Project 8: Strategy Evaluation

Chad Cortvriendt
ccortvriendt3@gatech.edu

1 INTRODUCTION

In this report I will be comparing the performance of three different trading strategies used on a singular stock, JPM, over an in-sample trading period (1/1/2008 to 12/31/2009) and out-sample trading period (1/1/2010 to 12/31/2011). The 3 strategies that will be explored are a benchmark, manual strategy, and a learner that uses reinforcement learning via a Q-Learner. The benchmark strategy will be simple and will entail purchasing 1000 shares of the stock on the first trading day and holding it for the trading period. The manual strategy will be a trading strategy I design based on values from three trading indicators: SMA ratio, momentum, and Chaikin Money Flow. Lastly, the learner strategy will be developed by using a Q-Learner with the same 3 indicators used in my manual strategy to train on the in-sample trading data. The Q-Learner is a model free learner that chooses actions based on actions rewards and then creates an optimal policy based on maximizing and optimizing actions and rewards. When training the learner, it will continue to update the trading policy until the portfolio performance is no longer improving. The learner will receive positive rewards based on the amount the stock price goes up during a long position or the amount the stock price goes down during a short position. Experiment one will compare the portfolio performance of the three trading strategies across the in-sample trading period. I believe that the benchmark strategy will be the worst performer and the learner strategy will be the best performer. For the second experiment I will explore how different impact values effect the Q-Learner strategy on the in-sample data. Impact is the effect that buying or selling a stock has on the stock price. Thus, if a market participant makes a large order the stock price will move during the order which will incur an impact cost on the market participant (What is impact cost?). For this experiment I will create and train three different Q-Learners one with an impact value of 0, one with an impact value of .02, and one with an impact value of .1 and will test each against the in-sample data. My hypothesis is that the lower the impact the better the portfolio performance will be.

2 INDICATOR OVERVIEW

The 3 indicators that I used to develop my manual and learner strategies are SMA to price ratio, momentum, and Chaikin Money Flow. The SMA ratio is an indicator that calculates the average closing price within a period or range (Hayes). For this project I set the SMA period to 20 days and I calculate the SMA ratio by dividing the closing price of a stock by the simple moving average on that day and then subtracting 1. The SMA ratio can be used with the current price of a stock to indicate if a stock price is a certain percentage above or below the moving average. If the stock price is much greater than the SMA than it might be overbought, so you could short the stock anticipating a price decrease and vice-versa if the stock price is less than the SMA. The second indicator, momentum, is designed to provide insight on a stock price's strength or weakness based on past data. To calculate momentum, you take the closing price of a stock divided by the closing price of the stock from 20 days prior and then subtract 1 (Staff). Momentum can be used to indicate the direction the stock is moving, so if you see momentum shift from negative to positive it could indicate a buying opportunity. The last indicator, Chaikin Money Flow, is designed to use volume and price data to create an oscillator that identifies changes to money flow and trends. The CMF is calculated by taking the 20-day sum of money flow volume divided by the 20-day sum of volume. This oscillator might work as an indicator because not only does it consider the stock price, but it incorporates volume as well and volume can be a strong indicator of buying and selling interest. The oscillator can be used to identify price trends if it begins a negative or positive slope. This could indicate buying and selling triggers. Additionally, these indicators can be used together to confirm if a stock is oversold or overbought. A stock with a high CMF value, near .5, a momentum value greater than 0, and an SMA ratio greater than 0 could indicate an overbought stock and a stock with a low CMF value, near -.5, a negative momentum, and a SMA ratio less than 0 could indicate an oversold stock.

3 MANUAL STRATEGY

In developing my manual trading strategy, I looked at the CMF indicator and noticed that extreme CMF values could be a strong indicator that the stock was overbought or oversold. A CMF value higher than .55 might be followed by a price drop and a CMF value lower than -.55 might be followed by a price increase. I used the SMA indicator as confirmation of stock being overbought or oversold by only initiating a buy if the SMA is less than -.05 which would indicate the stock is oversold and only initiating a short position if the SMA is greater than .05 which would indicate the stock is overbought. Lastly, I used momentum as a third indicator of the stock being overbought or oversold. If the momentum is greater than .1 in addition to the CMF value being greater than .55 and the SMA value being greater than .05 then I concluded the stock was overbought and it was time to initiate a short position. If the momentum is less than -.1 in addition to the CMF value being less than -.55 and the SMA value being less than -.05 then I concluded the stock was oversold and it was time to initiate a long position. I believe that this is an effective strategy for this trading period because these indicators are all confirming the same truth about the stock price while also all using different metrics to get to that conclusion. By using three metrics and not initiating a position until all three reach a certain value prevents the manual strategy from initiating risky positions. Additionally, the fact that the CMF value needs to be greater than .55 or less than -.55 is a very strong indicator that the stock is either overbought or oversold because it is such an extreme value for the period. However, I do not have enough evidence to prove that this is an effective strategy to use outside of this trading period. Since this manual strategy was designed and optimized for the specific in-sample dates, there may be biases or occurrences that only happen in this set of data. Figure 1 shows the portfolio values for the in-sample and out-sample trading periods. As you can see the manual strategy performs about 22%

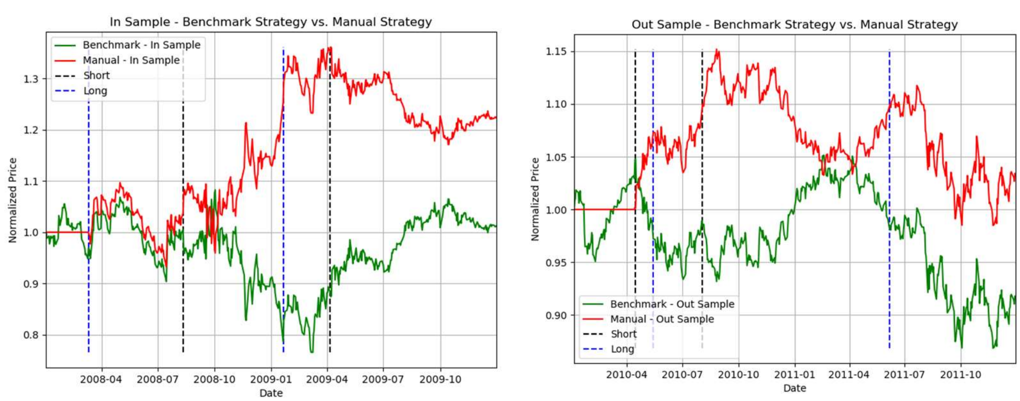

Figure 1—In-sample and out-sample portfolio value comparison.

better than the benchmark strategy on the in-sample period. One of the challenges of developing the manual strategy was that after a substantial loss between 2008-10 and around 2009-01 the strategy begins a long position at the bottom, but it does not hold the position long enough for the stock price to return to the levels it was previously at and it implements a short position around 2009-04 assuming that the stock has been overbought/oversold. On the right chart we see the performance for the out-sample trading period. The out-sample manual strategy also outperforms the benchmark, but by a smaller margin than in the in-sample. The manual strategy is very effective early on, but it fails to implement a long position between 2010-07 and 2011-04. During this window, the benchmark strategy performs significantly better than the manual strategy since the three indicators did not trigger a long position in the manual strategy during this time. Figure 2, below, shows the cumulative return, standard deviation, mean of daily returns, and final portfolio value for the benchmark and manual strategies. In the table you can see that the manual strategy clearly outperformed the benchmark strategy in cumulative return, mean of daily returns, and the final portfolio value. These differences occur because the manual strategy was more effective at anticipating positive and negative JPM price movements than the benchmark, buy and hold strategy. The manual strategies also had a lower standard deviation than the

Strategy	Cumulative Return	Standard Deviation	Mean of Daily Returns	Final Portfolio Value
In-Sample Benchmark	0.012324	0.017041	0.000168	1.012324
In-Sample Manual	0.223189	0.013970	0.000496	1.223189
Out-Sample Benchmark	-0.083579	0.008500	-0.000137	0.916420
Out-Sample Manual	0.032648	0.007279	-0.000090	1.032648

Figure 2 — Table comparing the performance of each portfolio.

benchmark strategies which indicates that the manual strategy is less volatile. The reason for this is that the manual strategy has a predictive approach which attempts to make positive portfolio gains during both long and short positions instead of just holding. Since the manual strategy was able to anticipate price movement better than the benchmark each day was more likely to see a gain in value when compared to the benchmark strategy which results in a lower standard deviation and a more consistent daily return. There are also noticeable differences between the performance of the portfolios between the in-sample and out-sample trading periods. As you can see the in-sample manual strategy performed significantly better than the out-sample manual strategy in terms of cumulative return, mean of daily returns, and final portfolio value. The reason for this is because the manual strategy was designed to fit the in-sample trading data. The key indicator values that I used to define a buy and sell trigger were designed to fit the peaks and valleys of the in-sample data. The out-sample data has different peak and valley values for the indicators and is not identical to the in-sample data. However, it should be noted that the manual strategy did outperform the benchmark strategy in both the in-sample and out-sample periods so while the performance is not nearly as impressive for the out-sample period it did outperform the benchmark and that's likely because the indicators provided decent insight on when the stock was oversold or overbought. Also, it is worth noting that the manual strategy for the out-sample had a lower standard deviation than the in-sample which implies it was a less volatile strategy. This is because over the trading period the cumulative return was only 3% higher than the initial investment which means the mean of daily returns is closer to zero and the standard deviation is closer to zero when compared to the in-sample manual strategy which saw a 22% gain in value and has a higher standard deviation and mean of daily returns.

4 STRATEGY LEARNER

To develop a Q-Learner that could be used to create and learn an optimal trading strategy I had to wrap the trading data and indicator data in a format that the learner could interpret and optimize during the add evidence phase. As mentioned before the indicator values that I used for this project are SMA, Momentum, and CMF. The first step in this process was gathering the indicator values into Pandas DataFrames, one DataFrame for each indicator. From there I needed to discretize the data and save the discretization bins used on the in-sample data to be used on any out-sample data. I retain the bin values so that I can see how the learned strategy could be applied to any out of sample trading periods. Skipping this step and re-discretizing the data in the test policy function on the out-sample data would be modifying the learner and would not actually test the learner's policy. The reason discretization of the indicators is important in developing the Q-Learner is because this learner is dependent on three things: states, actions, and rewards. Actions are the decisions that the learner can make, in developing a trading strategy and they are: buy, sell, and do nothing. The reward is the change in value of the portfolio each day. And the state is the current observed position or status of the system. The Q-Table that is created during the learning process of a Q-Learner depends on the state as way to keep track of the current status of the system and will use it to see what actions yielded what rewards in previous experiences to determine the optimal action for a given state to maximize rewards. To create a usable state value based on the indicators I needed to convert the daily real number value of each indicator to an integer which could then be concatenated into a state ID and this is done via discretization. For example, for a given trading day if the SMA value is 0.051, the momentum value is -0.034, and the CMF value is -0.156, the discretization process on the in-sample data will return a bin value of 15 for SMA, 7 for momentum, and 7 for CMF. This will then be combined into state 1577 which can be passed to the learner as the current state on this trading day. The actions and rewards that learner optimizes are all based on this state value. As mentioned, the Q-Learner uses the combination of these three states and corresponding actions to determine which action at a given state will yield the highest reward. To calculate the rewards for a given day, I take the action generated from the learner and evaluate the change in portfolio value based on that action being performed. For example, if the learner returns a buy action then the reward is the difference between today's price of the stock and the impact price of the stock on the day the learner performed the buy action (previous day) times the number of shares purchased minus the commission rate. This reward calculation computes the impact price of the stock by multiplying the purchase price times 1 + the impact value. The reward calculation also takes into consideration whether we are already in a long or short position in case the Q-Learner returns a 'do nothing' action so that the change in portfolio value and reward is still returned for positions held over multiple trading days. The reward calculation for positions held during the 'do nothing' action does not need to account for impact price or commission because those costs have already been incurred and calculated into the reward during the original buy or sell action. Once the state and reward for the previous action are calculated these are passed into the query function of the learner to update the Q-Table and continue learning until convergence is reached and the portfolio value is no longer increasing. There are several hyperparameters that needed to be established when developing the Q-Learner. The first is the number of states. This value is calculated based on the number of indicators used and the number of bins created in the discretization process for each indicator. In developing the Q-Learner I set my initial bin size for discretization to 10, however after some experimentation I concluded that 20 bins yielded better portfolio performance. With a bin size of 20 and three indicators the number of states for the learner is 191920. The number of actions for the learner is based on the possible decisions to be made on a given trading day which are long, short, or do nothing so 3 actions. Another hyperparameter of the Q-Learner is the alpha value which determines the learning rate, or the weight given to new experiences compared with past Q-values. I ran several different experiments with this value on the in-sample trading period and concluded that a value of .4 returned the best portfolio results from the values I tested. The gamma hyperparameter is the discount factor or the impact that future rewards have on the learner. A low value of gamma means later rewards are valued less significantly than a reward right now. A high value of gamma meats later rewards are valued just as much as a reward right now. In my research on the in-sample data a gamma value of .7 returned slightly better performance than .9 which means more emphasis is being put on current rewards than later rewards when compared to the .9 gamma value. The rar hyperparameter of the Q-Learner determines the probability of selecting a random action for any given state and the radr parameter is the random action decay rate which determines how the probability is affected after each update. A low radr means the probability of random actions gets smaller more quickly as the learner is updated. A high radr means the probability of random actions does not decrease as much as the learner is updated. I set the rar and radr hyperparameters to their typical values which is .98 and .999, respectively. The last hyperparameter of the Q-Learner is dyna which is designed to hallucinate a given number of experiences instead of going through each experience to improve learning time when training the learner. Since learning time is not a substantial focus for my experiments I decided to omit the dyna function of the Q-Learner by passing a 0 value for this hyperparameter.

5 EXPERIMENT 1

In the first experiment I will compare the portfolio values of the benchmark strategy, manual strategy, and strategy learner over the in-sample trading period. My hypothesis for this experiment is that my strategy learner, the Q-Learner, will be the best performing strategy and the benchmark strategy will be the worst performing strategy. The details for this experiment are as follows. The strategies will be tested against stock symbol JPM, the in-sample trading period is January 1, 2008 to December 31, 2009. The starting cash for the experiment is $100,000 and the only allowable positions for the strategies are 1000 shares long, 1000 shares short, and 0 shares. The strategies can place a daily maximum order of 2000 shares when switching from a short position to a long position or a minimum order of -2000 shares when switching from a long position to a short position. The only types of trades allowed are buy and sell and selling after holding no position with the stock or selling 2000 shares to exit a long position is equivalent to shorting. There is no limit on leverage and the commission for each trade is $9.95 and the impact is .005. When placing an order, the price of the stock on a given day is the adjusted close value on that day. As mentioned above, impact is the effect your trade has on the price of the stock. So, when one of the strategies places a buy order the cost of the order is the adjusted close on that day plus 1 plus .005 (impact value) + 9.95 (commission). The benchmark strategy is buying 1000 shares of JPM on the first trading day and holding that position for the trading period. The manual strategy and strategy learner use the following indicators and parameters to make trade decisions. A 20-day SMA to price ratio, a 20-day momentum calculation, and a 20-day Chaikin money flow (CMF) calculation. In my calculations for CMF, I added some smoothing by setting the daily low and daily high to the period low over 20 days and the period high over 20 days. Figure 3, below, shows the equations used for these indicators. The manual strategy executes a long

SIMPLE MOVING AVERAGE

$$SMA = \frac{A_1 + A_2 + A_3 + \cdots + A_n}{n}$$

MOMENTUM

Momentum = Adjusted close / Adjusted Close 20 Days Prior

CMF

Money Flow Multiplier = [(Close-Low) – (High-Close)]/(High-Low)

Money Flow Volume = Money Flow Multiplier * Volume for the Period

20 Period CMF = 20 Period Sum of Money Flow Volume / 20 Period Sum of Volume

Figure 3 — The equations used to generate daily indicator values.

position if the CMF value for a day is less than -.55, the momentum on the same day is less than -.1 and the SMA to price ratio on the same day is less than -.05. To initiate a long position, if there is currently no position held on the stock then it executes a buy order of 1000 shares. If there is currently a short position then it initiates a buy order of 2000 shares. If there is currently a long position then it does nothing. The manual strategy executes a short position if the CMF value for a day is greater than .55, the momentum on the same day is greater than .1 and the SMA to price ratio on the same day is greater than .05. To initiate a short position, if there is currently no position held on the stock then it executes a sell order of 1000 shares. If there is currently a long position then it initiates a sell order of 2000 shares. If there is currently a short position then it does nothing.

Lastly, the strategy learner uses a Q-Learner designed with the specifications mentioned earlier in the report. The hyperparameters for the learner are as follows: Discretized SMA ratio, momentum, and CMF values into 20 bins. 191920 for the number of states, 3 actions, .4 alpha, .7 gamma, .98 rar, .999 radr, and dyna of 0. The reward is the change in value of the portfolio each day and the actions are buy, sell, and do nothing. Additional details for the Q-Learner are outlined in the Q-Learner section of the report. Figure 4, below, shows the portfolio values based on this experiment. As you can see the learner far outpaced the benchmark and manual trading strategies. The learner is designed to find optimal strategies based on the hyperparameters and it clearly does a better job than the manual and benchmark strategy. The number of experiences and amount of information that the learner can process is far superior to my data analytics ability used in developing a manual strategy. I would expect the strategy learner to outperform the manual and benchmark strategy for almost any in-sample period passed through it and the reason is because it is able to iterate through the data numerous times to try to find the best action and reward for each daily state. It will

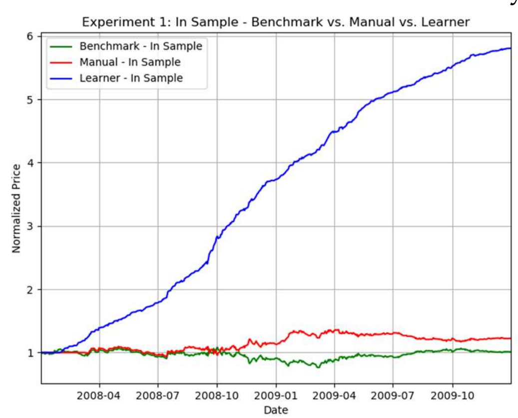

Figure 4—The portfolio value of the benchmark, manual, and learner strategies for the in-sample trading period.

ultimately find a more optimal strategy than the manual and benchmark. Additionally, my manual strategy was designed to statically fit this in-sample dataset and would unlikely perform as well on other in-sample trading periods because it was not designed to fit the data.

6 EXPERIMENT 2

For this experiment I will evaluate how different impact values will affect the Q-Learner trading strategy. I hypothesize that the lower the impact value the better the Q-Learner will perform. The reason for this is because a lower impact value will encourage the learner to make more frequent trades because the cost of the trade will be less which will result in a higher reward. And based on Grinold's Fundamental Law of Active Management where performance equals skill times the square root of breadth where breadth is the number of investments then I suspect the higher number of trades will equal greater breadth and greater performance. To perform this experiment, I will train three different Q-Learners on the JPM stock, between 1/1/2008 to 12/31/2009, with a commission of 0, starting value of $100,000. The only allowable positions for the learners are 1000 shares long, 1000 shares short, and 0 shares. The strategies can place a daily maximum order of 2000 shares when switching from a short position to a long position or a minimum order of -2000 shares when switching from a long position to a short position. The only types of trades allowed are buy and sell and selling after holding no position with the stock or selling 2000 shares to exit a long position is equivalent to shorting. The first learner will have an impact of 0, the second learner will have an impact of .02 and the third learner will have an impact of .1. The hyperparameters of the learners are the same as experiment 1. The 3 learners, our measurements, will be evaluated on two metrics, portfolio value and trade frequency. Figure 5, below, shows the portfolio values and the trade frequency of the 3 learners. As you can see the learner with an impact of 0 outperformed the

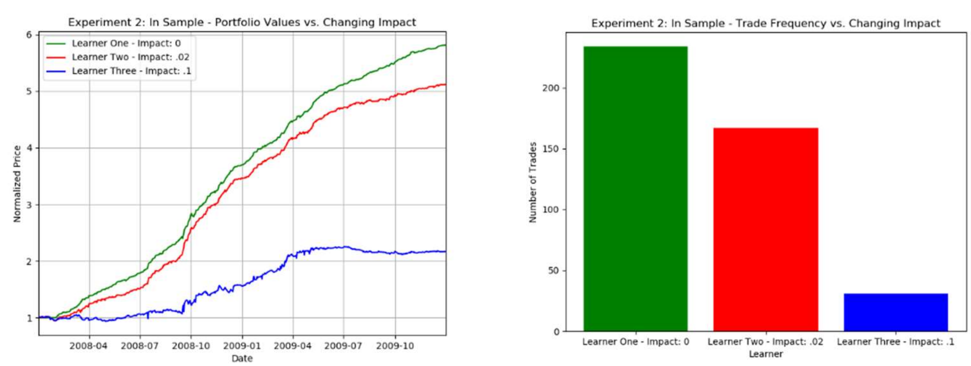

Figure 5—Portfolio value and trade frequency of the learners.

other learners and had a much higher frequency of trades. This confirms my hypothesis as well as Grinold's Fundamental Law of Active Management equation that a higher frequency of trades will result in better performance.

7 REFERENCES

1. Hayes, A. (2021, March 01). Simple moving average (sma) definition. Retrieved March 02, 2021, from https://www.investopedia.com/terms/s/sma.asp
2. Staff, I. (2021, January 07). Momentum indicates stock price strength. Retrieved March 03, 2021, from https://www.investopedia.com/articles/technical/081501.asp#:~:text=Calculating%20Momentum&text=Market%20momentum%20is%20measured%20by,plotted%20around%20a%20zero%20line.
3. What is impact cost? Definition of impact cost, impact cost meaning. (n.d.). Retrieved April 11, 2021, from https://economictimes.indiatimes.com/definition/impact-cost