Crypto Whales - Report



Yin Kiat Javier Tan Lakshya Agarwal Daniel Li

INTRODUCTION

Decentralised Finance (DeFi) has witnessed consistent growth, with Uniswap V3 standing as a crucial protocol in the ecosystem. In order to improve capital market efficiency, Uniswap V3 further introduced the concept of "Concentrated Liquidity", where LPs are able to set a upper and lower price range for their liquidity to be used. While this significantly improves capital market efficiency, it exacerbates the problem of impermanent loss defined as the divergence in value between providing liquidity and simply holding assets.

As such, Uniswap V3 provides LPs both the opportunity to reap greater annual percentage rates (APR) through greater utilisation of capital and potential for large impermanent losses. It is essential for LPs to be able to minimise their impermanent losses through hedging.

In this report, we will explore a comprehensive approach to addressing the challenges posed by applying advanced quantitative techniques and model impermanent loss with a hybrid ARIMA-GARCH model, predict future pool prices and provide a foundation for estimating impermanent loss in a forward-looking perspective.

ASSET SELECTION

Impermanent Loss refers to a temporary decrease in the value of an LP's (Liquidity Provider) deposited assets caused by changes in the relative prices of the tokens within the liquidity pool.

The impermanent loss occurs when the price of the two tokens in the pool diverges from the initial price ratio at which the LP provided liquidity. Since the LP's funds are exposed to changes in token prices, the impermanent loss is a result of the difference in value between holding the tokens in the pool and simply holding those tokens outside the pool.

Intuitively, this suggests that pools which token prices generally move together ie. have similar price trends and tend to be less volatile in terms of impermanent loss.

$$IL = 1 - rac{2\sqrt{r}}{1+r}$$

Figure 1: Impermanent Loss Formula

This means that a crucial step in achieving lower overall impermanent loss is choosing a suitable non-volatile pool, whereby its pool tokens have similar price trends and have significant trading volume for fees to be earned. We use a simple correlation matrix of top tokens in CoinGecko to generate rankings of highly correlated pairs in terms of their returns.

The top ten pool tokens in Uniswap V3 based on past 24H Volume are ETH, WBTC, LINK and DAI

	LINK Price CoinGecko	WBTC Price CoinGecko	DAI Price CoinGecko	ETH Price CoinGecko
LINK Price CoinGecko	1.000000	0.803190	-0.083678	0.828505
WBTC Price CoinGecko	0.803190	1.000000	-0.161683	0.896426
DAI Price CoinGecko	-0.083678	-0.161683	1.000000	-0.184570
ETH Price CoinGecko	0.828505	0.896426	-0.184570	1.000000

Figure 2: Correlation Matrix of Token Returns from CoinGecko

From the matrix analysis, we concluded that the WBTC/ETH token pairing exhibits the highest correlation in terms of returns. Consequently, it is anticipated to have the least price ratio volatility and the lowest hypothetical Impermanent Loss volatility. The high price correlation between Bitcoin and Ethereum is likely due to market sentiment and perception. As the first and most well-known cryptocurrency, Bitcoin often sets the tone for market sentiment whereas Ethereum, being the second-largest cryptocurrency by market capitalization, is influenced by similar market dynamics, thus likely contributing to a high correlation.

A key criteria in determining a suitable LP is the amount of fees an LP can earn by providing liquidity. Maximising the fees within the set range and pair to offset the potential impermanent loss is key to a LP investment. After filtering out stablecoin-stablecoin and ETH-Stablecoin pairs due to the low price correlation, the following are the pairs with the highest traded volume.

Pool	Trading Volume (7D)	TVL
WBTC/ETH (0.05%)	\$295.78m	\$69.75m
WBTC/ETH (0.3%)	\$113.48m	\$228.24m

Figure 3: Comparing the two WBTC/ETH Pools

Given our proposed strategy with a medium to long term horizon, we would use the WBTC/ETH (0.3%) Pool with a higher TVL. Total Value Locked (TVL) is an important indicator of pool health as it signifies greater pool health and investor confidence over a medium to long term investment period. Hence, the WBTC/ETH pool with a 0.3% fee tier emerges as the superior selection for enhancing returns while reducing impermanent loss.

MODEL SELECTION AND PROCESS: ARIMA-GARCH

For our model, we decided to use a hybrid ARIMA-GARCH Model to model and predict daily Uniswap Pool Close Prices, training the model on the WETH/WBTC pool. The rationale for the chosen model primarily lies on two factors:

- 1. Ability to capture temporal trends
- 2. Statistical Interpretability

ARIMA is a popular time series forecasting method that combines autoregression, differencing, and moving averages. Similar to financial markets, we hypothesise that Uniswap V3 pools often exhibit patterns over time, influenced by factors such as market trends, demand, and seasonality. Hence, we deemed the usage of ARIMA to be appropriate.

Preliminary Analysis of Univariate ARIMA

For our analysis, we conducted a stationarity check, seasonal decomposition as well as a Dickey-Fuller Test of our variable Close (WETH) of the Uniswap V3 Pool. These generate the following results

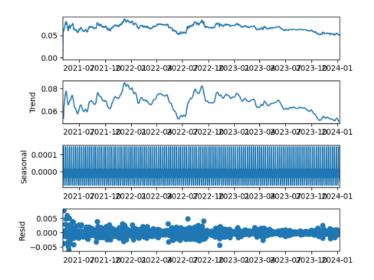


Figure 3: Dickey-Fuller Test Results

This shows that while surprising, Close (WETH) displays a rather stationary type trend, ascertained by a p-value of 0.000000 for the Dickey-Fuller Test, indicating stationarity and no need for transformation to create stationary variables as one would typically. That being said, further analysis on transformed variables would be required.

PACF and ACF Plots

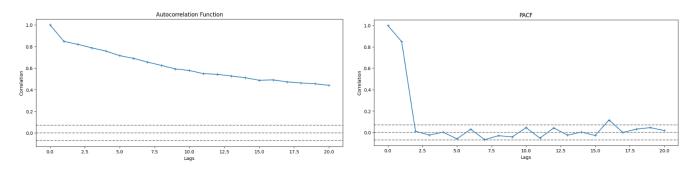


Figure 4: ACF and PACF Plots

Generating Orders

In order to determine the best p, q, d orders for the ARIMA model, we decided to adopt the Grid Search Optimisation procedure for optimising machine learning model parameters, optimising based on AIC, BIC and Log-Likelihood model results. This gives us our optimal ARIMA model: ARIMA(9, 0, 2)

		SARI	MAX Results	;		
Dep. Varia			-	No. Observations:		784
Model:		ARIMA(9, 0,	, 0	kelihood		3531.723
Date:	T	ue, 09 Jan 20	24 AIC			-7037.446
Time:		14:05:	00 BIC			-6976.809
Sample:		05-04-20	21 HQIC			-7014.130
		- 06-26-20	23			
Covariance	Type:	0	pg			
	coef	std err	z	P> z	[0.025	0.975]
const		0.003				
ar.L1	0.5957			0.001	0.235	0.956
ar.L2	0.2398	0.192	1.252	0.211	-0.136	0.619
ar.L3	0.1317	0.186	0.708	0.479	-0.233	0.496
ar.L4	0.1180	0.056	2.089	0.037	0.007	0.229
ar.L5	-0.1905	0.043	-4.438	0.000	-0.275	-0.106
ar.L6	0.2141	0.049	4.377	0.000	0.118	0.310
ar.L7	-0.0065	0.062	-0.106	0.916	-0.128	0.115
ar.L8	0.0271	0.052	0.521	0.603	-0.075	0.129
ar.L9	-0.2030	0.047	-4.349	0.000	-0.294	-0.111
ma.L1	0.3677	0.182	2.015	0.044	0.010	0.72
ma.L2	0.1765	0.173	1.018	0.308	-0.163	0.516
sigma2	7.15e-06	1.26e-07	56.944	0.000	6.9e-06	7.4e-06

Figure 5: ARIMA(9, 0, 2) summary

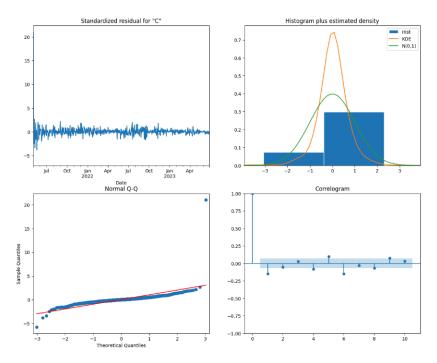


Figure 6: Model Diagnostics

Test on residuals	p-value	Interpretation
Jarque-Bera	0.000	Strong evidence against the null hypothesis of normality. In other words, the residuals are not normally distributed.
Ljung-Box	6.65e-22	Residuals exhibit statistically significant autocorrelation
Box-Pierce	9.02e-22	Residuals exhibit no significant autocorrelation

The Jarque-Bera test suggests that the residuals deviate from the skewness and kurtosis expected in a normal distribution. Although the model diagnostics show a strong fit and accuracy, we opted to integrate a GARCH (Generalized Autoregressive Conditional Heteroskedasticity) model. This decision stems from the recognition that the ARIMA model assumes constant volatility throughout the entire period, an assumption that contradicts the realistic nature of financial markets. The GARCH model is introduced to account for varying volatilities in the market, acknowledging the dynamic and changing nature of volatility over time.

GARCH Model Summary

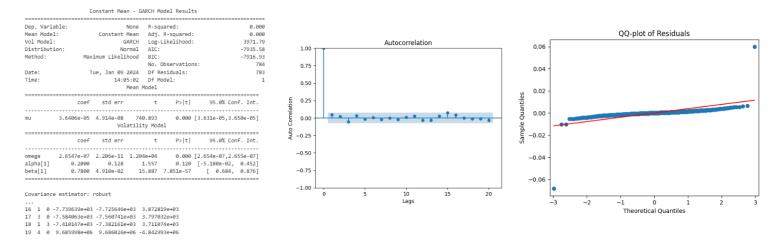


Figure 7: GARCH Model Diagnostics

In-Sample Prediction Plots

Below depicts the figures of the ARIMA plot as well as the ARIMA + GARCH plot taking into account the model's volatility. Using the GARCH's forecasted conditional volatilities, a 95% confidence interval for the predicted price was generated, serving as our upper and lower price range boundaries.

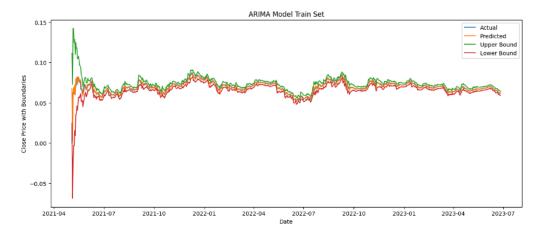


Figure 8: ARIMA-GARCH In-Sample Predictions

Final Results and Performance

Using the Walk-Forward Validation technique and a (80-20) train-test ratio of the WETH/WBTC Pool data from 2021-01-06 to 2023-12-24, we are able to successfully model and predict next step Close (WETH) as well as its upper and lower price boundaries using the conditional variance of the GARCH Model.

RMSE	0.0008553
Mean Absolute Percentage Error	1.03676
Mean Absolute Error	0.000585
Times Outside Bounds	2 out of 196

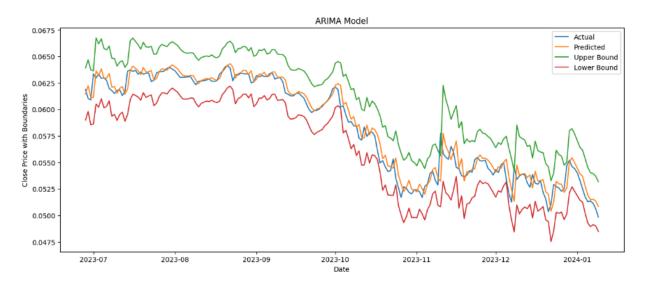


Figure 9: Out of Sample Prediction Results

STRATEGY - Investigating Impermanent Loss

Impermanent loss in the pool refers to the loss suffered by LPs when providing liquidity into the pool as compared to simply holding the assets. The main reason LPs will be willing to subject themselves into the risk of impermanent loss is to be able to farm the transaction fees from swaps made with the pool provided liquidity to.

As such, a useful benchmark for comparing strategies in providing liquidity would be a HODL 50-50 portfolio. Such benchmark comparisons will be made in future backtesting of our strategy.

As LPs, one has to consider how to minimise impermanent loss, as well as optimise returns. We will separate these 2 ideas in the formulation of our strategy by introducing ways of hedging our LP positions as well as how to optimise our returns by setting appropriate upper and lower price boundaries, tested through backtesting and simulations.

As shown in the figure below, the IL curve of a LP position has a negative convexity, whereby we will experience impermanent loss with price changes in both directions. The main idea behind the hedging strategy, would simply to balance out this impermanent loss through the use of an hedging instrument with an ideal payoff equal and opposite to that of our LP position, creating an overall delta-neutral position.

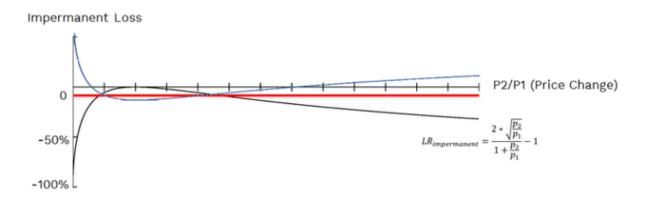


Figure 10: Ideal Hedging Payoff of Uniswap V3 position

STRATEGY - Hedging

In our hedging strategy, we start with a notional value of \$1 million USD, evenly split between WBTC and ETH, based on a price ratio of 0.0538 ETH/WBTC. This translates to holding 219 ETH and 11.7786 WBTC, forming our liquidity provision in the ETH/WBTC pool.

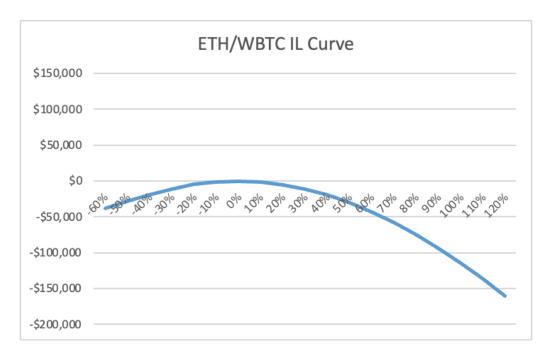


Figure 11: IL Curve for ETH/WBTC

Referring to Figure 11, which illustrates the Impermanent Loss (IL) Curve for ETH/WBTC within a range of -60% to +120%, we observe significant IL exposure.

To mitigate this risk, we adopt a hedging approach involving the underlying assets, BTC-USD and ETH-USD. Our aim is to achieve a delta-neutral position through the utilisation of long put and call options on Deribit.

Deribit, now the largest cryptocurrency derivatives exchange, specialises in high-speed trading of Bitcoin and Ethereum options and futures. Its dominance in the market leads to higher trading volumes, offering better price discovery and a broader range of strike prices for options. Notably, in contrast to other major exchanges like Binance, Deribit offers fractional options, allowing for greater flexibility in hedging strategies.

In Figure 12, the orange line depicts the inverse payoff using ETH and BTC options. Our strategy employs monthly option contracts, positioned 20% to 40% away from current prices at 10% intervals. For a notional deposit value of 1m, we long put and call contracts for each interval. The ratio of the amount of ETH options versus BTC options to

purchase will be based on the current price of ETH/BTC. At \$US 1 million notional, we will purchase 3 ETH contracts and the equivalent amount in BTC options based on price ratio. This effectively minimises the rate of loss, as illustrated by the orange line. This strategy is designed to create an inverse payoff structure aligned with the IL curve mentioned earlier. This approach contrasts with the IL rate (Blue line), highlighting the effectiveness of our hedging measures.

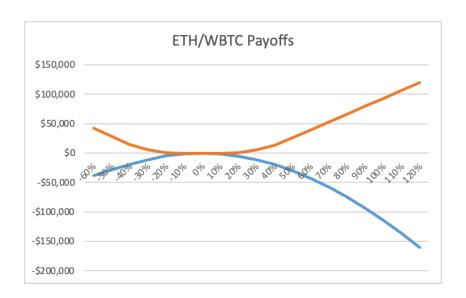


Figure 12: Inverse payoff using ETH and BTC Options (Orange)



Figure 13: Rate of loss when hedging (Orange), IL Rate (Blue)

Notably, the cost of premiums for these options amounts to \$2020.26 USD. This figure serves as a rough estimate, encompassing the expense associated with rolling the options on a monthly basis. Given the perpetual nature of LPs with no expiration, we could even explore the adoption of a perpetual option with a continuous variable premium, providing optimal term matching for our strategy. This ensures ongoing risk mitigation and cost-effective

management in line with the perpetual nature of liquidity provisions.

"Hold On Tight" - Basic One Boundary Strategy

Our trading strategy has a medium to long-term horizon, whereby we will aim to provide liquidity into a Uniswap V3 liquidity pool for a given duration of time. We will then use the price of a futures contract of our token pairs as a proxy for the expected future price at the end of our investment period.

Futures are used to gauge market consensus on potential price deviations from our entry point, serving as an indicator of market trends, supply-demand dynamics, interest rates, and other key economic factors. For our trades on the 0.3% WETH/BTC liquidity pool (ID:0xcbcdf9626bc03e24f779434178a73a0b4bad62ed), we use BTC-USD, and ETH-USD Futures prices on Deribit, comparing the two to find the ending price ratio.

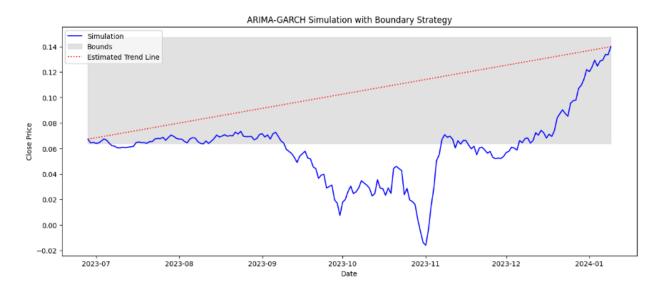


Figure 12: Range Setting 1 Boundary

Using our ARIMA-GARCH model, we will then forecast the estimated volatilities of our liquidity pool both at the start of our investment as well as in the future, plot a confidence interval price bound around both our start and end prices to create a "hypothesised" region where we will expect the price to fluctuate with high levels of confidence. By constructing a price boundary centred around our start and expected future price, we believe that this will allow us to generate substantial returns through fees.

Dynamic Boundaries and Monte Carlo Simulations

For LPs that wish to have rebalancing periods as they provide liquidity, the strategy can be further developed by deciding on the number of times they wish to rebalance. From there, the entire investment period will be further divided into individual windows, with their own estimated volatilities and expected future price based on the price

of the futures contract at the end of the window period. This will allow optimised boundaries for each specific window and 'farm' the fees with greater capital utilisation. Moreover, this will also allow us to introduce greater flexibility and customization with both our options hedging strategy as well as our investment strategy.

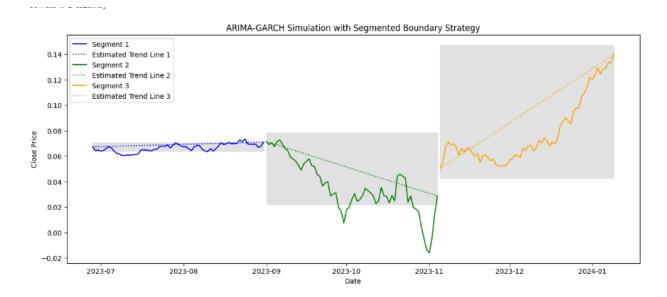


Figure 13: Range Setting 3 Boundaries

Using our ARIMA-GARCH model, we will generate Monte Carlo simulations to simulate our trading strategy, examine performance metrics and optimise our price boundaries in order to maximise our returns.

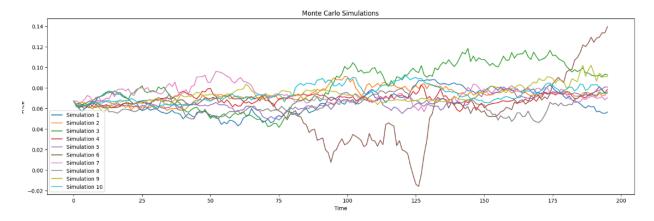


Figure 14: Monte Carlo Simulations

Phase 2: Developing Backtesting Engine: Fee Calculations

Taking inspiration from Defi Lab, we utilise a similar fee calculation strategy in order to accurately calculate how much fees would a simulated LP position earn, given a price boundaries and initial investment amount. The main premise of this calculation procedure is to analyse historical fees earned by 1 unit of unbounded liquidity when the contract was initialised, and multiplying that by the concentrated liquidity multiplier defined by our price range, as

well as the time spent in range.

- 1. Extract Historical Data for a specific Uniswap V3 Pool using TheGraph and calculate fee earned by 1 unit of unbounded liquidity in the period
- 2. Initialise liquidity positions for both a simulated strategy and unbounded strategy, determine the liquidity provided for both cases
- 3. Calculate fees earned by an unbounded position, determined by the amount of liquidity it provided * fees earned per unit of unbounded liquidity. Multiply it on top of a multiplier (ratio of liquidity strategy to liquidity unbounded) and times spent in range

Black-Scholes Merton estimation of Historical Option Premiums

Usually traders use more sophisticated and complex models to price options. However, given resource constraints, we decided to adopt the Black-Scholes Option Pricing Model from the traditional finance sector in order to estimate our option premiums that our strategy will incur as we set up the hedging positions.

The Black-Scholes model is used for providing a good approximation for BTC and ETH options premiums, which is particularly valuable given the limited historical data for these cryptocurrencies. This model's ability to incorporate key variables such as asset price, strike price, and volatility makes it useful for estimating total hedging costs for liquidity provider positions on Uniswap.

Since historical option premium data is not available for BTC and ETH readily, we create our own data set. We use the Black-Scholes Merton model to compare and calibrate the implied volatility(IV) for an option associated with a particular underlying coin with current options data on Deribit. We then employ this IV to find the premium for historical dates and use this IV to estimate the cost of options in our hedging strategy.

Secured Overnight Financing Rate (SOFR) was used to price the BTC-USD and ETH-USD options, as it reflects a nearly risk-free rate based on actual transactions in the U.S. Treasury repurchase market. This rate is more reflective of current market conditions compared to traditional measures. In the volatile and rapidly evolving crypto market, SOFR offers a more accurate and up-to-date benchmark for option pricing, leading to more realistic valuations and better risk management in the inherently unpredictable cryptocurrency market.

Utilising Futures Data

In order to determine our estimated end price, we utilised Futures data for BTC/USD and ETH/USD from Yahoo Finance and obtained their relative ratios against each other. Given the lack of expired futures data for both coins on the internet, to estimate a forward futures data from our entry points, we utilised the futures pricing equation to backtrack an assumed spot price for any specific day, using the relevant coin futures contracts (BTCF.CME,

ETHF.CME) historical prices.

Average SOFR rate of 5.15% was assumed during this period for calculation of assumed spot prices. No carrying costs or dividends were used in the calculation. When backtracking using this calculation, the assumed spot price should theoretically capture future trends and sentiment, based on trade flow and price discovery of the January Futures contract, while accounting for interest rates.

This assumed spot price will be used as a placeholder for a futures contract with the same expiry date, and will be used as exit points of our windows.

Algorithm

- 1. Given a user-inputted number of windows, divide up the test period into windows with each window length minimally one month long. This is due to our use of options contract pricing that requires an appropriate term length.
- 2. For each window, given the start price of the pool close as well as the estimated end price given by Futures data, generate boundaries based on forecasted volatility of the ARIMA-GARCH model at the start and end price, as well as the confidence interval (risk level) that the user inputs.
- 3. Initialise long strangle hedging strategy (+- 20%, 30%, 40%), proportional to 1 million notional amount, calculating premiums using our Black-Scholes Merton estimation of Historical Option Premiums. Simulate LP investment into the liquidity pool
- 4. At the end of the window, exit the LP. Calculate fees generated within the window, payoffs from the hedging strategy and tabulate cumulative investment amount in USD.
- 5. Repeat Steps 2 4 for each window till we reach the end of the investment duration

Backtesting Preliminary Results (Windows=3, Investment=1 million, risk = 0.95)

Utilising hourly pool data for our test period 2023-05-25 to 2023-12-24, our backtesting engine was able to calculate fees based on the above fee calculation methodology. It was also able to accept user-input parameters (confidence interval, investment amount, number of windows) to successfully simulate a LP investment with our "Hold On Tight" strategy. Below shows the boundaries for each window in our strategy.

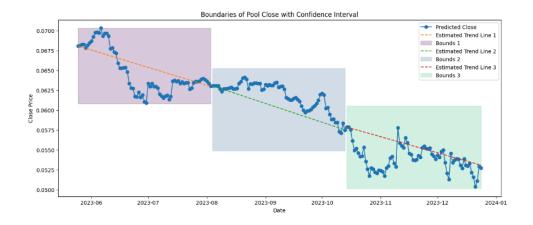


Figure 15: Boundaries of Pool Close with Confidence Interval

As shown, our ARIMA-GARCH model was able to predict volatilities with high accuracy, allowing us to set appropriate bounds that encapsulated the price movements with our forecasted price trend. Comparing with an unbounded strategy, our bounds setting strategy was able to achieve a significantly higher APR as compared to an unbounded one



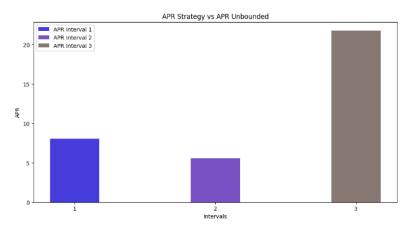


Figure 16: APR Results

Moreover, we were also able to maintain a reasonable level of impermanent loss throughout each window. This is attributed to the customised boundaries and trend line for each window.

Interval	Mean	Standard Deviation	Minimum	Maximum
1	-0.007212	0.006764	`-0.023981	0.0

2	-0.002082	0.003243	-0.013066	0.0
3	-0.007842	0.005862	-0.029509	0.0

Figure 17: IL Results

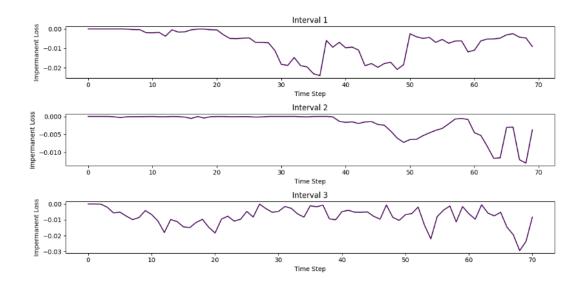
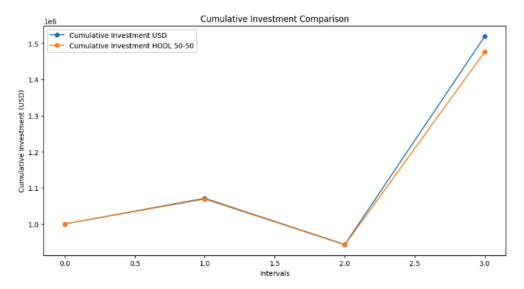


Figure 18: Tracking Impermanent Loss

Benchmark Comparison

Like all Uniswap V3 investment strategies, it is always important to benchmark our results against the HODL 50-50 performance. By tracking our fees accumulated together with our hedging costs and payoffs, we tracked our investment strategy against the HODL 50-50 portfolio, which comprises BTC and ETH of equal value at the start of our test period. Below shows the cumulative value of our strategy and the HODL 50-50 throughout our windows.



Interval	Cumulative Investment (USD)	HODL 50-50 (USD)	Excess Return (USD)
1	1071268.4775	1069101.3209	2167.1566
2	943154.3415	942150.9634	1003.3781
3	1520098.4625	1476980.3168	43118.1457

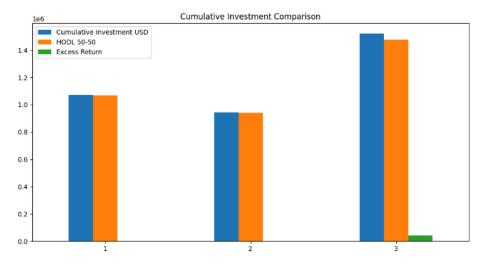


Figure 19: Uniswap V3 vs HODL 50-50 portfolio

Dissecting Fees, Hedging Costs and Payoff

Our results show that our strategy closely matches and slightly outperformed the HODL 50-50 portfolio in terms of cumulative investment. This is a positive sign in our strategy development and provides great motivation for further optimisation.

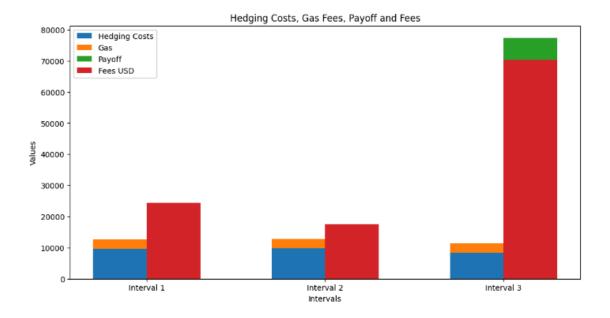


Figure 20: Dissecting Costs and Fees

Preliminary Conclusions

With our successful implementation of a backtesting engine, capable of simulating our "Hold On Tight" strategy as well as generating performance metrics against our benchmark comparison, this indicates that we are on track as per our projected roadmap.

Our results indicate that we were able to accurately set data-driven range boundaries for our investment windows and generate fees. That being said, there are further improvements on our strategy that we will work on in the next phase. These improvements are mainly further optimization of our hedging strategy based on price volatilities so that we are able to construct a data-driven hedging strategy that is able to hedge our impermanent loss optimally for our chosen liquidity pool.

Whats Next

We have successfully completed Phases 1 and 2 with the creation of the backtesting engine and Stream-lit dashboard. As per our projected roadmap, we will further optimise our hedging strategy through optimization methods such as Bayesian Optimization or Genetic Algorithms, with the aim being to create a dynamic hedging strategy that is able to strike a balance between impermanent loss hedging and cost, trained using historical data and price volatilities.

Moreover, we will delve into experimenting with dynamic window settings as part of our ongoing efforts to optimise our boundary-setting strategy. In doing so, our goal is to efficiently determine the most suitable window length for each interval. This approach aims to maximise fee generation within each window, contributing to a more effective overall strategy. Determining the best dynamic window settings could be done through forecasting price volatilities and the use of our Monte Carlo simulations

Projected Roadmap

Phase 1: Backtesting and Simulation Implementation

Objective: Validate the effectiveness and reliability of the hedging and investment strategy through a thorough backtesting process.

• Collect comprehensive historical data, such as historical Pools day data, Ticks day data as well as Swaps

- Model the impermanent loss and create a hedging strategy to hedge against the impermanent loss.
- Create a successful statistical model that would allow us to estimate volatilities and form the basis for boundary setting
- Utilised volatilities to form a fee generating strategy as well as perform preliminary analysis of strategies to test feasibility
- Simulate strategy with Monte Carlo Simulations

Phase 2: Backtesting Engine Creation

Objective: Create a robust backtesting engine that would allow us to simulate LP investments and hedging strategies. Generating meaningful and accurate results that would allow for further optimization of strategies.

- Incorporated ARIMA-GARCH model in our backtesting engine, successfully creating a backtesting engine that can test out our strategy.
- Generated preliminary analysis of results such as fees, hedging costs, as well as performance benchmark comparisons against the HODL 50-50 portfolio
- Created a Stream-lit Dashboard to allow us to visualise and backtest our strategy

Phase 3: Hedging Strategy & Investment Strategy Optimization

Objective: Enhance the sophistication of the hedging approach by identifying and implementing the most suitable combinations of call and put options. Analyse optimal parameters such as number of boundaries or investment windows for maximum fees

- Conduct a detailed analysis to identify optimal combinations of call and put options within predefined price boundaries and ranges.
- Implement a fine-tuning process to achieve an optimal balance between risk mitigation, capital efficiency, volatility of Impermanent Loss experienced as well as fees generated
- Usage of optimization algorithms such as Genetic Algorithms and Bayesian Optimization