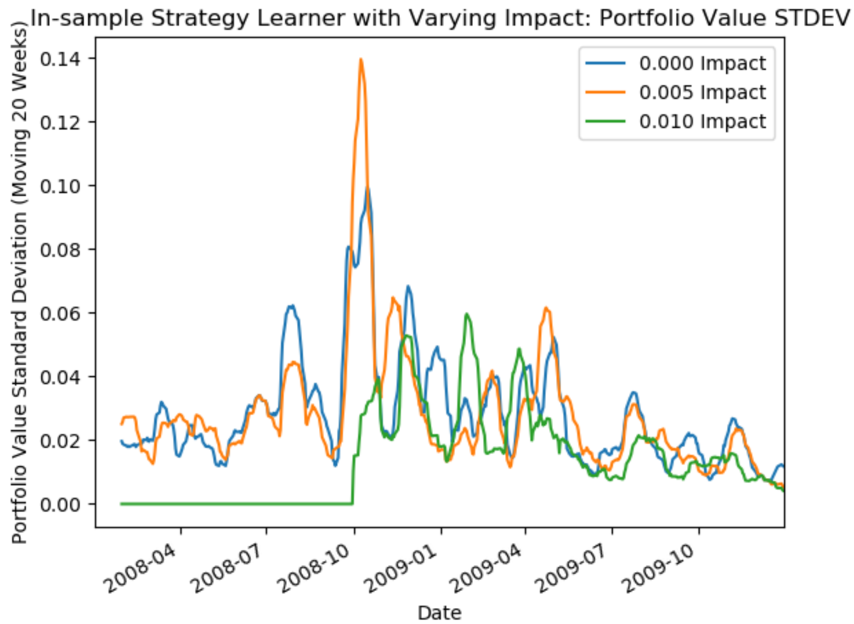

Figure 7— Strategy Learner returns' volatility at differing impact levels

REFERENCES

1. *Technical Indicators: Barchart.com Education*. Barchart.com. (n.d.). https://www.barchart.com/education/technical-indicators/bollinger_bands_percent.
2. Hayes, A. (2021, May 19). *Golden Cross*. Investopedia. https://www.investopedia.com/terms/g/goldencross.asp.