Constructing and Backtesting Trading Strategies

Analysis and Strategic Trading of the SPTL ETF

Anantajit Raja

March 2024

1 Introduction

This paper aims to craft and rigorously evaluate leveraged trading strategies tailored to the SPDR Portfolio Long Term Treasury exchange-traded fund (SPTL ETF) over the period from January 1, 2014, to December 31, 2019. The investigation is structured around a comprehensive framework that includes the preparation of time series data for the SPTL ETF and the Effective Fed Funds Rate (EFFR) as a proxy for the risk-free rate [3], the conception of three distinct leveraged trading strategies, and the application of several performance indicators to assess these strategies' efficacy and robustness.

ETFs have gained significant attention from both academic researchers and practitioners for their flexibility, efficiency, and the diversification benefits they offer to investment portfolios. Among the plethora of ETFs available in the market, the SPDR Portfolio Long Term Treasury ETF (SPTL) presents an intriguing opportunity for investors interested in the fixed income domain, particularly in U.S. Treasury bonds. The period from January 1, 2014, to December 31, 2019, offers a compelling study window due to its diverse economic cycles, interest rate environments, and significant geopolitical events that have the potential to impact the performance of long-term treasury instruments.

2 Methodology

We download our daily data for SPTL ETF adjusted closing prices from the Yahoo! Finance website [6], specifying the start and end dates we're interested in (01/01/14 and 31/12/19). Now note that the 'risk-free rate' represents the return on an investment with zero risk - we use as the risk-free rate the Effective Federal Funds rate (EFFR), whose daily data is provided by the Federal Reserve Bank of New York on the newyorkfed website for the same dates [2].

2.1 Data Cleaning, Analysis and Visualisation

We begin by using the pandas Python package to read in our data, filter it (e.g. for 'Adj Close' or 'Effective Rate' columns) and merge our dataframes for prices and rates into a single dataframe (where we kept the rows with the common dates). We see if our data has any missing values (though there were none to begin with) and drop any duplicate rows (of which we found were 55 that were duplicates).

Now to effectively construct trading strategies we must understand the impact of the EFFR. The EFFR is widely considered a benchmark for the overnight unsecured lending rate for banks in the United States. It's often used as the risk-free rate in financial models due to its stability and the backing of the U.S. government. The annual risk-free rate is adjusted to a daily rate to align with the daily analysis of financial instruments:

Definition 1. *(Daily risk-free rate) The daily risk-free rate can be calculated as* $r_t^f = EFFR(t) \cdot dc$ *where* $EFFR(t)$ *is our measurement of the annual risk-free rate and the day-count* dc *is approximately 1/252 (due to there being roughly 252 business days in a year).*

To get an idea of how the ETF price and the daily risk-free rate changes with time we produce plots (using matplotlib) seen in **Figure 1** and descriptive statistics (using pandas' describe function). We see from the ETF plot that prices trended upward initially, then there was a 4 year period where prices remained around the same level before trending upward again in the final year. The average price is around \$30 with moderate fluctuations over time, ranging from around \$23 to \$38. As for daily rates we see there has been a steady increase from 2016 which then down-turned from early 2019. The average daily risk-free rate is very low at around 0.0037%, with slight variations, remaining under 0.01%.

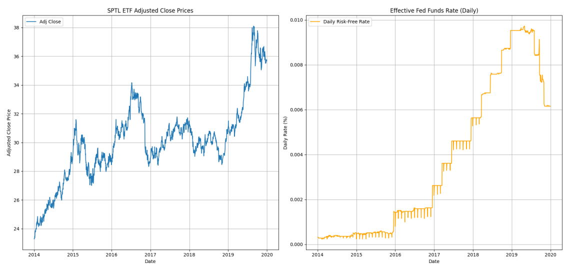


Figure 1: Plots of SPTL ETF adjusted closing prices (in dollars) and Daily EFFR (%) from 01/01/14 to 31/12/19

	Adj Close (\$)	Daily Rate (%)
count	1443	1443
mean	30.076009	0.003654
std	2.781817	0.003306
min	23.283545	0.000238
25%	28.683072	0.000516
50%	30.069080	0.001627
75%	31.190990	0.006726
max	38.106731	0.009722

Figure 2: Summary statistics of Adj Close (\$) and Daily Rate (%) from 01/01/14 to 31/12/19

When dealing with investments like ETFs, understanding and calculating the daily excess return is crucial to determine the profitability of our investment:

Definition 2. *(Daily excess return) The daily excess return per unit of an ETF reads* $r_t^e = \frac{\Delta p_t}{p_t} - r_t^f$ *where we define* $\Delta p_t / p_t$ *to be the gross linear return and* r_t^f *the daily risk-free rate in decimal form.*

Calculating the daily excess return allows investors to understand the performance of an investment after adjusting for its risk, as represented by the risk-free rate. This provides a clearer picture of the investment's value proposition. By using such a standardised measure, investors can compare the performance of different investments on a level playing field, taking both return and risk into account. In **Figure 3** we produce plots for the SPTL linear return time series, the daily EFFR, and the excess return per unit of SPTL over the period of concern.

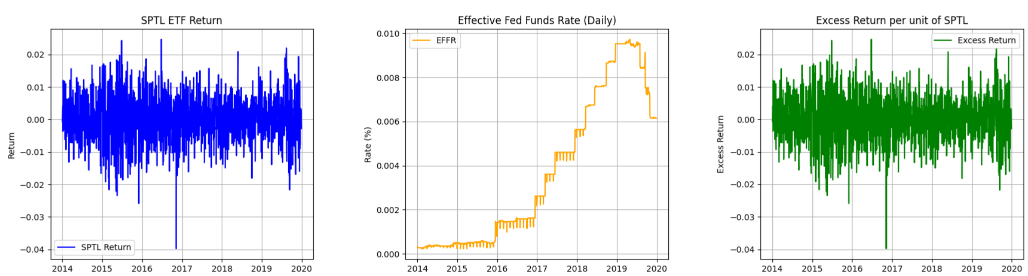


Figure 3: Plots of SPTL ETF returns, Daily EFFR (%) and excess return per unit of SPTL from 01/01/14 to 31/12/19

Note that there's barely any visual difference between the return and excess return plots since the daily rate, when converted to a decimal, becomes very small (since we divide it by 100 to get the decimal) and therefore has a negligible impact on the return when subtracted from it. In any case, we can see that that returns are mean-reverting and and they oscillate about the zero level, which is an indication of stationarity (can be formally confirmed using a unit root test, but omitted from this paper for brevity) [4].

2.2 Defining our trading strategies

We now transition to a critical aspect of our analysis: the formulation and assessment of leveraged trading strategies. Leveraged strategies inherently involve amplifying potential returns on investment by using borrowed funds, thereby also increasing the risk associated with the investment. This section aims to meticulously outline the development, rationale, and implementation of three distinct leveraged trading strategies tailored for the SPTL ETF, taking into consideration the unique market dynamics and risk-return profiles observed from January 1, 2014, to December 31, 2019.

Leverage is applied as a multiplier to the initial investment capital, which we define to be $V_0 = \$200000$, determining the maximum dollar value of SPTL that can be either long (bought) or short (sold). The parameter L signifies the leverage multiplier, dictating the aggressiveness of the strategy - we set $L = 10$. Formally, for each trading period t, the absolute dollar value invested in SPTL, θ_t, is constrained by the product of the initial capital and the leverage, expressed as $|\theta_t| \leq V_0 \cdot L$. These bounds directly control the maximum exposure to the SPTL, helping to prevent excessively risky positions that could lead to significant losses. It also helps mitigate the impact of sudden market movements, ensuring that the investment can withstand short-term fluctuations without incurring catastrophic losses.

We now define three long-only strategies that satisfy the above constraints. We use a train-test split of 70:30 (this split seems to be used most often in literature) where the training set is used to tune pre-defined parameters of the model, and we apply the model to the test set to evaluate the performance of the model on unseen data.

Strategy 1: A long-only medium-term trend-following strategy

From our analysis of the SPTL ETF price plot we saw that it is generally upward trending for certain extended periods. This suggests that a strategy which aims to capitalise on the price momentum by identifying optimal entry and exit points based on the moving average of the adjusted closing prices will hopefully work well [1].

Definition 3. *(Moving average) The moving average of a price time series with time window τ is the average of the current value and its τ - 1 preceding values. We let P(t) be the price at time t and τ our time window. The moving average at time t is computed as* $MAv\ (t) = \frac{1}{\tau}\sum_{i=t-\tau+1}^{t} P(i)$.

The strategy works as follows (we detail the reason for our choice of window later):

1. We start by using \$130000 of our initial capital as our initial cash amount for possible investment in the ETF (leveraged 10 times so it becomes \$1.3 million). Note this sufficiently low initial cash amount allows us to obey the theta (position value) constraint imposed earlier.

2. 75-Day-Window Moving Average Calculation: For each point in the dataset, a cumulative sum of the ETF's adjusted closing prices is maintained, enabling the efficient computation of the moving average once sufficient data points (equivalent to the time window) are available.

3. Position Adjustments: The strategy dynamically adjusts its position in the ETF based on the relationship between the moving average and the current adjusted closing price every day:

   - If the current price is equal to the moving average, the position is held constant.
   - If the current price exceeds the moving average, indicating upward momentum, the strategy converts all available cash into ETF shares, fully investing in the expectation of continued price appreciation.
   - If the current price is below the moving average, suggesting downward momentum, the strategy liquidates its ETF holdings, converting the entire position back to cash to mitigate potential losses.

Note that to obtain this strategy we used validation on the training set. A key aspect of this strategy involves experimenting with various time windows to determine the optimal period over which to calculate the moving average. Time windows range from 10 to 90 days, incremented by 5-day intervals, allowing for a thorough exploration of short-to-medium-term price trends. We find that a time window of length 75 maximises return over the period for the training set (see **Figure 4**), and we applied this same window for the test set as detailed above.

We also keep track of the amount borrowed at each stage, subtracting the total cost of leveraging at each time step to get a more accurate return evaluation for our train and test sets. This cost is calculated by multiplying the daily risk free rate with the amount borrowed.

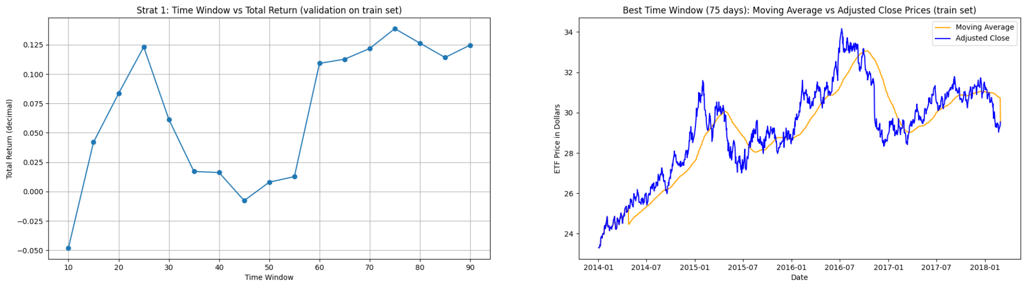


Figure 4: Plots of Strategy 1 Returns vs Time Window (left) and SPTL ETF price plotted with the best moving average for the training set (right)

Note that for an n-day moving average we need n data points, which is why the moving average curves start on the n^{th} day, and hence no trades are executed until at least the 75^{th} day for our strategy (which is acceptable given the size of our dataset is of order roughly ten times this).

Strategy 2: A long-only short-term mean-reversion strategy

Applying a mean reversion strategy to the SPTL ETF is also well-founded given some of the characteristics of long-term Treasury bonds, such as cyclical price movements driven by changes in interest rates and economic conditions. U.S. Treasury securities are known for their safety and relatively predictable interest rate sensitivity, which makes the ETF a suitable candidate for a strategy that expects prices to revert to their historical average over time.

The strategy works as follows (again, detailing the reason for choice of window later):

1. We start by using $130000 of our initial capital as our initial cash amount for possible investment in the ETF (leveraged 10 times so it becomes $1.3 million). As before, this sufficiently low initial cash amount allows us to obey the theta constraint imposed earlier.

2. 10-Day-Window Moving Average Calculation: For each point in the dataset, we calculate the moving average as before (but with the previous 10 data points instead) using the ETF's adjusted closing prices given that sufficient data points are available to do so.

3. Position Adjustments: The strategy dynamically adjusts its position in the ETF based on the relationship between the moving average and the current adjusted closing price every day:

   - If the current price is equal to the moving average, the strategy maintains its current allocation between cash and ETF shares, interpreting the price alignment as a sign of market equilibrium.
   - If the current price exceeds the moving average, this indicates a possibly overrvalued ETF which has deviated from its typical trading range, and the strategy sees this as a signal to sell. It liquidates its ETF holdings, converting the entire position to cash to capitalise on the current high prices before the expected reversion.
   - If the current price is below the moving average, this indicates a potentially undervalued ETF which has deviated from its typical trading range, and the strategy interprets this as a buying opportunity. It uses available cash to purchase ETF shares, betting on the price's eventual rebound to its mean.

Similar to before the optimal length of the moving average window was found using validation on the training set, though this time the window which produced the best return proved to be 10 days long (see **Figure 5**) - much shorter than for the trend-following strategy. This same window length is then used for the strategy on the test set as above.

Leveraging costs have also been factored in by multiplying the daily risk-free rate by the amount borrowed at each time step.

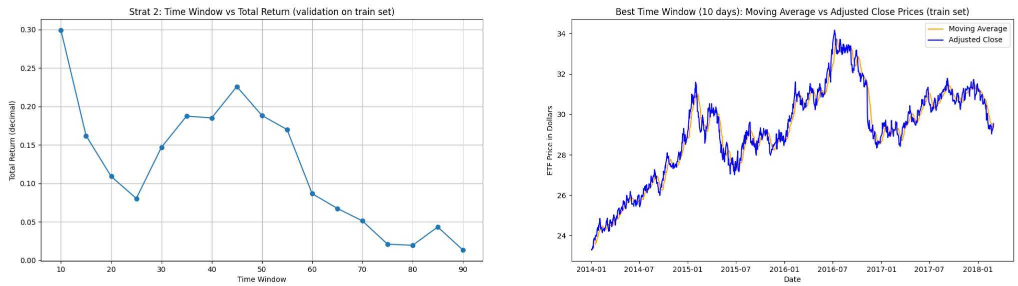


Figure 5: Plots of Strategy 2 Returns vs Time Window (left) and SPTL ETF price plotted with the best moving average for the training set (right)

Strategy 3: A long-only AR(1) strategy

We do some further statistical analyses on the training set to show that an autoregressive strategy can potentially be applied, and to tune the lag parameter for this strategy.

Definition 4. *(Autocorrelation) Autocorrelation is a measure of the correlation between a time series and its own past values over successive time intervals. Let X_t be a time series. The autocorrelation is the normalised sum:* $R(k) = \frac{\sum_{t=k+1}^{T}(X_t - \bar{X})(X_{t-k} - \bar{X})}{\sum_{t=1}^{T}(X_t - \bar{X})^2}$ *where T is the number of observations of the time series.*

Definition 5. *(Partial Autocorrelation) Partial autocorrelation quantifies the direct correlation between an observation in a time series X_t and another observation at a specific lag k away, X_{t+k} (eliminating the influence of the observations in between). The function at lag k is defined as:* $\phi(1) = corr(X_{t+1}, X_t)$ *for* $k = 1$, $\phi(k) = corr(X_t - \hat{X}_t, X_{t+k} - \hat{X}_{t+k})$ *for* $k \geq 2$ *where $\hat{X}_t$ and $\hat{X}_{t+k}$ are linear combinations of the smaller lags that minimise the mean squared error with respect to X_t and X_{t+k}*

We plot the SPTL ETF price autocorrelation and partial autocorrelation functions (ACF and PACF) in **Figure 6** using the training set. These plots are used to identify whether the price process has long memory (by examining the rate of decrease of ACF values), and if so identifies the lag parameter of the autoregressive process in question (by seeing for which lag there's a sharp drop in the PACF value).

The ACF plot for the ETF prices shows a gradual decline in autocorrelation as the lag increases, with the autocorrelation not reaching zero even at large lags of more than 50. This suggests that the ETF prices are autocorrelated over a long period, indicating a "long memory" effect where past prices have a prolonged influence on future prices. For the PACF plot there's a sharp drop in correlation after the first lag. This is a classic characteristic of an AR(1) process, indicating that the current price is primarily influenced by its immediately preceding value, with negligible influence from further past values.

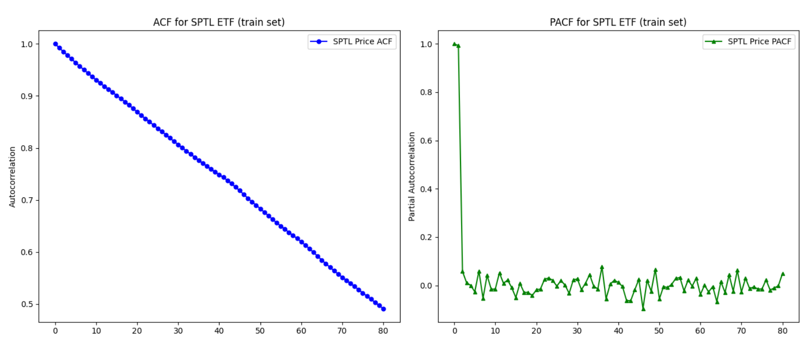


Figure 6: Plots of SPTL ETF price autocorrelation and partial autocorrelation functions for the first 7/10th of the data (the training set)

For an Autoregressive (AR(1)) trading strategy applied to the SPDR Portfolio Long Term Treasury ETF (SPTL), the decision-making process involves using predictions from an autoregressive model to guide trading actions. This strategy specifically considers the relationship between the current price and its predicted next value based on historical data.

Our strategy therefore works as follows:

1. We start by using \$130000 of our initial capital as our initial cash amount for possible investment in the ETF (leveraged 10 times so it becomes \$1.3 million). As before, this sufficiently low initial cash amount allows us to obey the theta constraint imposed earlier.

2. We fit an AR(1) model at each time step given information up until time t using the ETF's adjusted closing prices given that sufficient data points are available to do so (note that even though the test set is meant to be unseen, at each point in time in the test set we have information until that point in time so we can keep fitting a new AR model based on past data).

3. Position Adjustments: The strategy dynamically adjusts its position in the ETF based on the relationship between the predicted price and the current adjusted closing price every day:

   - If the predicted price is equal to the current price, the strategy does not alter its position, maintaining the existing allocation of ETF shares and cash. The model anticipates no significant change in price that could be exploited for profit, suggesting a hold decision to keep the portfolio's current structure.
   - If the predicted price is higher than the current price, indicating an expected price increase, the strategy converts all available cash into ETF shares to fully invest in anticipation of the price appreciation.
   - If the predicted price is lower than the current price, suggesting an expected decline, the strategy liquidates its ETF holdings, converting the entire position back to cash. This aims to lock in gains or minimize losses before the anticipated decrease in price.

At each point in time we keep track of the amount borrowed in order to account for leverage costs, using amount borrowed multiplied by the daily risk-free rate.

2.3 Analysis of trading strategies

We can now quantify the performance of these strategies through daily trading Profit and Loss (PnL). This assessment will hinge on the concept of excess return, which captures the profit or loss generated by each strategy beyond the risk-free rate of return.

Definition 6. *(Daily Trading PnL) The daily trading PnL ΔV_t is formally calculated as the difference in ETF price changes adjusted for the risk-free rate, multiplied by the dollar value of SPTL held at time t:* $\Delta V_t = (\frac{\Delta p_t}{p_t} - r_t^f)\theta_t$ *where θ_t represents the investment in SPTL at its current price p_t.*

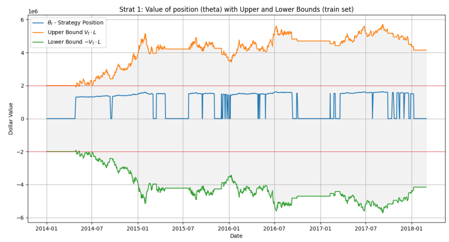
And so to delve deeper into the dynamics of these strategies we plot the position values (θ_t) alongside the designated leverage boundaries ($\pm\ V_t \cdot L$), providing a visual encapsulation of each strategy's adherence to its leverage constraints over time for both the train and the test sets - see **Figure 7**.

Note that all three strategies obey the imposed investment limits, i.e. for all time θ_t stays within bounds of $V_0 \cdot L$, which in our case is 2 million dollars. The reliance on a 75-day moving average for the first strategy means it reacts to longer-term market trends. The theta for all three strategies for all time is positive because our strategies are long-only (though they can involve selling only shares of SPTL that we already own).

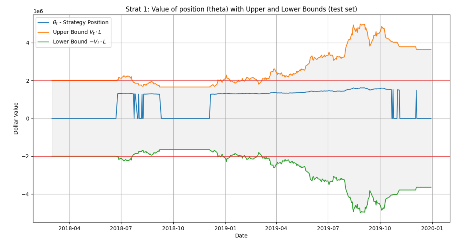
The infrequent changes in θ suggests that it's less sensitive to short-term market fluctuations, which could reduce transaction costs and potentially unnecessary exposure to brief market volatilities. However, there are points in time where significant adjustments occur, indicating a substantial reallocation of assets.

For the second strategy, utilising a 10-day moving average makes it more responsive to short-term market movements. There's increased frequency of significant θ adjustments which aims to capitalise on more immediate market opportunities. The presence of large jumps, similar to Strategy 1, indicates that despite its focus on the short term, the strategy does not shy away from making substantial position changes when the indicator suggests a strong market movement.

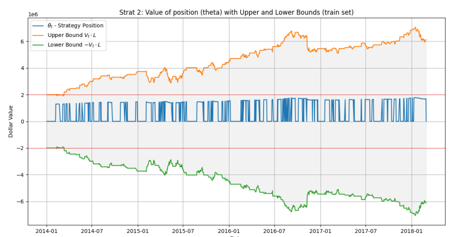
For the third strategy the use of the AR(1) model results in the most frequent significant adjustments of θ, a highly dynamic approach to portfolio management. This means the strategy is well-suited for capturing profits in volatile or rapidly changing markets, but it also bears the highest risk of significant losses and elevated transaction costs due to the frequent trading.



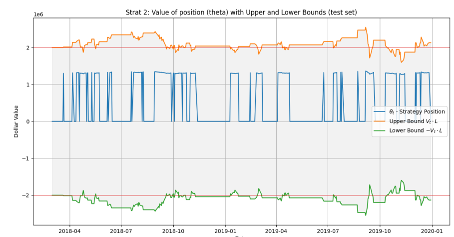
(a) Strat 1 train theta vs bounds



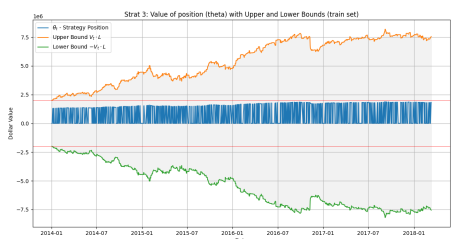
(b) Strat 1 test theta vs bounds



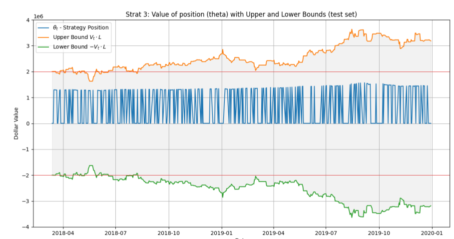
(c) Strat 2 train theta vs bounds



(d) Strat 2 test theta vs bounds



(e) Strat 3 train theta vs bounds



(f) Strat 3 test theta vs bounds

Figure 7: Strategies 1, 2, and 3 plots of the position value theta vs upper and lower leveraged Vt bounds

We now quantify the turnover in dollar value which reflects the aggregate amount of capital reallocated within the portfolio across the entire trading period. Mathematically, this is captured by:

Definition 7. *(Dollar value turnover)* $Turnover_{dollars} = \sum_{i=0}^{T} |\Delta\theta_t|$

This offers a direct measure of the strategy's market engagement and potential exposure to transaction costs.

We can also define the turnover in the number of units traded, which is computed as the sum of absolute changes in the portfolio's ETF holdings:

Definition 8. *(Unit turnover)* $Turnover_{units} = \sum_{i=0}^{T} |\frac{\theta_{t+1}}{p_{t+1}} - \frac{\theta_t}{p_t}|$

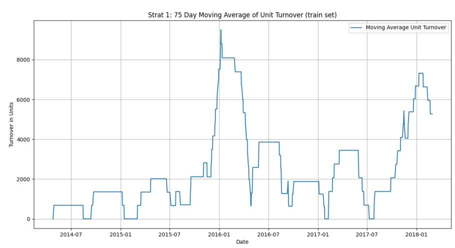
Unlike the dollar-value turnover, this metric offers insight into the physical volume of securities transacted, highlighting the strategy's liquidity demands and its potential influence on market prices. See **Table 1**.

We plot a 75-day moving average of the unit turnover for all 3 strategies for both the train and the test set to capture the general turnover trend - see **Figure 8** (note that the general shape of the turnover is the same for both metrics in dollar and units, so we just pick the unit turnover for the moving average plot).

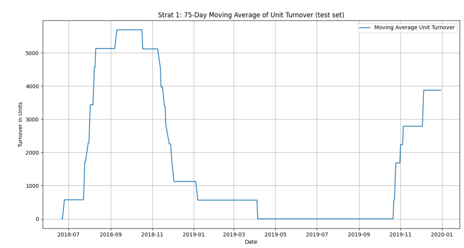
It seems the turnover is higher in certain periods compared to other periods, but this is dependent on the strategy in question. When there are clear market trends, either upward or downward, our first strategy shows more frequent trading to capitalise on these trends, namely a higher turnover. In highly volatile markets, price swings can be more pronounced and this causes signals for our first strategy to be triggered more frequently. This volatility can lead to a sequence of buy and sell signals as prices move sharply, increasing turnover.

Our mean reversion strategy (strat 2) tends to see more trading activity, i.e. higher turnover, during volatile periods because prices are more likely to deviate significantly from their means. Frequent crossing of the SPTL mean price level also leads to higher turnover in those periods (which explains why for the test set it's not as high for the up-trending period).

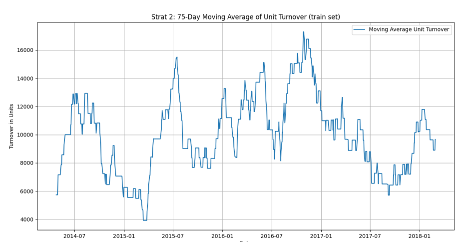
For the AR(1) strategy (strat 3) when SPTL experiences periods of high volatility, the price can diverge more significantly from the predicted AR(1) model value. This might lead to increased turnover when trading as the strategy attempts to capitalise on what it identifies as temporary mispricings.



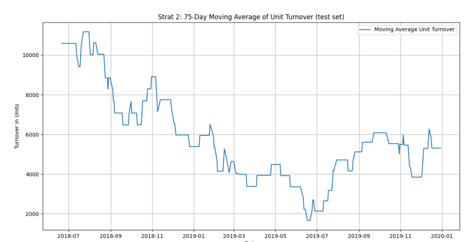
(a) Strat 1 train 75 day mavg unit turnover



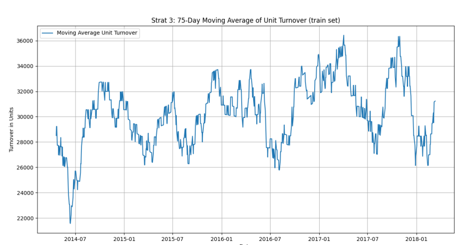
(b) Strat 1 test 75 day mavg unit turnover



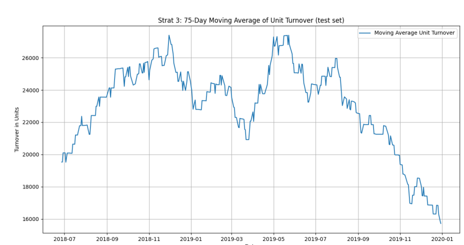
(c) Strat 2 train 75 day mavg unit turnover



(d) Strat 2 test 75 day mavg unit turnover



(e) Strat 3 train 75 day mavg unit turnover

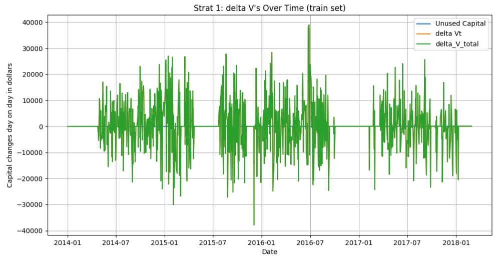


(f) Strat 3 test 75 day mavg unit turnover

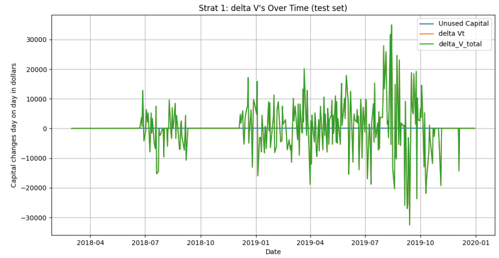
Figure 8: Strategies 1, 2, and 3 plots of the 75 day moving average unit turnover for the train and test sets

We will now develop a comprehensive framework for evaluating the profitability of our trading strategies over time. We aim to construct a total PnL series that reflects not only the direct trading gains and losses (ΔV_t) but also the opportunity costs associated with uninvested capital. We place this capital in a money-market account which accrues interest at the prevailing risk-free rate, contributing an additional factor to the account's value (ΔV_t^{cap}). The total value of the trading account (V_{total}) is therefore influenced by both the trading P&L and the growth of the money-market capital account, with the equation $V_{t+1}^{total} - V_t^{total} = \Delta V_t^{total} = \Delta V_t + \Delta V_t^{cap} = (\frac{\Delta p_t}{p_t} - r_t^f)\theta_t + (V_t^{total} - M_t)r_t^f$ where $M_t = \frac{|\theta_t|}{L}$ = margin used; $V_0 = \$200K$.

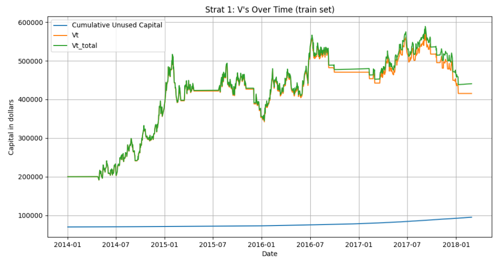
Figure 9: Strategies 1, 2, and 3 plots of day-on-day change in capital and accumulated capital over time for both training and test sets



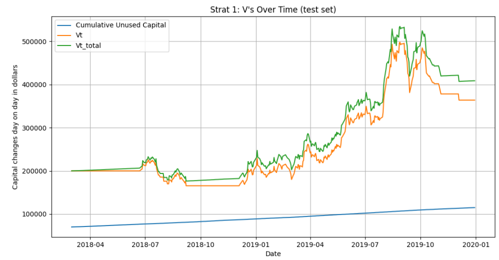
(a) Strat 1 train deltaV's



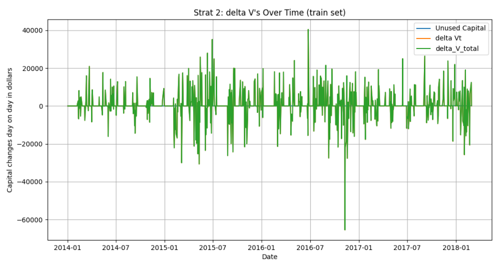
(b) Strat 1 test deltaV's



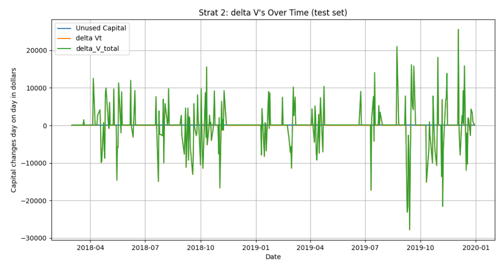
(c) Strat 1 train V's



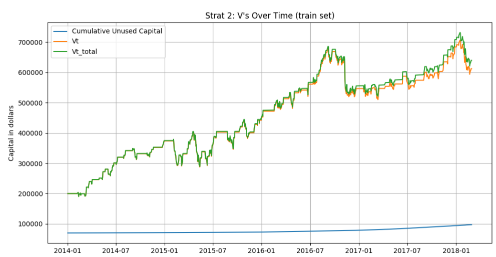
(d) Strat 1 test V's



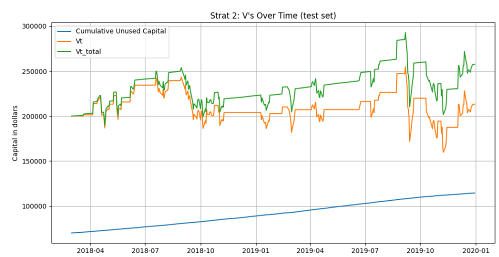
(e) Strat 2 train deltaV's



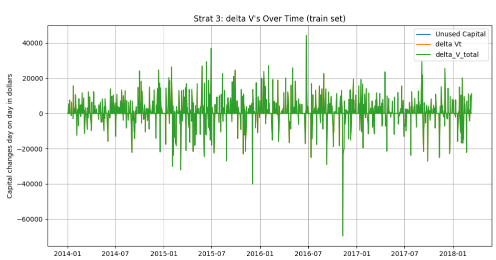
(f) Strat 2 test deltaV's



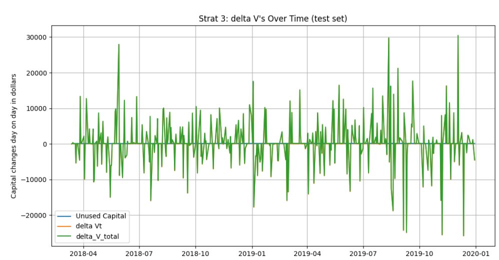
(g) Strat 2 train V's



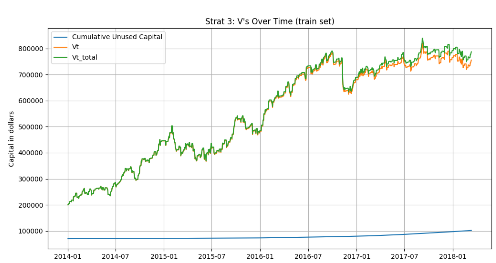
(h) Strat 2 test V's



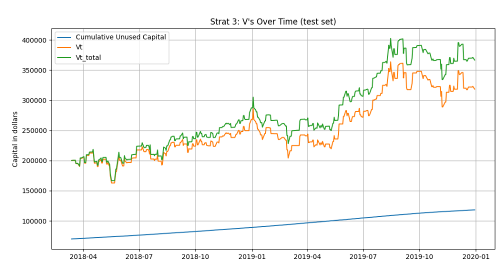
(i) Strat 3 train deltaV's



(j) Strat 3 test deltaV's



(k) Strat 3 train V's



(l) Strat 3 test V's

We plot the ΔV's over time and their accumulated values (V's) to get an idea of the contribution of each to the total capital - see **Figure 9** above.

From the plots we can see that from an initial capital allocation of \$200,000, from which \$130,000 is actively employed cash within our trading strategies, the remaining \$70,000 represents the initial value of the unutilised capital, which we allocate to a money-market account.

The ΔV_t graphs present a vivid depiction of the immediate financial consequences of our trading decisions, providing a granular view of performance over time. Though through the delta V graphs we can't see much of an impact of the money market account on day on day total capital changes, this is much more prevalent in the accumulated capital graphs where at the end of the investment period for both training and test sets we can see a positive sizeable difference in the final total capital for all three strategies.

Note that higher funding costs would mean higher interest payments on the borrowed funds used to maintain these leveraged positions. This would directly reduce the profitability of the trading strategies, as a greater portion of any gains would be used to cover the increased interest expenses. If funding costs increased by 150% then since the order of our position value in SPTL is around 10 times as much as the amount in our money market account, the losses from the SPTL position will massively outweigh (around 10 times as much) the gain from our money market account. This cost will therefore be around 1.5 times the previous borrowing cost (though slightly below due to the gain from the money market). Our strategies will still be profitable, but they won't be as profitable as previously mentioned.

3 Results

We now carry out a comprehensive performance evaluation of our trading strategies, employing a suite of well-regarded financial metrics. Our objective is to not only quantify the success and efficiency of our strategies but also to understand their behaviour under different market conditions.

We begin by defining the excess return of a trading strategy as its daily trading PnL (ΔV_t). Using this fundamental metric, we proceed to calculate and analyse the Sharpe Ratio (SR), Sortino Ratio, Maximum Drawdown, and Calmar Ratio for our strategies. These calculations are performed independently for both the training set and the test set, allowing us to scrutinise the strategies' performance during the periods of model fitting and subsequent validation [5].

Definition 9. *(Sharpe Ratio) The Sharpe Ratio (SR) is defined as the average excess return earned above the risk-free rate per unit of volatility or total risk. It is mathematically represented as* $SR = \frac{E[R_p - R_f]}{\sigma_p}$ *where* R_p *is the portfolio return,* R_f *is the risk-free rate,* $E[R_p - R_f]$ *is the expected excess return of the portfolio over the risk-free rate, and* σ_p *is the standard deviation of the portfolio excess return.*

Definition 10. *(Sortino Ratio) The Sortino Ratio enhances the Sharpe Ratio by considering only downside risk instead of total volatility. It is defined as* $Sortino\ Ratio = \frac{E[R_d - R_f]}{\sigma_d}$ *where* σ_d *represents the standard deviation of the portfolio's negative returns, focusing on the volatility of downside returns only.*

Definition 11. *(Maximum Drawdown) Maximum Drawdown (MDD) measures the largest peak-to-trough decline in the value of the portfolio, without considering the time frame. It is calculated as* $MDD = \max_{\tau \in (t,T)} \left(\max_{t \in (0,\tau)} P_t - P_\tau \right)$ *where* P_t *is the portfolio value at time* t*, and* T *is the total period observed. Note that the expression in the brackets can be used to obtain a drawdown chart.*

Definition 12. *(Calmar Ratio) The Calmar Ratio relates the annualised return of a portfolio to its maximum drawdown over a specific period. It is defined as* $Calmar\ Ratio = \frac{Annualized\ Return}{MDD}$*. This ratio offers an insight into the return earned per unit of the risk taken, as measured by the maximum drawdown.*

We get the following results for our 3 strategies:

Strategy 3 consistently shows the highest Sharpe Ratio in both training and testing sets, indicating a better risk-adjusted return compared to Strategies 1 and 2. Strategy 1 has the lowest Sharpe Ratio in the training set but shows a significant improvement in the test set. We can also plot a rolling Sharpe Ratio to visualise how it varies with time. See **Figure 10**.

Similar to the Sharpe Ratio but only considering downside risk, Strategy 3 again leads with the highest Sortino Ratio in both sets, suggesting it has the most favorable return considering only the downside volatility.

Metric	Strategy 1		Strategy 2		Strategy 3	
	Train	Test	Train	Test	Train	Test
Sharpe Ratio	0.465	1.11	0.805	0.916	0.874	1.12
Sortino Ratio	0.608	1.31	0.762	0.870	0.845	1.19
Maximum Drawdown	0.337	0.269	0.248	0.374	0.202	0.285
Calmar Ratio	0.597	1.58	0.909	1.30	1.15	2.19
Total turnover in dollars	77.1 million	26.5 million	292 million	85.3 million	930 million	320 million
Total turnover in units	2.41 million	760 thousand	9.82 million	2.66 million	31.8 million	10.0 million

Table 1: Performance metrics for strategies across training and testing sets.

Strategy 3 has the lowest drawdowns in both the training and testing sets, suggesting that it has the least amount of risk in terms of potential loss. Strategy 2 experiences a higher drawdown in the test set compared to the training set, indicating possible overfitting or a strategy that didn't perform as well out-of-sample.

Strategy 3 has the highest Calmar Ratio in the test set, indicating a better performance given the level of risk taken. Strategy 1 has the lowest Calmar Ratio in the training set but this improves significantly in the test set.

4 Discussion

All strategies show an increase in Sharpe Ratio, Sortino Ratio and Calmar Ratio from training to testing. This is probably due to either overfitting to specific market conditions (such as trending or mean-reverting) that did not persist, or due to sample bias where training set could have been drawn from a period of adverse market conditions while the test set experienced more benign conditions. To bring the results of the training and test sets closer and ensure more robust strategies we could analyse performance across different economic regimes to ensure the strategy can handle various environments. Building on this, we could have used techniques such as walk-forward analysis or k-fold cross-validation to ensure the strategy is tested on multiple out-of-sample subsets.

The rolling sharpe ratio isn't consistently high for any of the strategies (see **Figure 10**), so they aren't particularly robust when it comes to adapting to changing market conditions - we can change this in future by imposing stricter risk management measures e.g. using a stop-loss condition.

Note the higher turnovers for strategy 2 and 3 are due to the frequency of trading: strategy 1 only uses a 75 day moving average whereas strategy 2 uses a 10-day moving average for the mean reversion and strategy 3 uses just the previous lag, resulting in several more points of trade for the latter.

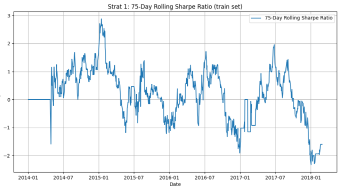
Note that we can analyse the drawdowns in more detail by looking at the drawdown chart for each of the strategies and plotting this against the rolling 90-day volatility of the SPTL price (measured by standard deviation of returns). See **Figure 10**.

Note that the drawdown chart roughly follows the shape of the 90-day rolling volatility chart for most of our graphs. Some of the biggest drawdowns occur when volatility is very high - significant price swings can lead to larger and more frequent drawdowns. Note also that the use of leverage for all of our strategies means amplification of both gains and losses. These amplified losses result in deeper drawdowns too - in times of high volatility liquidity can "dry up" making it difficult to exit positions at favourable prices.

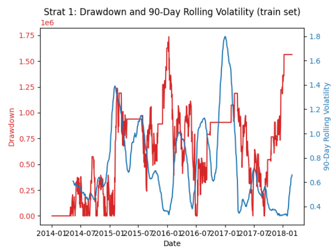
Some safer strategies could involve tailoring margin use based on underlying market volatility which can lead to safer strategies, managing risk more effectively:

- when volatility is high, reducing margin can limit exposure to large price swings, helping to control the size of potential drawdowns
- when volatility is low and price movements are more predictable, a strategy might safely increase margin to leverage smaller price movements for greater profit

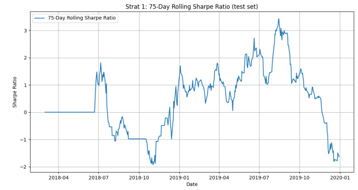
Alternatively we could also use volatility based position-sizing or volatility-adjusted stop losses to take advantage of stable market movements while maintaining positions during volatile periods without excessive risk.



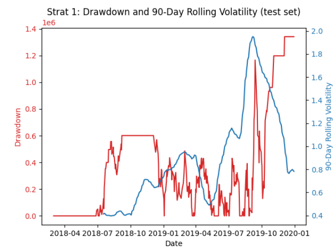
(a) Strat 1 train rolling Sharpe (75 day)



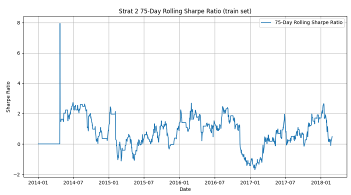
(b) Strat 1 train drawdown vs rolling vol (90 day)



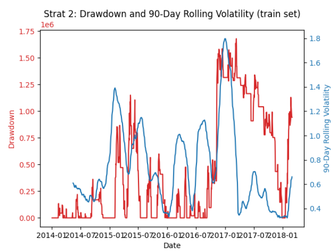
(c) Strat 1 test rolling Sharpe (75 day)



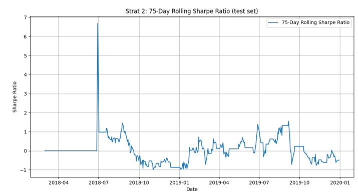
(d) Strat 1 test drawdown vs rolling vol (90 day)



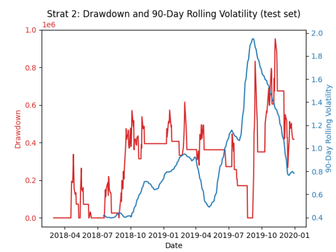
(e) Strat 2 train rolling Sharpe (75 day)



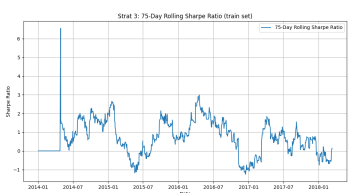
(f) Strat 2 train drawdown vs rolling vol (90 day)



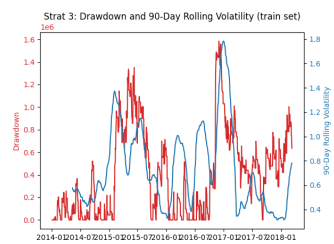
(g) Strat 2 test rolling Sharpe (75 day)



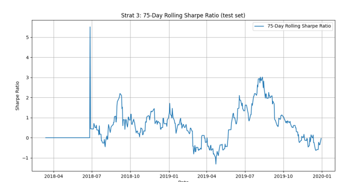
(h) Strat 2 test drawdown vs rolling vol (90 day)



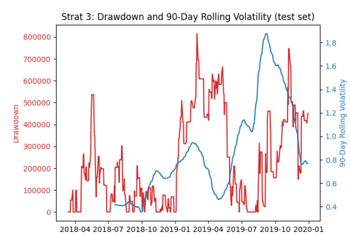
(i) Strat 3 train rolling Sharpe (75 day)



(j) Strat 3 train drawdown vs rolling vol (90 day)



(k) Strat 3 test rolling Sharpe (75 day)



(l) Strat 3 test drawdown vs rolling vol (90 day)

Figure 10: Strategies 1, 2, and 3: 75-day rolling Sharpe ratio, and drawdown vs. 90-day rolling volatility for both train and test sets.

References

[1] Craig A. Ellis and Simon A. Parbery. Is smarter better? a comparison of adaptive, and simple moving average trading strategies. *Research in International Business and Finance*, 19(3):399–411, 2005.

[2] Federal Reserve Bank of New York. Effective federal funds rate, accessed 2024-03-18, 2024.

[3] Karol Gellert and Erik Schloegl. Short rate dynamics: A fed funds and sofr perspective, 2021. FIRN Research Paper Forthcoming, Available at SSRN: https://ssrn.com/abstract=3763589 or http://dx.doi.org/10.2139/ssrn.3763589.

[4] Deniz Ilalan and Özgür Özel. Unit root testing in the presence of mean reverting jumps: Evidence from us t-bond yields. *International Journal of Nonlinear Sciences and Numerical Simulation*, 20(2):145–152, 2019.

[5] Jaydip Sen. A comparative study on the sharpe ratio, sortino ratio, and calmar ratio in portfolio optimization, 12 2022.

[6] Yahoo Finance. Sptl etf prices. https://finance.yahoo.com/quote/SPTL, 2024. Accessed: 2024-03-11.