Project 8: Strategy Evaluation

Cameron Railton
crailton3@gatech.edu

Project 8 builds a framework to test how different technical indicators and trading strategies complement each other. In my implementation, I use a set of indicators derived from the price history: a Weighted Moving Average (WMA) ratio, Bollinger Bands percentage (BBP), momentum over a fixed lookback window, Relative Strength Index (RSI), and a Moving Average Convergence Divergence (MACD) signal. These indicator series form a daily feature vector that drives both strategies. The Manual Strategy applies a weighted voting scheme across these indicators to map each day into a target position of -1000, 0, or +1000 shares of JPM, while the Strategy Learner uses the same features to learn this mapping directly from data. My initial hypothesis is that, in-sample, the Strategy Learner will achieve the highest cumulative return, followed by the Manual Strategy and then the buy-and-hold benchmark and for the out of sample the learned strategy should still perform best, followed by the manual strategy, and the benchmark performing worst.

Indicator Overview

This project focuses on a single, consistent set of technical indicators that are used in both the Manual Strategy and Strategy Learner. The Manual Strategy turns these indicators into discrete trading rules via a weighted vote, while the Strategy Learner consumes the same indicator values as features. Here are the indicators chosen from project 6.

Weighted Moving Average (WMA) Ratio

The WMA gives more weight to recent prices, making it more responsive than a simple moving average. I compute a 12-day and 30-day WMA, take their difference, and divide by the current price to get a relative spread, then pass it through **tanh to k**eep it roughly in [−1, 1]. In the Manual Strategy this normalized trend score is one of the weighted inputs to the composite signal; in the Strategy Learner it is discretized into three bins (downtrend, neutral, uptrend) and used as part of the Q-learner's state.

Bollinger Bands Percentage (BBP)

Bollinger Bands measure how far price is from its recent average in units of standard deviation. I compute a 20-day BBP, which is near 0 at the lower band and near 1 at the upper band, then transform it with 1 − 2 * BBP so values near the lower band map to +1 (long), mid-range to 0 (cash), and near the upper band to −1 (short). In the Strategy Learner, this same BBP-based signal is discretized into low / neutral / high bins and used as another feature in the state.

Momentum

Momentum measures how fast price is changing. I use a 7-day momentum defined as price(t) / price(t − 7) − 1, then multiply by 10 and apply tanh to keep it in roughly [−1, 1] and comparable to the other indicators. In the Manual Strategy this signal is computed but given weight 0 in the best configuration, so it does not affect the final composite; in the Strategy Learner, the raw momentum value is still used as a feature and binned into low / neutral / high for the Q-learner's state.

Relative Strength Index (RSI)

The RSI measures how quickly prices are rising or falling to identify when a stock may be overbought or oversold. It compares the size of recent gains to recent losses over a set period usually 14 days and converts that information into a number between 0 and 100. In the manual strategy I first compute a 10-day RSI from my indicators file, which gives a value between 0 and 100. I then transform it into an RSI score using (50–RSI)/20 so that values near 50 map to 0 (neutral), oversold levels around 30 map up toward 1 (long), and overbought levels around 70 map down toward −1 (short). I clip this score to stay between −1 and 1 and fill any missing values with 0 so early-window noise doesn't distort the signal. This normalized RSI score is then one of the weighted components in my Manual

Strategy. In the Strategy Learner, where it gets discretized into three bins (long, neutral, short) as part of the Q-learner's state.

Moving Average Convergence Divergence (MACD)

MACD tracks the strength and direction of the trend. I compute the MACD line (fast EMA − slow EMA), a signal line (EMA of MACD), and use the MACD - signal histogram as a trend-strength measure. I normalize this histogram by a 30-day rolling standard deviation and apply tanh so it lies roughly in [−1, 1]. A grid search set its Manual Strategy weight to 0, so it does not contribute to the composite there, but the MACD score is still computed and passed to the Strategy Learner, where it is binned into short / neutral / long and included in the discrete state.

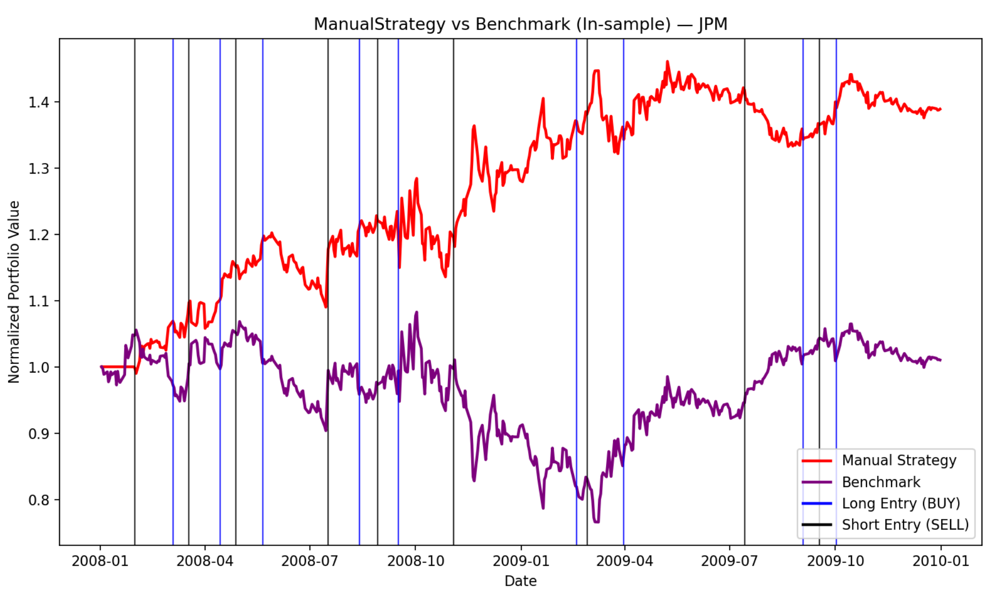
Manual Strategy

My manual strategy is built around a single composite signal that aggregates the five technical indicators mentioned above into a normalized score, which is then translated into discrete trading positions of −1000, 0, or +1000 shares. I calculate all five Project 6 indicators (WMA trend, Bollinger %B, RSI, momentum, and MACD), but I don't hand-pick their weights. Instead, I run a grid search over different weight combinations and buy/sell thresholds. I then choose the setup that gives a good in-sample Sharpe ratio and cumulative return, while also avoiding strategies that trade too aggressively. The optimal weighting found by this strategy is (WMA:0.35, BBP:0.2, RSI:0.2). Interestingly the optimal weights for Momentum and MACD were set to 0 and do not contribute to the final signal. Momentum and MACD likely received zero weight because, in this setup, they didn't add useful information beyond what WMA trend, Bollinger %B, and RSI were already capturing. All three of those core indicators already provide a strong mix of trend, volatility, and overbought/oversold signals. By contrast, momentum and MACD are somewhat redundant with the trend signal and, depending on their parameters, can introduce extra noise or trigger unnecessary trades. Based on this composite value I apply a threshold also found using grid search which was found to be 0.25 and -0.25. Where a value greater than 0.25

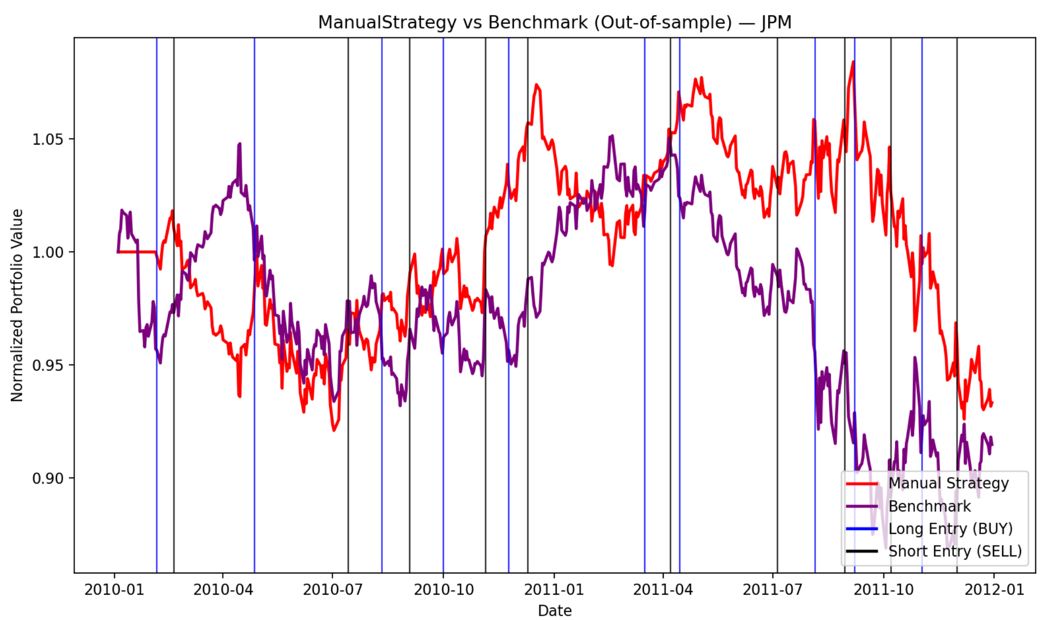
would make the strategy hold a long position, less than -0.25 would cause the strategy to go short, and a value in between would cause the strategy to hold cash. Intuitively, this design tries to go long when the stock is cheap and recovering in an improving trend (low BBP, low/normal RSI, positive WMA spread) and go short when the stock is expensive and weakening (high BBP, high/normal RSI, negative WMA spread), while the dead zone reduces churn from minor, conflicting indicator movements, which I expect to make the rule more robust than acting on any single indicator in isolation.

In Sample Performance

The in-sample period includes the financial crisis and sharp rebound in an environment where mean reversion and trend-following signals tend to be highly predictable. As expected, the manual strategy clearly outperforms the benchmark. The manual strategy achieves a cumulative return of about 38.9%, compared to just 1% for the buy-and-hold benchmark, while also exhibiting lower daily volatility and more than four times the mean daily return of the benchmark. In the in-sample equity chart below, the red line represents the Manual Strategy, the purple line represents the benchmark, and the blue and black vertical lines indicate long and short entries, respectively.



There are a few periods where the strategy clearly seems to guess wrong. It goes long right before a sharp drop and the price continues to fall, and it also sometimes goes short right at the start of a sustained upward move. These mistakes are largely driven by how the indicators are combined. When the strategy enters a long into a continuing sell-off, it's usually because the mean-reversion signals are strong: Bollinger BBP is very low, RSI is oversold, and recent momentum has turned negative for several days. Taken together, these suggest the stock is over stretched to the downside and due for a bounce, even though the broader trend (as captured by the WMA-based trend filter) is still down. In a true downtrend, those oversold conditions can persist, so the strategy ends up trying to catch a falling knife. The opposite happens on the shorts: the stock becomes overbought (high BBP, high RSI, extended positive momentum), and the indicators flag it as a good mean-reversion short before the trend filter fully flips to an uptrend. In a strong bull leg, price can stay overbought for a long time, so the strategy shorts too early and then gets squeezed as the upward move continues.



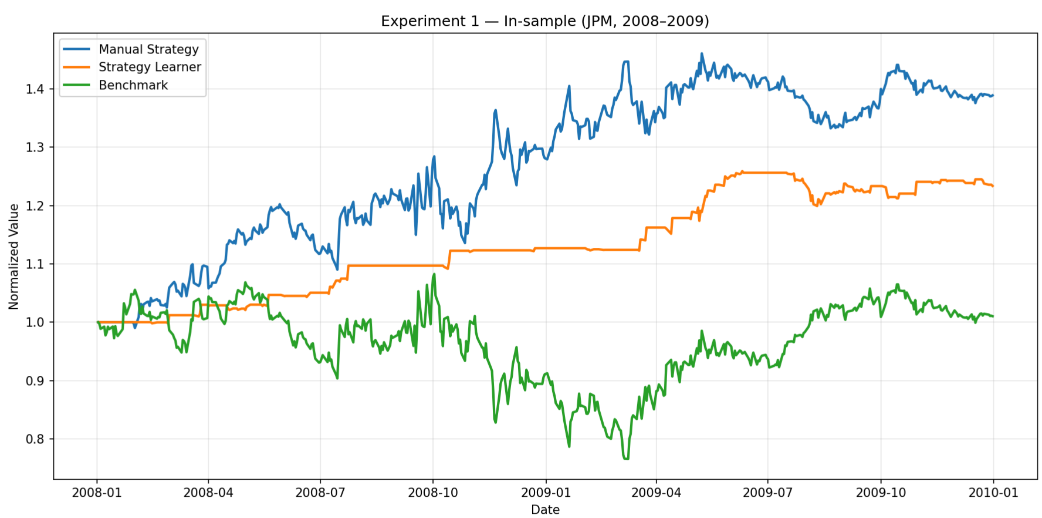
For out of sample testing (2010 - 2011), all parameters are frozen with no tuning or retraining and it is applied to JPM in a calmer, more sideways market. In this mildly downward regime both strategies lose money, but the manual strategy loses less, outperforming the benchmark by about %1.8 while also maintaining slightly lower volatility. The out-of-sample equity curves show that the manual strategy often reduces exposure when the benchmark remains fully long and occasionally flips short during declining windows, but it does not capture trends as cleanly as it did in the in sample data. We see similar issues in the out of sample test as the strategy seems to keep trying to catch the falling knife. The strategy could be improved by strengthening its trend and confirmation logic and adding basic risk controls. For example, you could require the WMA trend and MACD to agree on an uptrend before taking long mean-reversion signals (and similarly for shorts in a downtrend), and use Bollinger band width as a volatility filter so trades in very high-vol regimes only trigger on more extreme RSI/BBP values. Finally, adding simple stop loss or time-based exits could potentially help limit losses in periods where oversold or overbought conditions persist within a strong trend and the indicators stay wrong for longer.

Strategy Learner

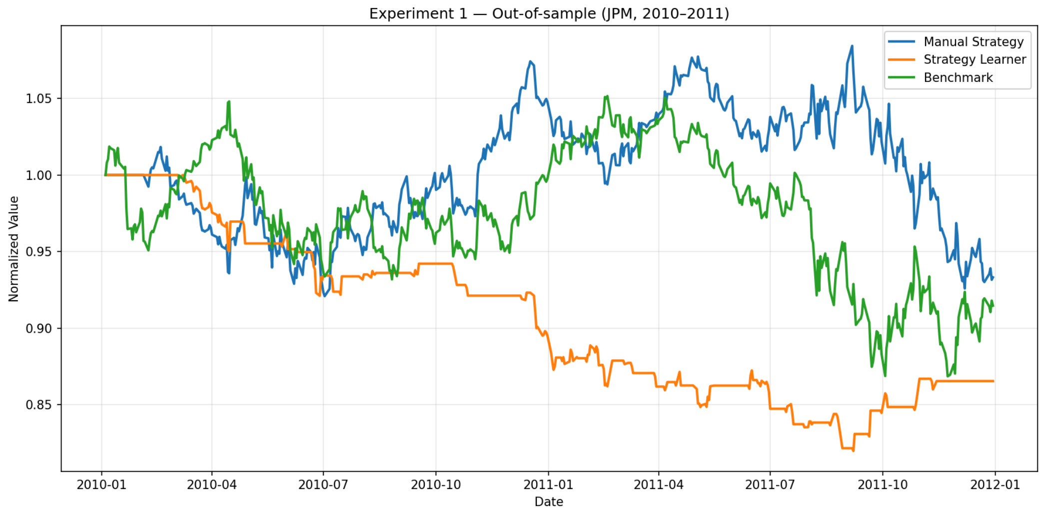
To train the Strategy Learner, I represent the market as a discrete state built from the same indicators used in the Manual Strategy. Each day I compute a trend signal (fast vs. slow WMA), Bollinger Band BBP, RSI, momentum, MACD, and the current position (short / neutral / long). Because Q-learning needs a finite state space, I discretize each continuous indicator using quantile-based bins from its in-sample distribution (very low RSI = low bin, mid RSI = middle bin, very high RSI = high bin). I apply the same bucketing to BBP, trend, momentum, and MACD so that extreme overbought/oversold or strong-trend conditions fall into outer bins and neutral conditions into middle bins. The final state is a tuple like (trend_bin, bbp_bin, rsi_bin, mom_bin, macd_bin, position), which is mapped to a single integer index in the Q-table, allowing the learner to treat different market regimes as distinct states and learn separate value estimates for going long, short, or to cash. I use 3 bins per indicator and train over multiple passes on the 2008–2009 in-sample period, grid-searching alpha in [0.1, 0.2, 0.3], gamma in [0.8, 0.9], and epsilon in [0.05, 0.10, 0.20] and choosing the combination that maximizes in-sample Sharpe and cumulative return. In the final configuration I use alpha = 0.2, gamma = 0.9, and an epsilon-greedy policy that decays from 0.10 toward 0.01, with the reward at each step equal to the daily portfolio return (change in portfolio value divided by the previous day's value), including transaction costs and impact. During testPolicy, learning is turned off. On each day I compute the current discrete state from the indicators and position, pick the action (long, short, or cash) with the highest Q-value, and convert that into a +1000, −1000, or 0 trade while respecting the holding constraints.

Experiment 1

Experiment 1 compares my rule-based Manual Strategy and the Q-learning StrategyLearner when they both trade JPM under identical assumptions: an initial portfolio value of $100,000, a $9.95 commission per trade, market impact of 0.005 applied symmetrically, and allowable positions of −1000, 0, or +1000 shares (with ±2000 share flips only when switching directly between long and short).



In-sample (2008–2009), both strategies comfortably beat the buy-and-hold benchmark. As shown in the table above, the Manual Strategy achieves the highest cumulative return and final portfolio value, but with higher daily volatility and a lower Sharpe ratio than the StrategyLearner. The learner earns a more modest cumulative return but delivers better risk-adjusted performance, reflecting its tendency to hold positions longer in regimes it has learned to favor.

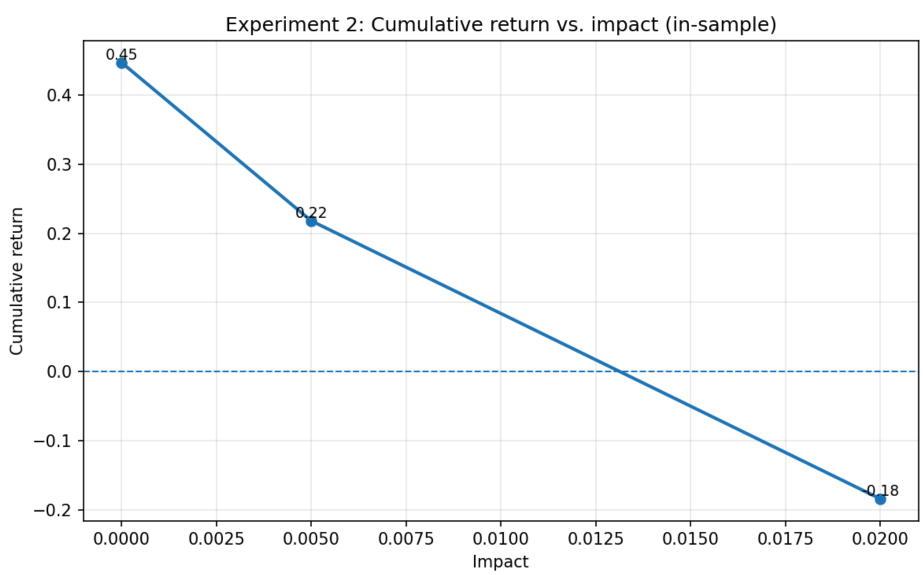


Out-of-sample (2010–2011), the environment shifts to a slightly downward, quieter market. The Manual Strategy again holds up better than both the benchmark and the StrategyLearner: its simple, trend-aware rules (WMA for direction, BBP and RSI for mean reversion, with momentum and MACD effectively downweighted) keep it mostly long or in cash and only briefly short, so it loses less than buy-and-hold. The StrategyLearner, by contrast, appears overfit to the 2008–2009 crisis, where extended shorting during large sell-offs was heavily rewarded. In the calmer 2010–2011 regime, the same indicator states no longer imply major crashes, but the learner still tends to stay too bearish; combined with transaction costs and noisy Q-values over a large discrete state space, this leads to a persistent short bias and a drifting downward equity curve. I would not expect the Manual Strategy to always beat the StrategyLearner in every in-sample period, since the Q-learner's performance is highly sensitive to the training window, exploration path, and how well the state space is covered.

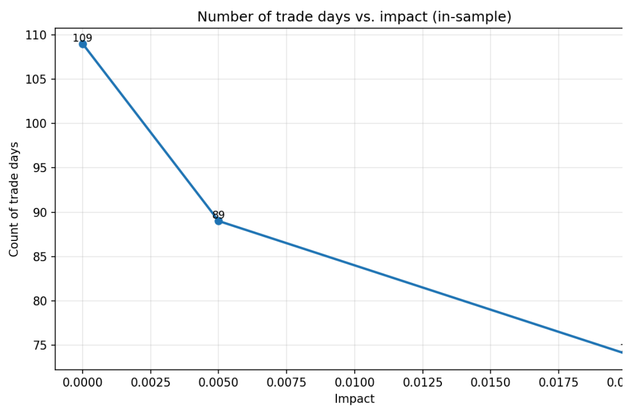
Cumulative Return	Average Daily Return	Std Daily Return	Sharpe Ratio	Final Value
0.388793	0.000731	0.012589	0.921540	138879.300000
0.222735	0.000411	0.004908	1.329319	122273.450000
0.010236	0.000165	0.017041	0.153387	100819.250000

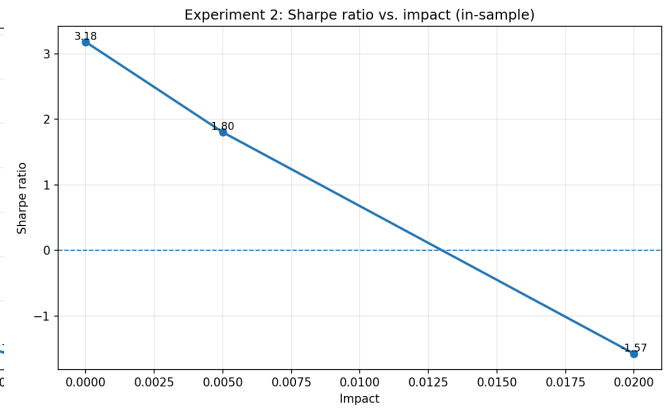
Experiment 2

I hypothesize that as market impact increases, the StrategyLearner will trade less often and deliver worse performance. Higher impact makes each trade more expensive, so the learner should become more selective about acting on signals, leading to fewer trade days and smaller position changes. At the same time, the higher effective costs will erode profits on the trades it does make, causing cumulative return to decline and potentially turn negative at high impact. Because returns fall faster than risk, the Sharpe ratio should also drop as impact increases, reflecting weaker risk adjusted performance in more expensive trading environments. From the experiment, we see a clear degradation in performance as impact rises from 0.0 to 0.005 to 0.02. Cumulative return drops from about 0.45 to 0.22 and then to −0.18, and the final portfolio value falls from roughly $144k to $122k to $82k. The Sharpe ratio declines from around 3.18 to 1.80 and finally to −1.57, while average daily return moves from a small positive to a smaller positive and then to a negative value. Daily volatility (std of daily returns) stays in a similar range, and the number of trades falls from 109 to 89 to 74, showing the learner trades less as costs rise. This behavior is expected as higher impact means each trade executes at a worse effective price, so trading costs eat into profits, shrinking both raw and risk-adjusted returns.



At high enough impact, the cumulative drag from these costs overwhelms the learner's edge, turning a previously profitable policy into a losing one, and the learner responds by trading less frequently and being more conservative about flipping positions.





impact	n_trades	cum_return	sharpe	avg_daily_return	std_daily_return	final_value
0.000000	109	0.446600	3.181676	0.000740	0.003690	144660.000000
0.005000	89	0.217958	1.804047	0.000397	0.003496	121795.750000
0.020000	74	-0.183606	-1.574591	-0.000395	0.003978	81639.400000