SCUOLA DI MANAGEMENT E ECONOMIA

Corso di laurea in
Economia & Data Science

Harvesting the Volatility Risk Premium in the Crypto-Options Market

Relatore:
Prof. Lorenzo Shoenleber

Candidato:
Giulio Bellini
Matricola 955544

Anno accademico 2022 - 2023

Introduction 2
1 The Crypto-Options Market 2
1.1 Cryptocurrencies 2
1.2 Options Contracts 4
1.3 Crypto Options and Deribit.com 8
2 Harvesting the Volatility Risk Premium 10
2.1 Implied Volatility 10
2.2 Volatility Risk Premium 12
2.3 Delta-Neutral Options Strategies 14
3 Trading Strategies Implementation 16
3.1 Data Analysis 16
3.2 Backtesting Algorithm 20
3.3 Multiple Regression Analysis 22
3.4 Best Model Implementation 28
Conclusion 30

Introduction

With the emergence of cryptocurrencies as a new asset class, financial innovation has led to the creation of derivative instruments such as crypto-options and crypto-futures. In this thesis I present these financial instruments and analyse their use for monetising the volatility risk premium (VRP), a phenomenon already present in the stock market and widely described in the financial literature. In the first part of the paper, I cover basic topics such as cryptocurrencies and options, and then delve into more advanced aspects such as implied volatility and VRP, to provide the theoretical knowledge necessary to understand the following chapters. In the second phase, I will conduct an empirical study based on high-frequency data from Deribit.com to select the best crypto-options trading strategies to monetise VRP. For this purpose, I use a backtesting algorithm I developed in Python and implement regression models.

1 The Crypto-Options Market

1.1 Cryptocurrencies

Wikipedia.com provides the following definition: "A cryptocurrency is a digital currency designed to work as a medium of exchange through a computer network that is not reliant on any central authority, such as government or bank, to uphold or maintain it."[1] Cryptocurrencies, also called digital currencies or digital assets, are born as a medium of exchange, exactly like FIAT currencies, but with the peculiarity of being based on a decentralised system for the validation of transactions, allowing traditional financial intermediaries, such as banks, to be eliminated. They, therefore, constitute a peer-to-peer electronic cash system.

Every cryptocurrency works thanks to blockchain technology, i.e. a distributed ledger that acts as a public database of financial transactions. Currency ownership and transaction security are guaranteed by the blockchain, which, using cryptography, controls the creation and exchange of currencies. In the blockchain, authorised transactions are stored in blocks of records, each of which contains a timestamp, transactional data, and a hash, i.e. an address identifying the block. The hash of each block is calculated by a hash algorithm on the basis of all the data contained in the block and the previous block’s hash. This creates a blockchain that grows over time and records transactions permanently. For the blockchain to be effectively a distributed ledger, it is often administered by a peer-to-peer network of computers, called nodes, which adhere to a protocol, i.e. a set of rules established by the creator

[1] Definition from en.wikipedia.org/wiki/Cryptocurrency

of the blockchain, and perform the function of mining, i.e. verifying transactions and creating new blocks by receiving cryptocurrencies as a reward for their work. The network guarantees the security of the blockchain. Each node owns a copy of the blockchain and validates each new transaction, which is only recorded if the majority of nodes consent to its legitimacy. Transactional data stored in the blockchain, moreover, are unchangeable. If a node tried to modify a previously validated transaction mistakenly or fraudulently, it would change the hash of the block in which it is contained and consequently that of all subsequent blocks. In that case, the copy of the blockchain owned by the node would be different from those owned by others, and the offending node would be excluded from the network. Changing historical transactions would only be possible with the collusion or hacking of most nodes, which is computationally impossible for large networks. The blockchain's design, therefore, makes it a distributed, secure, and reliable system for exchanging value.

The first cryptocurrency created and the most famous is Bitcoin (BTC). The whitepaper containing Bitcoin's protocol was published on 31 October 2008 by pseudonymous developer Satoshi Nakamoto as a response to the financial crisis and a sign of distrust in the traditional financial system. The cryptocurrency officially came into use on 9 January 2009 when the first block, the genesis block, was created. Bitcoin's blockchain uses the SHA-256 hash algorithm and proof-of-work to validate transactions and ensure their security. Bitcoin has a supply limit that sets a cap of 21 million on the creation of new coins. Bitcoin is currently the world's most popular cryptocurrency, with 19,340,166 existing coins, 784,501 mined blocks, a value of USD 28,044 per coin, an average of 13,791 transactions per hour, 17,712 reachable nodes and an estimated 200 million users (wallets).[2]

The creation of Bitcoin started a wave of innovation aimed at creating new cryptocurrencies, finding new applications, solving problems related to security, transaction speed, and cost, and promoting the industry. There are currently more than 9,000 cryptocurrencies[3] that position themselves as alternatives to Bitcoin and are collectively known as 'altcoins' for this reason. The most important altcoin by market capitalisation and celebrity is Ether (ETH), which is based on the Ethereum blockchain. Ethereum was created in 2013 by Russian programmer Vitalik Buterin and enables the deployment of decentralised applications based on smart contracts. Decentralised finance applications (DeFi) allow cryptocurrencies to be exchanged,

[2] According to bitinfocharts.com and bitnodes.io, data on April 8, 2023
[3] According to statista.com, data on April 8, 2023

traded, borrowed, lent, and preserved without the need for financial intermediaries such as banks, brokers, and exchanges.

Despite their name and origin as an alternative medium of exchange and store of value to traditional currencies, cryptocurrencies are often regarded as an independent asset class. The main reasons for this classification are the high volatility of their market price and their low correlation with traditional asset classes such as FIAT currencies, stocks, commodities, and bonds. This makes cryptos a highly speculative asset but at the same time attractive from a portfolio diversification perspective. In the following chapters, I will focus on selecting the best trading strategies to monetise the volatility of cryptocurrencies using new financial instruments available to retail investors.

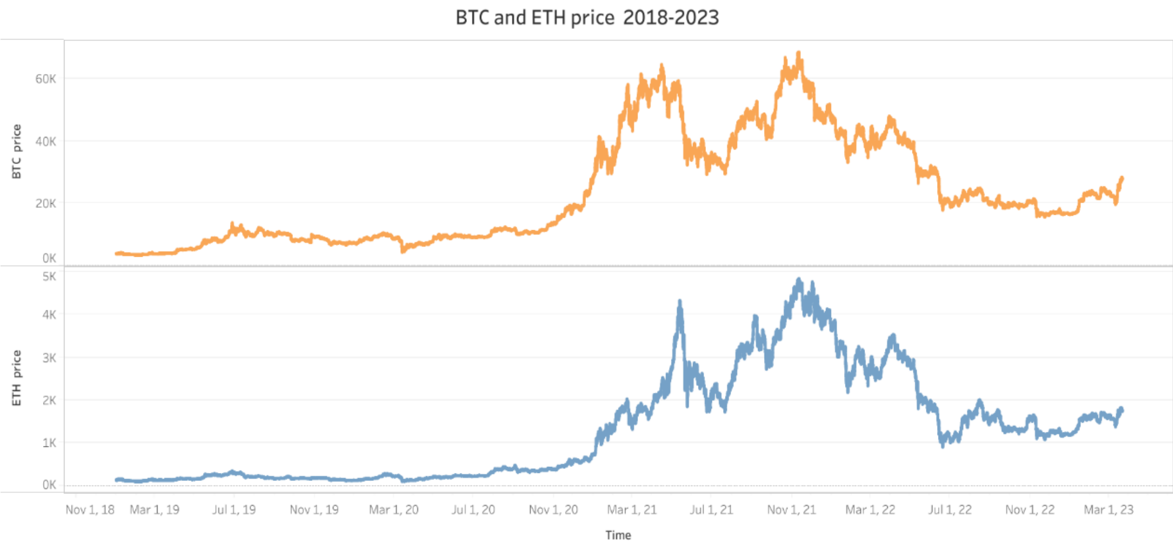

Figure 1: Evolution of BTC price (+590%) and ETH price (+1190%)

1.2 Options Contracts

Options are derivative contracts, i.e., financial instruments whose value is based on that of underlying securities such as shares, commodities, or cryptocurrencies. Options give the buyer the right, but not the obligation, to buy or sell a certain quantity (multiplier) of the underlying asset at a predetermined price based on the type of contract. Each option has an expiration date by which the holder can exercise its right. European-type options can only be exercised at expiration while American-type options can be exercised at any time before expiration. The price at which options can be exercised, i.e., at which the holder has the right to buy or sell the underlying asset, is called the strike price. Depending on the position of the strike relative to the current price of the underlying, options are divided into At-the-Money (ATM), In-the-money (ITM), and Out-of-the-Money (OTM). Typically, an option represents 100 shares of the underlying security (multiplier equal to 100) and thus allows the buyer to hold a leveraged position in the underlying at a low cost.

AAPL Option Chain

Date	Option	Calls & Puts	Moneyness	Type
June 2023	Composite	Calls & Puts	Near the Money	All (Types)

	Calls							Puts					
Exp. Date	Last	Change	Bid	Ask	Volume	Open Int.	Strike	Last	Change	Bid	Ask	Volume	Open Int.
June 2, 2023													
Jun 2	15.70	+0.20 ▲	–	–	193	709	160.00	0.15	-0.03 ▼	–	–	571	7258
Jun 2	13.49	+1.02 ▲	–	–	108	231	162.50	0.21	-0.03 ▼	–	–	100	507
Jun 2	10.75	-0.05 ▼	–	–	155	967	165.00	0.29	-0.05 ▼	–	–	1746	7905
Jun 2	7.90	-0.52 ▼	–	–	93	265	167.50	0.45	-0.08 ▼	–	–	955	2285
Jun 2	6.40	+0.10 ▲	–	–	581	4044	170.00	0.76	-0.09 ▼	–	–	3110	5526
Jun 2	4.40	-0.05 ▼	–	–	631	1046	172.50	1.29	-0.10 ▼	–	–	2118	2998
Jun 2	2.64	-0.12 ▼	–	–	5063	10059	175.00	2.12	-0.15 ▼	–	–	5572	10402
Jun 2	1.44	-0.05 ▼	–	–	7576	2022	177.50	3.45	-0.15 ▼	–	–	275	143
Jun 2	0.67	-0.02 ▼	–	–	10702	8555	180.00	5.19	-0.16 ▼	–	–	293	942
Jun 2	0.30	–	–	–	3069	1073	182.50	7.15	-0.40 ▼	–	–	27	8
Jun 2	0.14	+0.02 ▲	–	–	2895	8873	185.00	10.54	–	–	–	–	2
Jun 2	0.07	+0.01 ▲	–	–	732	341	187.50	–	–	–	–	–	–
Jun 2	0.06	+0.02 ▲	–	–	757	6858	190.00	16.00	–	–	–	–	2
Jun 2	0.04	+0.01 ▲	–	–	302	21	192.50	–	–	–	–	–	–

Figure 2: AAPL option chain for the June 2, 2023 expiry on May 22, 2023, data from nasdaq.com

Options always involve two parties: a buyer (holder) and a seller (writer). The buyer has the right but not the obligation to exercise the option by expiry at the strike price but must pay a premium for the purchase of the derivative contract. The seller, on the other hand, has the obligation to sell or buy the underlying asset from the buyer if the buyer decides to sell, but collects the premium. In each option contract, therefore, the buyer and seller bet on opposite price movements of the underlying asset. They are said to take a long or short position on the underlying security respectively.

There are mainly two types of options: call options and put options.
Call options give the buyer the right to buy a share of the underlying security at the strike price by the expiration date. Their value increases if the price of the underlying rises. *Figure 3* shows the payoff structures of a long (bought) call option and a short (sold) call option. The image shows how the buyer's loss is limited to the premium paid to purchase the contract and how the profit is potentially infinite, while the seller's gain is limited to the premium received and the loss is potentially unlimited.

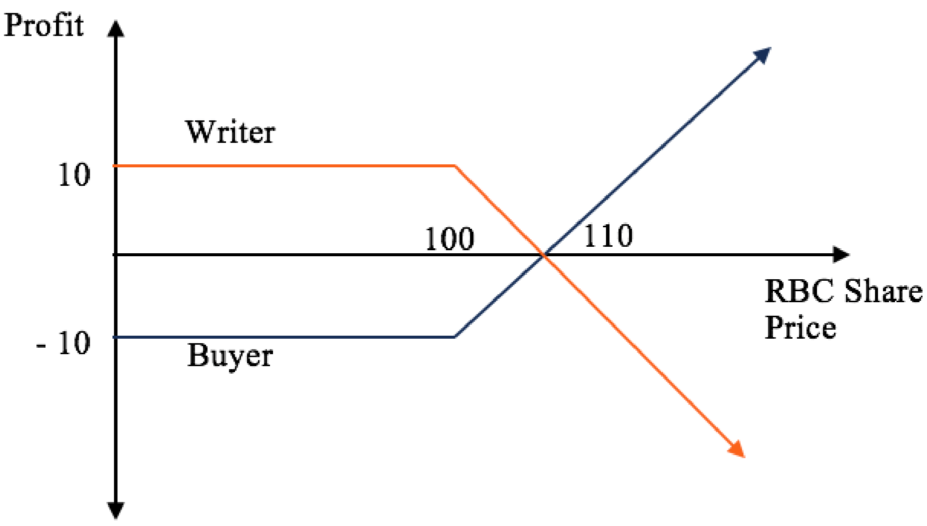

Figure 3: Long and Short call option payoff structures

Put options give the buyer the right to sell a share of the underlying security at the strike price by the expiration date. Their value increases if the price of the underlying falls. *Figure 4* shows the payoff structures of a long and short put option. Again, the buyer has unlimited gains and limited losses, while the seller has limited gains and unlimited losses.

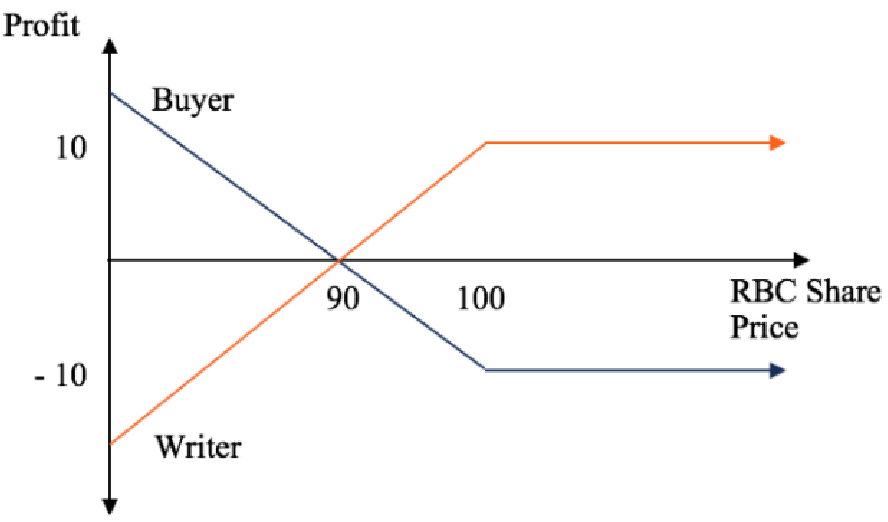

Figure 4: Long and Short put option payoff structures

By their very nature, options have several risk dimensions that are calculated by important risk metrics called 'Greeks'. The Greeks measure an option contract's exposure to the price movement of the underlying asset, the market's expected future volatility of the underlying asset over the life of the contract (implied volatility), and the time remaining to expiration. The most important Greeks are delta and gamma. The delta measures the change in the option price following a $1 increase in the price of the underlying. It, therefore, measures the sensitivity of the option price to the price of the underlying. Call options have a positive delta, while put options have a negative delta. Gamma measures the change in the delta value following a $1 increase in the price of the underlying. Gamma, therefore, is the second-order price sensitivity. Gamma determines the stability of the delta. It is higher for options whose strike price is close to the current underlying price and increases as expiry approaches.

Options are liquid and cheap instruments for portfolio hedging or speculating on the underlying price. Investors can also buy and sell combinations of options, creating option spreads. With options, it is possible to create a multitude of different payoff structures that reflect the investor's risk and return objectives. With spreads, it is possible to control the risk of the position and its exposure to movements in the price of the underlying asset. In the following chapters, we will look specifically at how to construct strategies with zero delta and gamma that gain from changes in the volatility of the underlying.

Figure 5: combining a long call and a short call we reduce the premium, but the profit becomes limited, graph from optionalpha.com

1.3 Crypto Options and Deribit.com

Crypto options are derivative instruments in all respects similar to traditional options, but they have a digital asset such as BTC or ETH as their underlying. Crypto options are relatively new instruments whose trading volume started to increase in early 2020, exceeding $30 billion for BTC options during the peak of the bull run in April 2021[4]. *Figure 6* and *Figure 7* show the trading volumes of Bitcoin options and Ether options by month on the major exchanges.

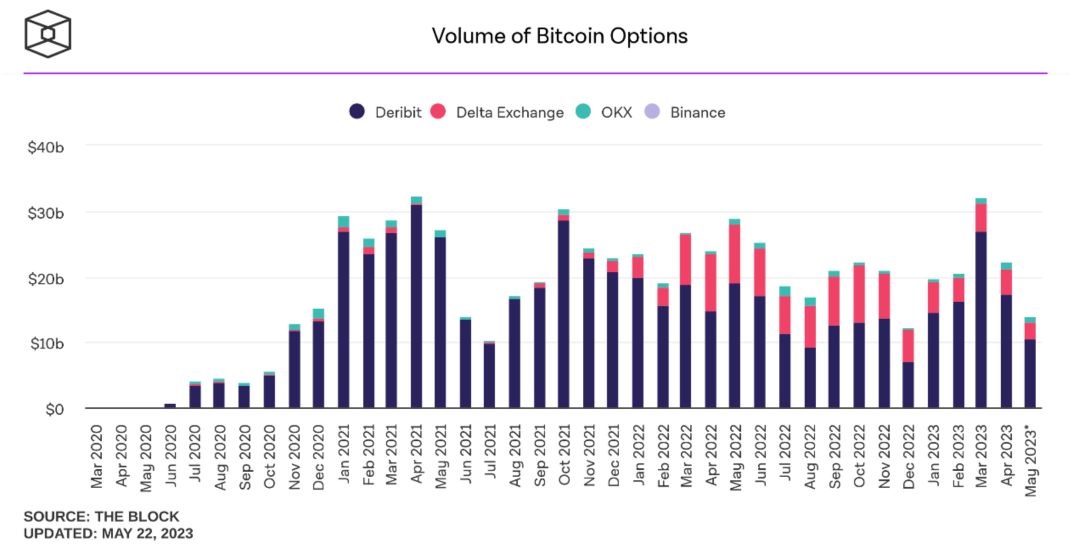

Figure 6: Volume of BTC options traded on major exchanges per month, data on May 22, 2023, from theblock.co

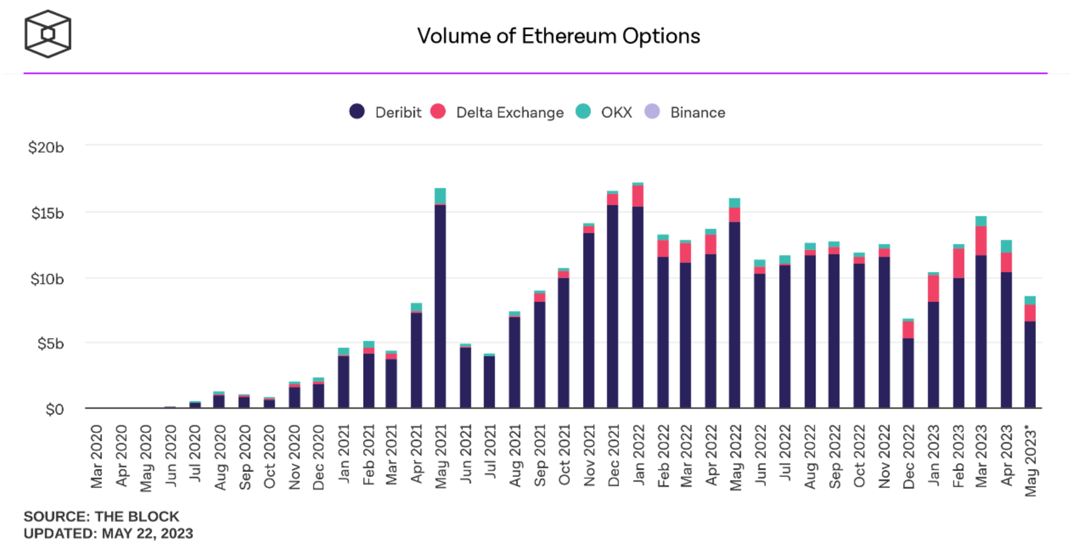

Figure 7: Volume of ETH options traded on major exchanges per month, data on May 22, 2023, from theblock.co

[4] Real-time data available at https://www.theblock.co/data/crypto-markets/options

The largest cryptocurrency derivatives exchange by trading volume is Deribit.com, which was established in 2016 as the world's first cryptocurrency options exchange. It currently offers perpetual, futures, and options on BTC and ETH tradable 24/7 securely through its proprietary technology. The platform was developed to handle large amounts of requests with low latency and is based on a unique matching engine created specifically for Deribit[5].

Deribit offers European-style cash-settled options. Options can therefore only be exercised at expiry when the writer of the contract pays any profit to the holder in BTC or ETH without transferring any assets. Options on BTC (ETH) have 1 BTC (ETH) as the underlying asset. The options are priced in BTC or ETH, but the price is also available in USD. It is important to emphasise, therefore, that by buying or selling a crypto-option on Deribit the investor, whose reference currency is USD, is exposing himself to an additional risk: the crypto-USD exchange rate risk.
Figure 8 and *Figure 9* show, respectively, the structure of the payoff at maturity for the same call option calculated in BTC and USD. Both payoffs take the exchange rate risk into account. Considering the P&L (profit and loss, on the vertical axis) in BTC, the marginal increase in the amount of BTC gained decreases as the price of BTC (on the horizontal axis) increases because a higher value of BTC translates into a lower amount paid at expiry. On the other hand, considering the P&L in USD, we note how the maximum loss (the premium paid) decreases as the price of BTC decreases since a smaller value of BTC results in a smaller loss in USD.

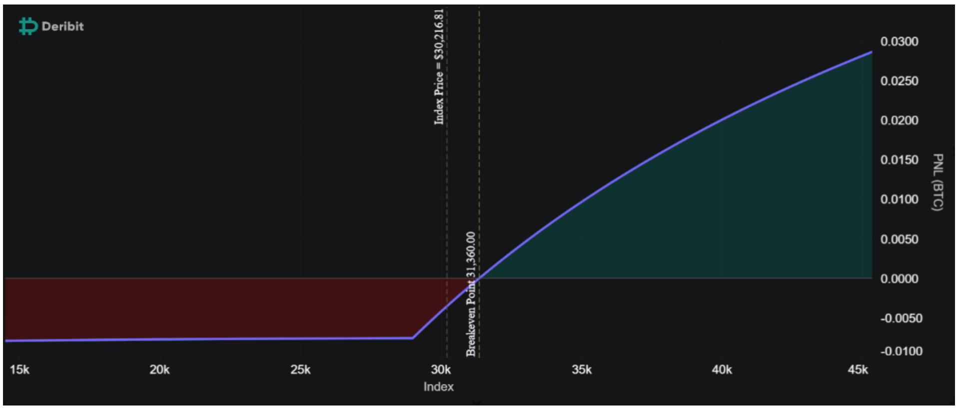

Figure 8: BTC long call option payoff in BTC

[5] All information available at https://www.deribit.com/kb/about-us

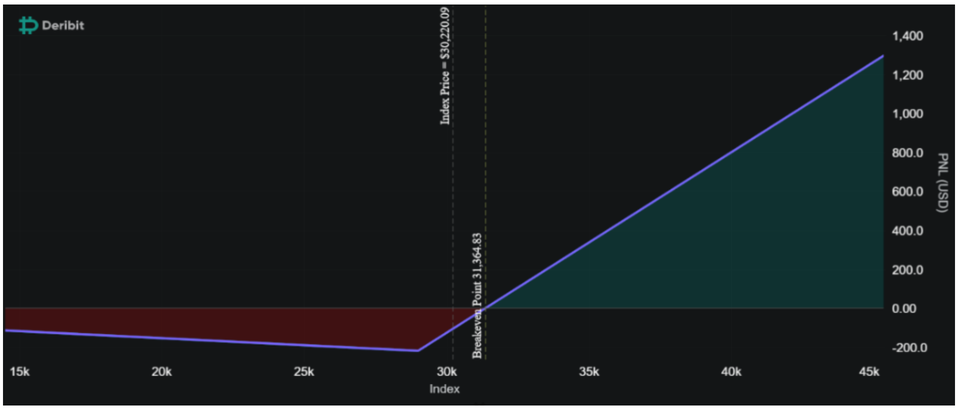

Figure 9: BTC long call option payoff in USD

2 Harvesting the Volatility Risk Premium

2.1 Implied Volatility

Implied volatility (IV) is a forward-looking metric that measures the market's expectation of the future price movement of a security. It is the future volatility estimated by the market and, as such, does not provide an indication of the future direction of the price, only the magnitude of the movement. Implied volatility is not a metric directly observable in the market, but is calculated from an options pricing model, a mathematical model that assigns an option its theoretical fair value (premium) from variables such as the current underlying price, strike price, moneyness, maturity, risk-free interest rate, and future volatility. Most of the previous inputs are dictated by the market, as is the actual option price, while future volatility is not. For this reason, option price models are often used to calculate the estimated future volatility (option implied volatility) from the other variables.
The most widely used option price model is the Black-Scholes-Merton (BSM) model, published in 1973 in an article called "The Pricing of Options and Corporate Liabilities" (Black, 1973). The BSM formula shown in *Figure 10* is the solution of the BSM differential equation and it’s used to price European options.

$$
\begin{aligned}
C(S_t, t) &= N(d_+)S_t - N(d_-)Ke^{-r(T-t)} \\
d_+ &= \frac{1}{\sigma\sqrt{T-t}}\left[\ln\left(\frac{S_t}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)(T-t)\right] \\
d_- &= d_+ - \sigma\sqrt{T-t}
\end{aligned}
$$

Figure 10: Value of a European call option on a non-dividend-paying stock

The BSM model is based on restrictive assumptions such as that the returns of the underlying asset are normally distributed, that there are no transaction fees, that the option is European style, and that the future volatility is known and constant over the option's life. Despite this, it is still widely used by brokers (such as Deribit), traders, and market makers, especially to calculate the implied volatility of options.
Each option contract, therefore, has its own implied volatility, which is univocally related to its price. The IV is in turn influenced by many factors (underlying price, strike price, maturity, etc.), which makes it a highly variable metric over time. To get an overall view of the market's implied volatility and follow its trend, the Chicago Board Options Exchange, the world's largest options exchange, created the Cboe Volatility Index (VIX)[6], which tracks the 30-day IV of the entire S&P 500 Index (SPX). The VIX extends implied volatility to the market level by aggregating the IV of options on the SPX with a maturity of around 30 days and different strike prices. The VIX is also called the "fear index" because the IV tends to be negatively correlated with the price of the underlying and, therefore, market crashes are often associated with spikes in the value of the VIX, as shown in *Figure 11*. Formulas like the one used by the Cboe to calculate the VIX are currently used to calculate the IV of any underlying.

Figure 11: SP500 and VIX are negatively correlated (-0.7). The 2020 crash (-34%) coincides with the peak in the VIX (82.7)

In the era of crypto-options, indices have been created to track the implied volatility of BTC and ETH. In the following chapters, I use BitVol and EthVol[7], two indices

[6] Real-time data available on https://www.cboe.com/us/indices/dashboard/VIX
[7] Real-time data available on https://t3index.com/indexes/bit-vol

developed by T3 Volatility Indices in 2019 that measure the expected 30-day implied volatility in BTC and ETH, respectively.

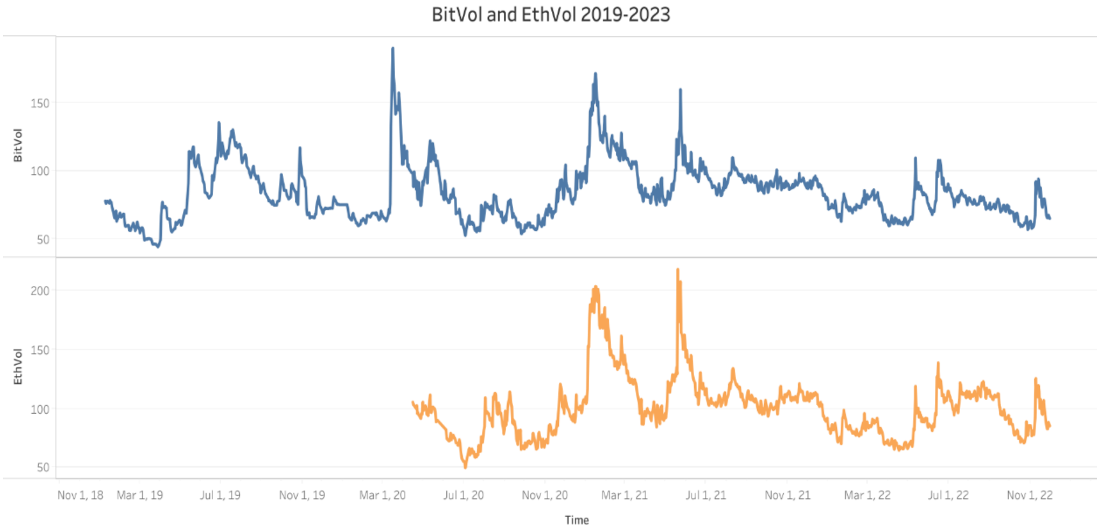

Figure 12: High volatility of BTC and ETH results in large values of BitVol and EthVol with respect to the VIX

2.2 Volatility Risk Premium

The Volatility Risk Premium (VRP) is defined as the difference between the realised volatility (RV) of an asset and its implied volatility (IV) over the same period. Considering a period ΔT and time scales, therefore, realised volatility is backward looking over ΔT, while implied volatility is forward looking over ΔT. Therefore, to visually compare them and calculate the historical VRP we move RV backward by the ΔT period.

$$\mathrm{VRP}_{\Delta\mathrm{T}} = \mathrm{RV}_{\Delta\mathrm{T}} - \mathrm{IV}_{\Delta\mathrm{T}}$$

Literature shows the empirical existence of VRP in equity index options (Bollerslev, 2009) and the fact that it assumes low negative values most of the time (Kozhan, 2013) and high positive values during rare extreme events (market crashes). In most cases, therefore, implied volatility tends to overestimate realised volatility. This discrepancy results in market prices of options that are generally higher than their theoretical prices (obtained by applying an option pricing model). Option writers can profit from this risk premium by selling the contracts at a price higher than their fair value. At the same time, however, a short option position can cause huge losses in the event of large movements in the price of the underlying asset. The VRP can be partly justified by thinking of option contracts as insurance policies: the buyer pays a premium to cover himself against the risk of volatility during periods of market turmoil. Goetzmann and Kim (Goetzmann, 2017) showed how market participants tend to overestimate the probability of a market crash due to the influence of the media. This heightened perception of risk leads to an increase in the demand for options and thus their price. Carr and Wu (Carr, 2008) show that investors are willing to pay extra money to buy overpriced options and protect themselves from volatility. The VRP, therefore, is typically negative because investors pay to include variance in their portfolios due to its negative correlation with the stock market.

The literature concerning the VRP in the crypto-options market is limited because the necessary data on option prices are only available from 2019. Carol Alexander (Alexander, 2020), after calculating her own VIX-style BTC implied volatility index (BVIN) for different maturities using options data from Deribit.com, shows how the Bitcoin VRP took predominantly negative values between March 2019 and March 2020 and how the longer-term VRPs are less variable. On top of that, since November 2019 the Bitcoin VRP has become more stable, taking small and negative values most of the time. *Figure 13* compares 30-day BTC RV and 30-day BTC IV over time, while *Figure 14* shows the Bitcoin 30-days VRP.

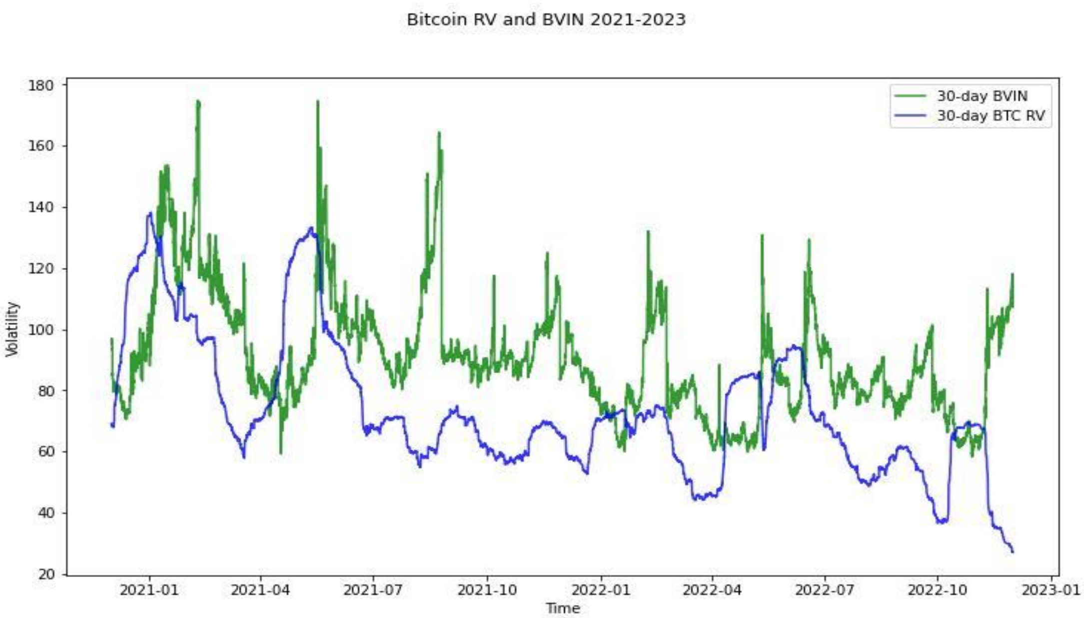

Figure 13: Comparison between RV moved back 30 days and BVIN 2021-2023

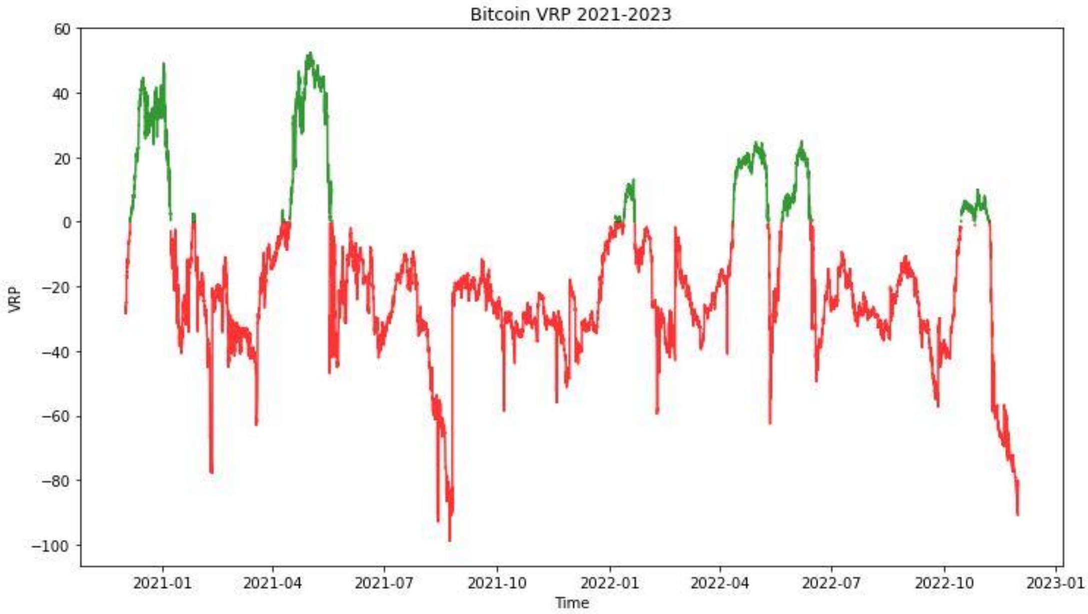

Figure 14: Bitcoin VRP was negative 78.5% of the time between 2021 and 2023

2.3 Delta-Neutral Options Strategies

To profit from the existence of a predominantly negative VRP, it is necessary to open delta-neutral short option positions, i.e. with zero exposure to the movements of the underlying. Since the present VRP cannot be calculated without knowing the future volatility, we can limit ourselves to basing our trading strategy on the fact that in the past the implied volatility has consistently overestimated the realised volatility. We can then proceed by selling options that we believe to be overpriced by combining them in a way that minimises exposure to factors other than volatility. The most used delta-neutral option strategies are straddles and strangles. Short straddles are constructed by selling one ATM call and one ATM put, thus with deltas close to 0.5 in absolute terms. Short strangles, on the other hand, combine a short OTM call and a short OTM put with the same delta in absolute terms. In both cases, the strategy has zero delta ($Delta_{call} + Delta_{put} = 0$) and is profitable when the price of the underlying remains in the range ($Strike_{put} - Price_{put}$; $Strike_{call} + Price_{call}$) until the trade is closed. By selling a straddle, a higher premium ($Price_{call} + Price_{put}$) is collected since ATM options are generally more expensive, but the probability of incurring a loss is equally high due to the proximity of the strike prices. It is necessary to seek a trade-off between the premium and the options deltas/strikes.

It must be underlined that straddles and strangles are delta-neutral only at the exact moment of the trade because, due to the gamma effect, as the price of the underlying moves, the delta of the position also changes. To maintain a delta-neutral position over time it is necessary to hedge it. This can be done by adding other contracts or shares of the underlying to the position to set the total delta to zero or by rolling it out, i.e. closing and reopening the position periodically following the price of the underlying. In the former case, however, more exposure to the underlying is added to the portfolio, while in the latter the transaction costs could heavily impact performance. As usual, the key is to seek a trade-off.

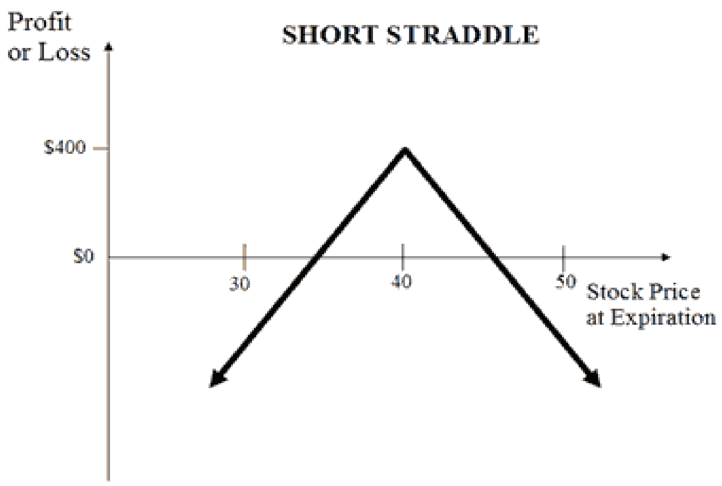

Figure 15: Short Straddle payoff structure

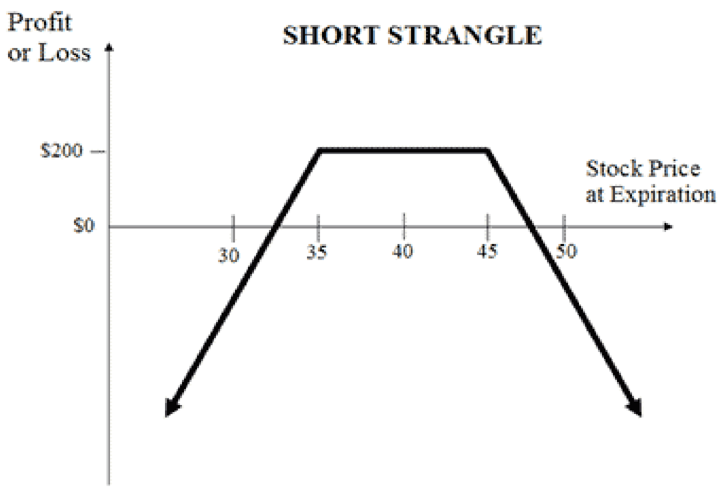

Figure 16: Short Strangle payoff structure

3 Trading Strategies Implementation

Historical data show the existence of a negative VRP in the BTC options market between 2021 and 2023. In this chapter, I implement a backtesting algorithm to study the performance of strategies consisting of periodic selling of straddles and strangles. The objective is to understand the effect of factors such as implied volatility, liquidity, maturity, and delta of short positions on the performance of the strategies to select the best trading criteria for monetising VRP.

3.1 Data Analysis

In my empirical study, I use options chains high-frequency historical data from Deribit.com covering 15 months from January 2021 to March 2022, including the 2021 boom and part of the 2022 crash in the cryptocurrency market. The high-frequency data contains key variables (symbol, maturity, delta, gamma, etc.) on all option contracts traded on Deribit at any given time, resulting in gigabytes of data available. Since this study does not require extreme temporal detail, I simply record information for all available contracts at one-hour intervals, resulting in a 2.2 GB dataset with 4.7 million observations.

To select the options with which to create the straddles and strangles to sell, I select the contracts I consider "tradable", discarding all those that do not fall within the limits dictated by Deribit's market-making obligations[8], according to which each market maker:

- must quote all expiries and 90% of all options contracts with delta between 0.1 and 0.9 in absolute terms.
- must guarantee a bid-ask spread of no more than $max(0.01; |delta| \times 0.04)$.

The application of these delta and bid-ask spread constraints reduced the size of the dataset by 34% (to 1.62 million observations) but resulted in the elimination of 95% of the missing values. This can be partially explained by the fact that far-out-of-the-money options, which are exempt from market-making obligations, are also the least liquid options for which there are often no quotes available (meaning missing values for the bid price, ask price and Greeks).

Starting from the data set thus obtained, I performed an exploratory analysis of the data to get an overview of the market, in terms of heterogeneity and liquidity of the contracts, which would provide me with useful information for backtesting.

[8] All information on deribit.com/kb/options

Deribit offers a wide range of expiry dates[9]. Daily options (1, 2, and 3 days) expire every day at 8 a.m. Weekly options (1, 2, and 3 weeks) expire every Friday at 8 a.m. Monthly options (at 1, 2, and 3 months) as well as quarterly options (at 6, 9, and 12 months), as well as quarterly options (at 6, 9, and 12 months), expire on the last Friday of the month at 8:00 am. During the period covered by the data, Deribit offered options with strikes between $4000 and $400000 of which only those with strikes between $15000 and $300000 are included in the filtered dataset. This wide range of strike prices is rather surprising considering that during the same period, the price of BTC reached a low of $29251 and a high of $68624. This shows the extreme uncertainty about the future trend of the BTC price due to its high historical volatility.

Regarding liquidity, I considered two indicators: the bid-ask spread and the open interest of options.
The bid-ask spread is the difference between the bid price (price at which you can buy) and the ask price (price at which you can sell). It is always positive since the bid price is always greater than the ask price, which means that by buying and selling an asset without its price changing you incur a loss, which depends on the size of the spread. The bid-ask spread is an important indicator of liquidity: in a liquid market with many competing participants, the bid and ask tend to get closer as market makers are willing to make smaller profits to see their orders executed. This plays to the advantage of investors, who in a liquid market can trade without incurring frequent losses due to a wide bid-ask spread. I analysed the average trend of the bid-ask spread during the hours of the day and over the life of the contracts from the filtered dataset. Having applied market-making obligations, I expect the bid-ask spread to always be less than 0.036 = 0.9 * 0.04. Over the period considered, the average bid-ask spread in percentage terms (w.r.t. ask price) is 6.3%, which makes the options market on BTC not very liquid, especially considering that ETFs listed on US exchanges have an average spread of 0.52%[10] and that BTC has an average spread of 0.03%. (Aleti). *Figure 17* shows the mean and standard deviation of the spread in percentage terms during the hours of the day. Both the mean and standard deviation of the percentage bid-ask spread tend to take on higher values during the night and lower values during the day with a reversal of the trend around 8 am, the time when options expire and new expiry dates are introduced on Deribit. The lowest values are touched between 11 a.m. and 5 p.m. when I execute trades in backtesting.

[9] All information on deribit.com/kb/deribit-introduction-policy
[10] According to etf.com

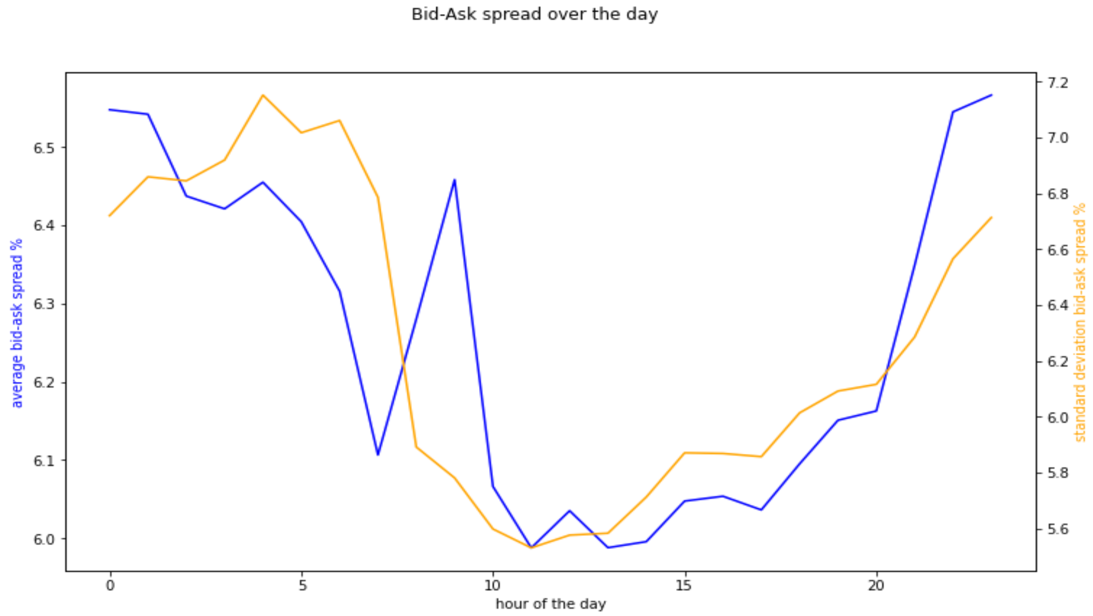

Figure 17: Bid-Ask Spread average and standard deviation over the day

Figure 18 shows the average bid-ask spread in both absolute and percentage terms over the life of the contracts. The two lines follow opposite trends. As maturity approaches, the spread decreases linearly in absolute terms but increases exponentially in relative terms. This may be due to 'theta decay', i.e. the loss of value of options, especially out-of-the-money options, as expiry approaches

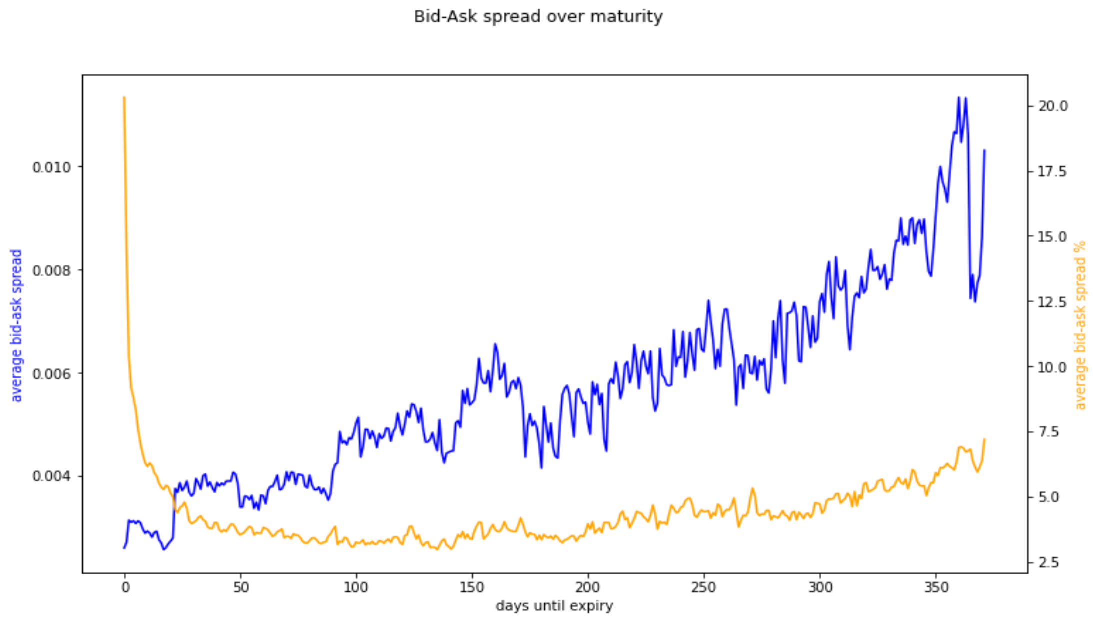

Figure 18: Bid-Ask Spread absolute and percentage over maturity

The second liquidity indicator I consider is open interest. Open interest is the total number of outstanding option contracts that have not been settled for an asset (underlying). It keeps track of every open position in a particular contract, rather than tracking the total volume traded in it, which may also include netting or closing positions. The higher the open interest of an option, the more liquid it is since many market participants have open positions in it. Investors prefer liquid contracts not only because they have a lower bid-ask spread, but also because they can be traded faster. In the period under consideration, the average open interest is 544 open contracts, which proves that the options market on BTC is still not very liquid. In *Figure 19*, one can see how the open interest varies during the day. Interestingly, the average open interest gradually increases from 10 am until 8 am, when there is a collapse in the number of open contracts. This is because options on Deribit expire and are automatically exercised at 8.00 am.

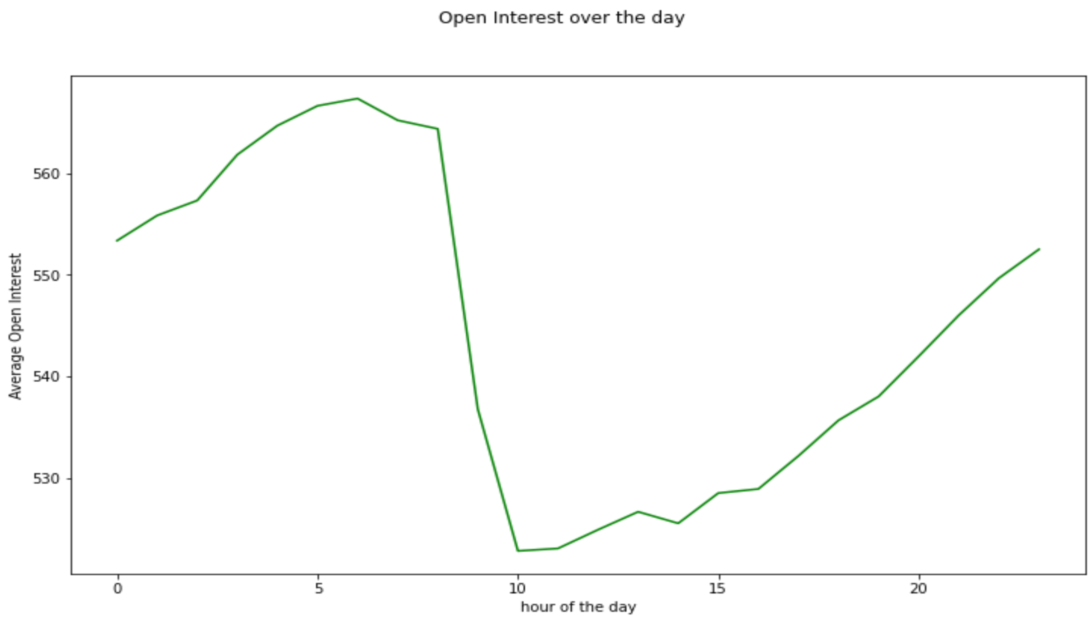

Figure 19: Open Interest average over the day

Figure 20 shows the trend in average open interest for different maturities. It appears that investors tend to have few open positions in options very close to or very far from expiration. We also see a peak in liquidity in contracts with maturities around 25 days, i.e. with maturities around one month away.

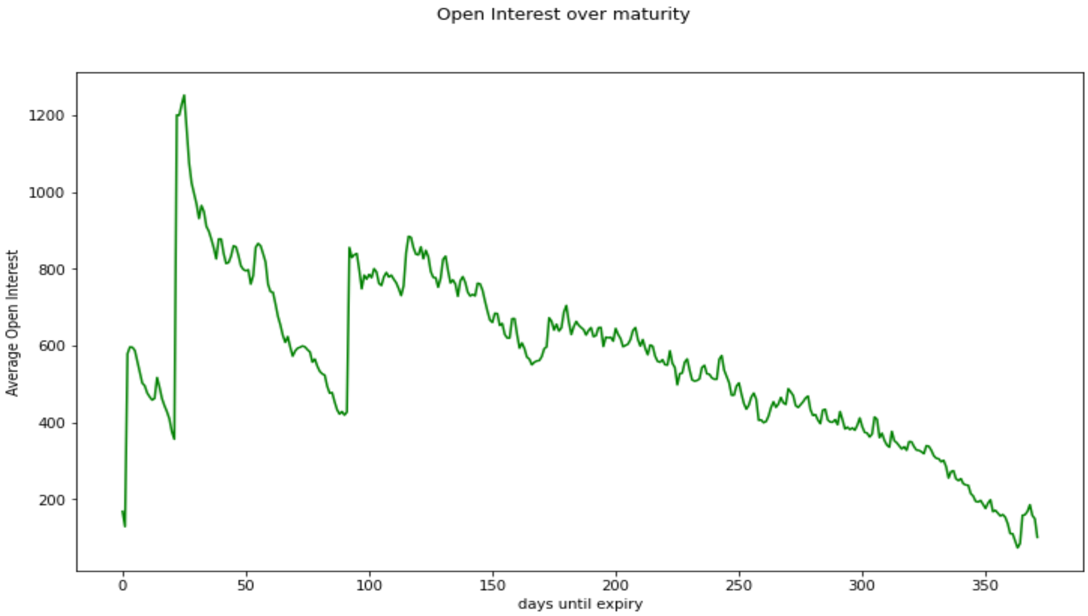

Figure 20: Open Interest average over maturity

3.2 Backtesting Algorithm

In the first version of the backtesting algorithm, I sell every last Friday of the month at 08:00 a straddle or a strangle with maturity at exactly one month and hold it until expiration. At this point, I disregard the margin requirements and assume I always have enough capital to open and maintain the positions. I choose options from the filtered dataset by selecting the contracts so that they have deltas close to the desired delta (0.2, 0.3, 0.4, or 0.5) and so that the resulting position has the lowest possible delta and gamma since to monetise the volatility risk premium, I need to build positions as delta neutral as possible. I then proceed to plot the performance of individual positions and the entire portfolio over time in terms of hourly P&L $(initial\ premium\ -\ current\ mark\ price)$ expressed in both BTC and USD. Compounding the portfolio hourly P&L I get the equity curve, i.e. the P&L accumulated over time.

I decided to use the mark price in the P&L calculation for reasons of interpretability of the results. It would be more realistic to use the ask price, as this is the price at which it is possible to close a short position at a given time. This conservative approach, however, causes problems when the liquidity of options decreases during periods of high volatility, such as in May 2021, when BTC lost 35% of its value and the 30-day annualised volatility exceeded 120%. In those days many options with deltas between 0.1 and 0.9 in absolute terms suddenly became far out-of-the-money

and therefore illiquid. As a result, the bid and ask prices of these options no longer quoted by market makers changed sharply. *Figure 21* shows the price of BTC and its 30-day volatility in May 2021. It also shows the bid and ask prices of a call option between the 16th and 19th of May 2021. It can be seen how at 4 am on 17th May the bid-ask spread spiked to over $1000 and then recovered in the following hours. This would result in a flash crash which would undermine my test.
The mark price, on the other hand, is the current value of the option as calculated by the Deribit risk management system as $Mark\ Price = Deribit\ Index + 30\ seconds\ EMA\ (Fair\ Price — Deribit\ Index)$ As the mark price is calculated with a dampener it follows the orderbook, which makes it a more reliable measure of the value of an option. Deribit uses it in the calculation of the P&L and to check any open position for the risk of a possible liquidation. To avoid outliers, I also decided to use the mark price in my calculations and displays.

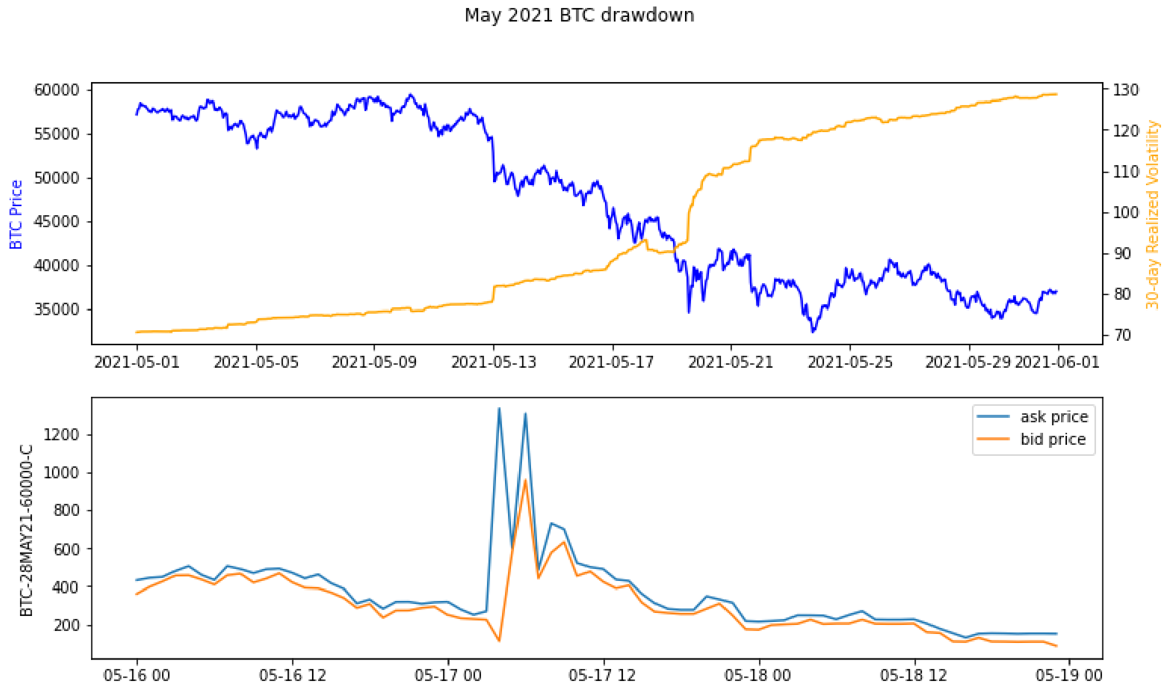

Figure 21: The collapse in the BTC price in May 2021 and option prices

Figure 22 shows the equity curves for different values of the delta of the options sold. For each curve, the final value is shown, i.e. the cumulative P&L at the end of the test. Strategies with higher deltas (0.4 and 0.5) performed better than those with lower deltas (0.3 and 0.2). This may be caused by the fact that selling Strangles with deltas close to 0.5 is a strategy that earns a high premium on average since at-the-money options are more expensive. This strategy incurs more frequent (30.4% of hourly P&Ls are negative) and larger losses (the maximum hourly drawdown is $22.9k on

19th May 2021) as the probability of an at-the-money option becoming in-the-money is high (theoretically 50%). The high average premium, however, allows for faster recovery from huge drawdowns, such as the one in February 2021 or the one in May 2021, and for better returns in the long run. On the other hand, selling Strangles with a delta close to 0.2 turns out to be a strategy that earns a low premium on average, as out-of-the-money options are less expensive. This strategy incurs less frequent (24.6% of hourly P&Ls are negative) and smaller losses (the maximum hourly drawdown is $20k on 19th May 2021) because the distance between strike prices is greater and out-of-the-money options are less likely to become in-the-money (theoretically 20% in this case). The low premium accumulated over time cannot cover the losses, which, although smaller and less frequent, significantly reduce the long-term return.

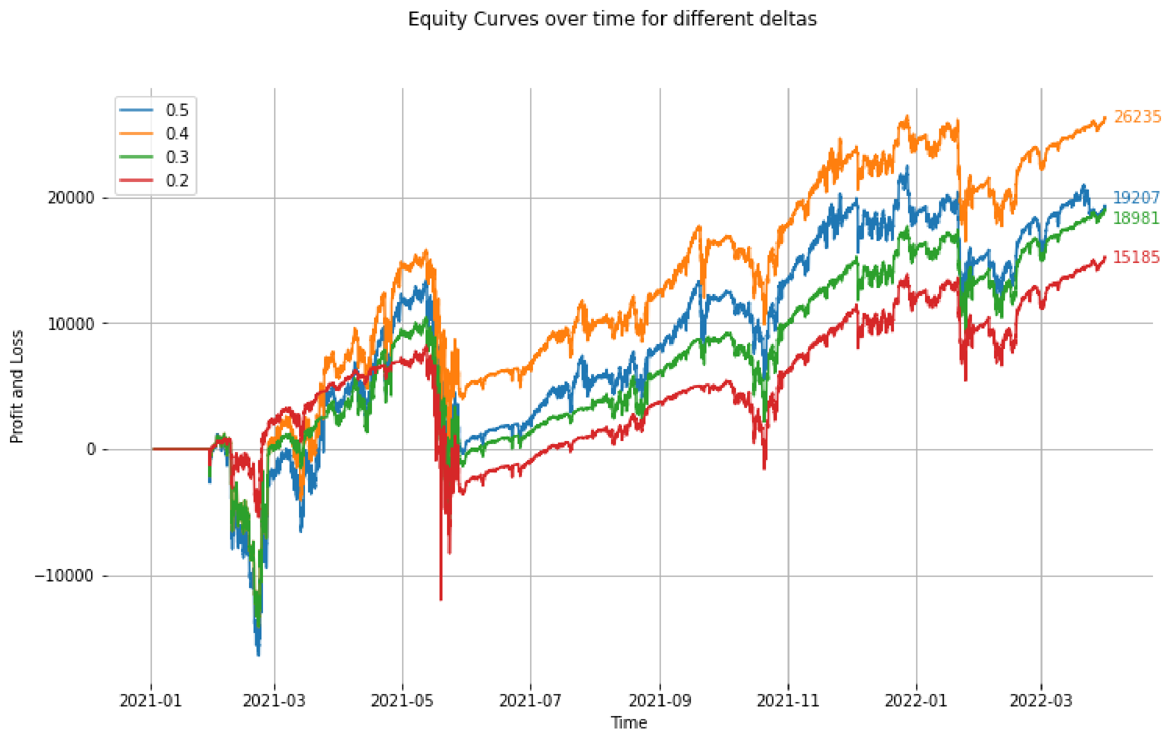

Figure 22: Equity curves and final P&L for different delta

3.3 Multiple Regression Analysis

In this chapter I focus on fully understanding the effect of different variables on strategy performance by creating multiple regression models based on individual trades. For this purpose, I construct a new backtesting algorithm in which I sell every day at 08:00 a Straddle or a Strangle with a given maturity and delta and hold it until expiration. For each position, I track the hourly Sharpe Ratio, the P&L at expiration, and 10 other useful variables described in *Table 1*. I repeat the test for all delta values (0.2, 0.3, 0.4, and 0.5) and all 13 maturities available on Deribit. I then merge the results into a single dataframe. This dataframe only includes trades closed by 26th August 2021 (8659 instances). I perform the multiple regression analysis on this training dataframe to select a model that predicts the PnL of a position given the variables in *Table 1* (except index, date, and Sharpe Ratio). In the next chapter I u,se the test set, which includes the data from 27th August 2021 the (last Friday of the month) to 30th March 2022 (about 50% of the option chain data), to observe the effect on the equity curve of using the model found above as a trading criterion: I open the position with the largest predicted P&L.

Table 1: Regressions variables

Variable	Description	Variable	Description
index	Index of the position	**ivr**	Implied volatility rank at the opening.
date	Opening date of the position	**ivp**	Implied volatility percentile at the opening.
pos_delta	Call delta + put delta. It is the delta of the position.	**init_marg**	Initial margin required to open the position.
mean_delta	Mean of absolute call and put delta. Average delta of the legs	**sharpe_ratio**	Hourly Sharpe Ratio of the position
pos_gamma	Call gamma + put gamma. It is the gamma of the position.	**premia**	The premium collected at the opening.
maturity	Days until the expiry of options.	**realized_pnl**	Premium – closing price.
bvin	BTC volatility index level at the opening.	**rv_30**	BTC price 30-day realized volatility

The calculation of the Sharpe Ratio for an option selling strategy is not straightforward as in the case of long positions. When a long position in options is opened, the returns are calculated based on the current value of the position and the capital invested, i.e. the premium paid at the opening of the trade, which also coincides with the maximum loss that can be incurred. In the case of short positions in options, on the other hand, no capital is invested when the trade is opened, but the premium is collected. It is therefore difficult to identify a method for calculating the returns of these strategies, especially since the possible loss, and thus the capital at risk is theoretically infinite. One possible approach is to base calculations on margin requirements, as explained in (Murray, 2012). Exchanges such as Deribit require a certain amount of liquidity at the opening of a short position (initial margin) and throughout its duration (maintenance margin). This amount of capital, which is locked in and therefore cannot be used until the trade is closed, is calculated to be sufficient to cover any reasonably probable loss the position may incur at expiration. The margin is not fixed, but changes over time as the value of the position changes. If the maintenance margin at any point exceeds the trading account's liquidity, a margin call occurs: the exchange liquidates (closes) part or all of the position until the margin requirement is met again. Since the margin is locked in, we can consider it as the capital 'invested' in the position. I therefore calculate the hourly returns as $hourly\ P\&L/current\ margin$. Since the margin is not constant over time, the return increases both as the P&L increases (as in long positions) and as the maintenance margin decreases. This results in high variability of returns and extreme Sharpe Ratios. This is why I use the realised P&L in the regression analyses as a response variable instead of the average return or Sharpe Ratio. *Figure 23* shows the distribution of hourly returns and hourly P&L for a strategy selling Straddles with one-month maturity as a meaningful example. The distribution of hourly returns is positively skewed and is platykurtic (negative excess kurtosis w.r.t. a value of 3 for the normal distribution). The distribution of hourly P&Ls, on the other hand, is negative skewed and leptokurtic (positive excess kurtosis). It is evident that in this case the rare extreme positive returns are not caused by equally large P&Ls, but by a sharp drop in the maintenance margin. As proof of the above, the annualised Sharpe Ratio is 35.15 (the hourly Sharpe is 0.38), which is an outlier value considering that the current 12-month Sharpe Ratio of the S&P500 is 1.10[11].

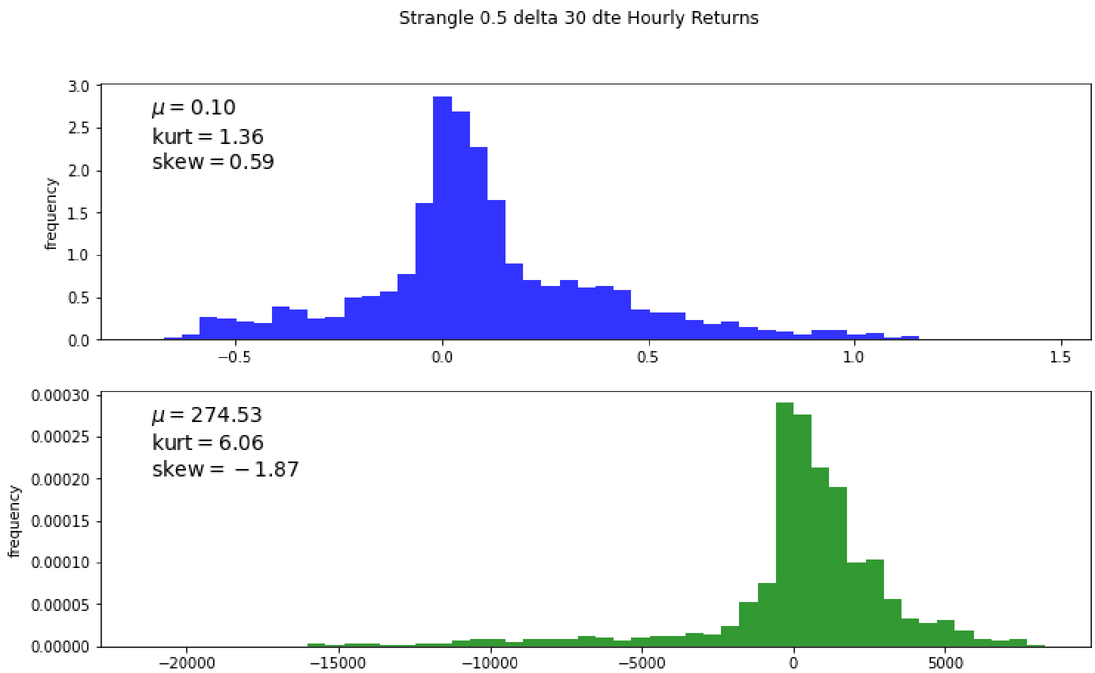

Figure 23: Hourly returns and P&L of selling 30-day straddles every month holding until expiry

[11] According to portfolioslab.com data on June 22, 2023

After the important clarifications concerning the definition of margin, the calculation of returns and the choice of the realised P&L as the response variable instead of the commonly used Sharpe Ratio, I proceed with the analysis of the correlation between the variables. My objective is to check for multicollinearity. Multicollinearity is the presence of high intercorrelations between two or more independent variables. It can lead to misleading results when trying to determine which regressors can be used most effectively to predict the response variable. In particular, it can inflate the variance of regression coefficient estimates making them non-significant. The use of highly correlated predictors, therefore, leads to less reliable inferences. To determine the presence of multicollinearity between the variables under consideration, I begin by analysing the correlation matrix in *Figure 24*.

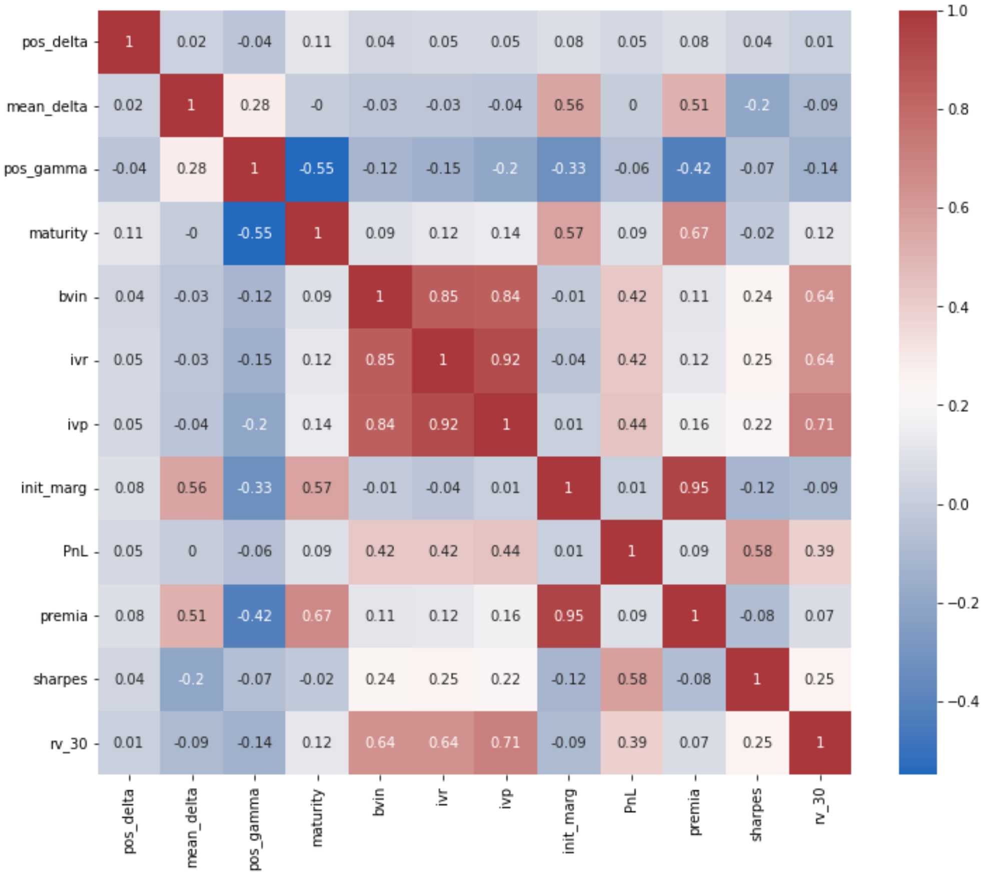

Figure 24: Correlation matrix of considered variables

Here are the most significant results:

- bvin, ivr, and ivp are positively correlated because all three are indicators of implied volatility. Their value increases as IV increases.

- Historical volatility is positively correlated with implied volatility as if the past performance of the BTC price influences expectations of future volatility.
- The premium received coincides with the sum of the bid prices of the options at the opening of the positions. It is positively related to margin, this being a function of option price, mean delta and maturity because options with higher delta and longer maturity are more expensive. In general, all these variables are highly positively correlated with each other.
- The gamma is negatively correlated with all the other variables except the mean delta of the options because the price of strangles with shorter maturities and closer legs is more volatile.

From the values obtained in the correlation matrix, I recognise the possibility of multicollinearity between the regressors. The Variance Inflation Factor (VIF) provides a useful measure of the degree of multicollinearity among the independent variables by measuring how much the variance of a regression coefficient is inflated by the correlation between the regressor and the other variables. Below are the formula and interpretation of the VIF.

$$VIF_i = \frac{1}{1 - R_i^2}$$

where:

$R_i^2 = unadjusted\ R^2\ for\ regressing\ i^{th}\ regressor\ on\ the\ remaining\ ones$

In general:

$VIF\ equal\ to\ 1 = variables\ are\ not\ correlated$

$VIF\ between\ 1\ and\ 5 = variables\ are\ moderately\ correlated$

$VIF\ greater\ than\ 5 = variables\ are\ highly\ correlated$

Below are the VIF values for all independent variables except opening date and sharpes, which I do not use in multiple regressions.

pos_delta	mean_delta	pos_gamma	maturity	bvin
1.05	2.36	1.95	2.49	2.40
ivr	**ivp**	**init_marg**	**premia**	**rv_30**
8.97	8.42	17.14	20.00	2.60

4 out of 10 variables are highly correlated. Having ascertained the existence of multicollinearity between the variables, I decide to proceed with multiple regression analysis by creating and comparing 4 models:

1. An OLS regression model in which I consider all variables as regressors, ignoring multicollinearity, and the P&L as the response variable.
2. An OLS regression model in which I do not consider the highly correlated independent variables, i.e. with VIF greater than 5, and use the P&L as response variable.
3. A LASSO regression model in which all variables are included, but in which multicollinearity is taken into account when calculating the coefficients.
4. An XGBoost regression model, an implementation of gradient boosting decision trees that is by nature immune to multicollinearity.

Table 2: Regression Analysis Results

	Dependent variable: PnL			
	OLS	OLS (VIF < 5)	LASSO	XGBOOST
Observations	8659	8659	8659	8659
N. of Predictors	10	6	10	10
Multiple R^2	0.227	0.211		
Adjusted R^2	0.226	0.210		
10-fold CV RMSE	4342.59	4385.46		
Validation Set RMSE			4394.86	599.96

It is evident that excluding variables with high VIF in the OLS linear regression or implementing a LASSO linear regression model does not improve performance in terms of validation root mean squared error (RMSE). On the contrary, despite the presence of multicollinearity, the OLS regression model has the lowest validation RMSE among the linear models. In general, the eXtreme Gradient Boosting model performs significantly better than the other models in terms of test error estimation.

3.4 Best Model Implementation

In this last section, I implement the eXtreme Gradient Boosting model in the sale of Straddles and Strangles for VRP monetisation. This model provides the best performance on the training set, which includes trades closed by 26th August 2021. I now want to test its performance on the test set, which includes data from 27th August 2021 to 30th March 2022. *Figure 25* compares the P&Ls predicted by the XGB model, on the horizontal axis, with the actual P&Ls, on the vertical axis. The RSME on this new unseen dataset is 7095.33 which, as is usually the case, is much greater than the validation-set test error estimate shown in *Table 2*. If the prediction error were zero, the points in the graph would be distributed on the 45-degree line. It is noticeable, however, that the model tends to predict P&Ls significantly lower than the actual ones on average. Most of the points lie on the left-hand side of the 45-degree line. The higher RMSE indicates the possibility of overfitting on the training set, although such conservative predictions could be advantageous from a trading perspective.

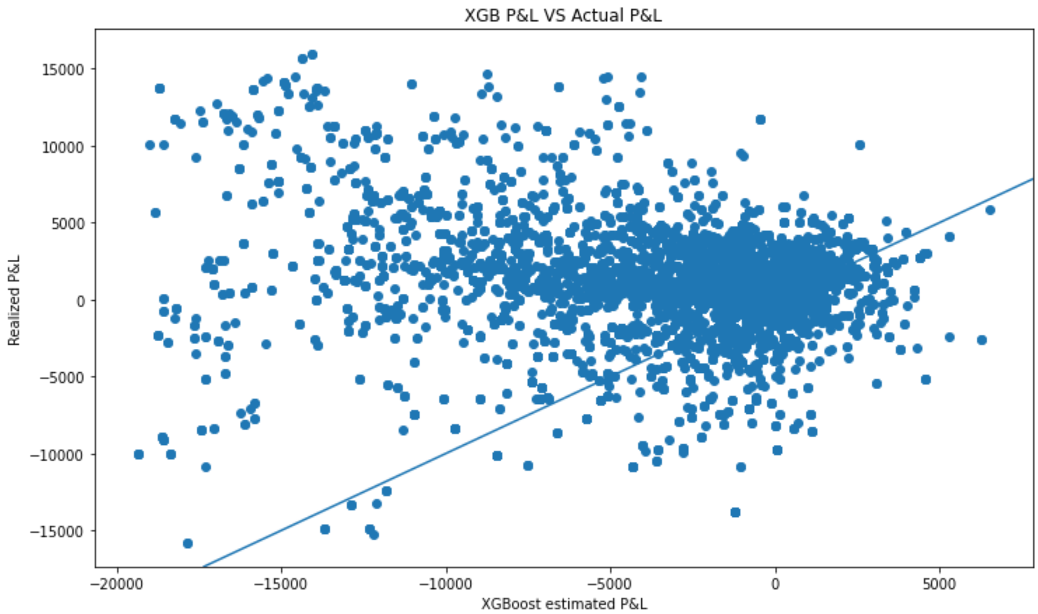

Figure 25: XGBoost model forecasted P&L and actual realized P&L in the test set

The XGBoost model has the disadvantage that it is not as easily interpretable as linear models. In finance, however, the interpretability of models is important for decision-making. To base a trading strategy on the predictions offered by this black-box model, it is necessary to understand how it works. The most advanced method for interpreting the results of tree-based (non-linear) models is to use SHAP (Shapley Additive exPlanations) values. This model-agnostic technique provides a global understanding of feature importance by quantifying each feature's contribution to predictions. *Figure 26* shows the SHAP values of the features in the XGBoost model. We can see that the value of the bvin index at the opening and the maturity of the contracts sold are the variables that contribute most to the model's predictions, while the delta of the position and the delta of the contracts matter little.

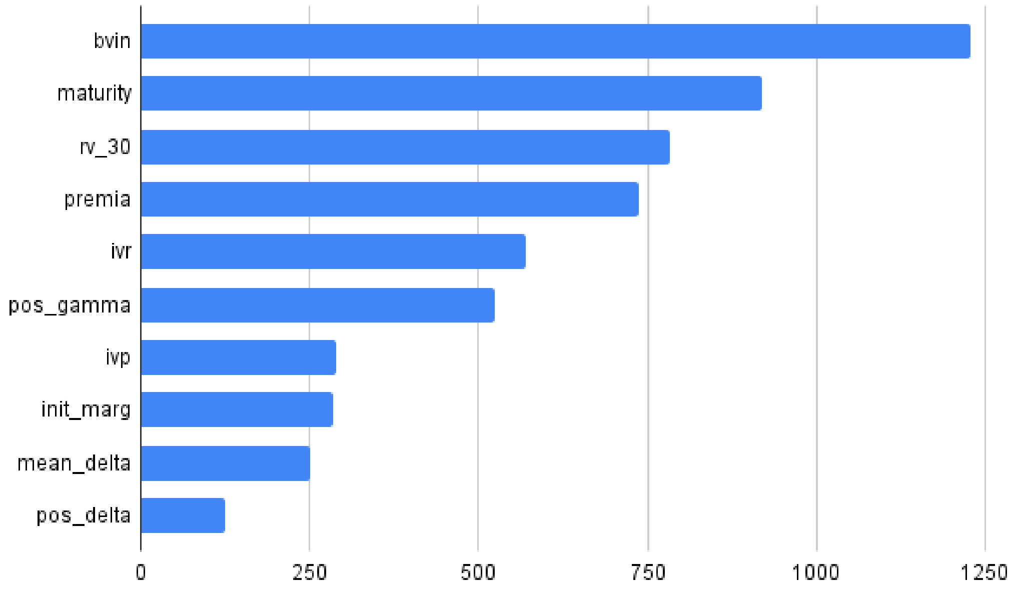

Figure 26: SHAP values of XGBoost model features

I now proceed to implement the XGBoost regression model in the sale of Strangles and Straddles. I train the model on the entire training set (trades until 26th August 2021) and use it to make trading decisions. In *Figure 27* I compare the equity curves of two strategies starting from 27th August 2021. The first strategy consists of selling a 30-day Strangle with a delta of 0.4 every last Friday of the month. This strategy proved to be the best strategy in Chapter 3.2, as shown in *Figure 22*. The second strategy consists of selling every last Friday of the month the Straddle/Strangle with the highest predicted P&L, considering all available deltas and maturities. One can see how using the XGBoost model decreases the final P&L by approximately 70%. During the 8 months in the test set, the model sells Strangles and Straddles with a maturity of less than 20 days and thus relatively cheap. This choice brought poor results, probably because the little premium collected was not sufficient to cover the losses. The poor performance of the trading strategy as well as the high RMSE in the test set could be due to the overfitting of the model, especially considering the low RMSE in the training set. It must be specified that training a model on only 8 months of option data is risky because temporary factors could compromise its robustness by creating noise in the data. To build a more robust model, therefore, I would have to collect more recent data and repeat both the regression analysis and the implementation in trading.

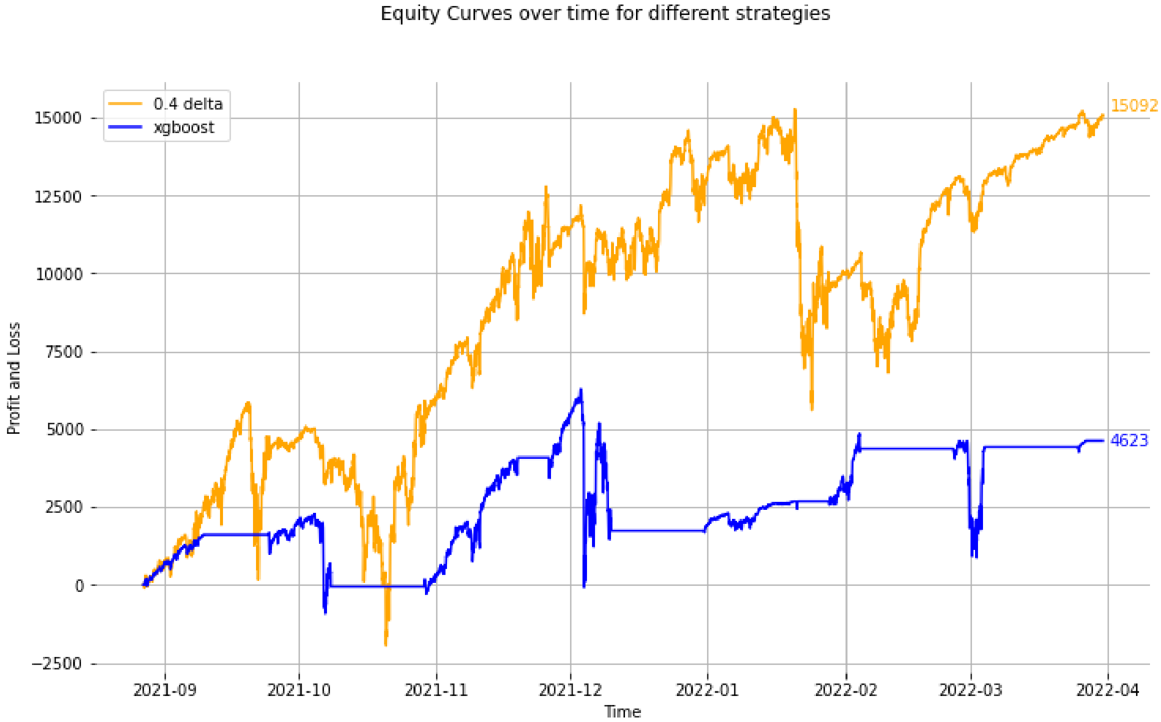

Figure 27: Equity curves and final P&L for 30-day 0.4 delta and XGBoost

Conclusion

In this thesis project, I presented the concepts of cryptocurrencies, options contracts, crypto-options, implied volatility, volatility risk premium, and delta-neutral option strategies and then carried out an in-depth analysis of the options market on BTC. This is a very young and fast-developing market and as such suffers from low liquidity problems. Through a self-made backtesting algorithm and regression analysis, I analysed the possibility of using options trading strategies such as Straddles and Strangles to monetise the negative VRP in this market. These strategies reported positive results even though they were accompanied by high P&L volatility and huge drawdowns. The regression model I selected provided poor out-of-sample results and its use in trading did not increase P&L or reduce volatility. This could mainly be due to the limited availability of option chain data (15 months). My plans include collecting a larger dataset and creating a new model to select the best option selling strategy to monetise the volatility risk premium.

Bibliography

Aleti, M. (Forthcoming). Bitcoin Spot and Futures Market Microstructure. *Journal of Futures Markets*.

Alexander, I. (2020). The Bitcoin VIX and its Variance Risk Premium. *Journal of Alternative Investments*.

Black, S. (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*.

Bollerslev, T. (2009). Expected Stock Returns and Variance Risk. *Oxford University Press*.

Carr, W. (2008). Variance Risk Premiums. *The Review of Financial Studies*.

Goetzmann, K. (2017). Negative bubbles: what happens after a crash. *National Bureau of Economic Research*.

Kozhan, N. S. (2013). The Skew Risk Premium in the Equity Index Market. *Review of Financial Studies*.

Murray. (2012). A Margin Requirement Based Return Calculation for Portfolios of Short Option Positions.

Ringraziamenti

Con il completamento della tesi si chiude il mio corso di laure in Economia e Data Science. Si concludono tre anni ricchi di cambiamenti ed esperienze che mi hanno arricchito di conoscenza e ricordi.
Ringrazio l'Università di Torino per l'educazione di qualità che mi ha fornito e di avermi aperto nuove strade verso il futuro che desidero.
Ringrazio il Collegio Carlo Alberto per avermi offerto l'opportunità di partecipare all'Allievi Honors Program, grazie a cui sono entrato in contatto con un ambiente accademico all'avanguardia e ho potuto stringere importanti amicizie.
Ringrazio Lorenzo Schoenleber per aver creduto in me tanto da condividere un progetto di ricerca accademica, per i preziosi consigli, per la disponibilità e per l'aiuto nella stesura di questa tesi.
Ringrazio la mia famiglia, i miei amici e Catarina per aver condiviso con me questi anni rendendoli così speciali.