U.PORTO

FEP FACULDADE DE ECONOMIA
UNIVERSIDADE DO PORTO

Mispricing of Bitcoin and Ethereum contracts under the cost-of-carry framework:

Evidence from FTX

Rodrigo Encarnação de Albuquerque

Dissertation

Master in Finance

Supervised by
Júlio Fernando Seara Sequeira da Mota Lobão, PhD

2021/2022

Acknowledgements

Firstly, I want to thank my mentor, Julio Lobão, for the support and guidance provided during the realization of this dissertation. The high standards demanded during the all the process are reflected on the quality of the document. I want to mention all the other professors that taught me during this master, increasing my knowledge on a broader range of topics in the field of Finance.

I want to express my gratitude to my family, especially to the ones who considered my formal education a priority. After numerous years of investment, I expect to demonstrate them how much human capital was created, first with this dissertation, then in the labour market.

Last but not the least, I want to acknowledge the importance of my friends. During my academic path, they were responsible for transforming this journey in an adventure of personal growth, teaching me invaluable lessons, and growing alongside me. I truly believe that I am an average of them. They are the reason I consider myself hedged on happiness, enabling me to grow and take new risks.

Abstract

The objective of this dissertation is to provide evidence on the existence, or non-existence of mispricing on BTC (Bitcoin) and ETH (Ethereum) future contracts, more specifically on the FTX Exchange, one of the most prominent retail-oriented exchanges in the cryptocurrency scene.

The motivation for such subject arises from personal interaction with the exchange in the middle of the chaotic cryptocurrency market, when returns yielded by simple arbitrage strategies such as cash-and-carry were persistently above what is expected according to the cost-of-carry framework, however dwarfed in comparison with delta positive returns on respective underlying assets.

The investigation process will relapse on using the theoretical framework on future contracts valuation, known as cost-of-carry, and comparing obtained values with observed values. The period analyzed starts on July 2019 and ends in March 2022, comprising 980 days of pricing data on contracts with BTC and ETH as underlying. In an attempt to justify the values obtained on mispricing, this dissertation uses exogenous variables, such as sentiment, to explain and observe how they affect or correlate with mispricing.

The results show that mispricing was persistent during the studied period, but only meaningful during specific sub-periods. During those sub-periods, when sentiment was high, the expected reward on arbitrage ascended to punctual periodic rates of 11.39% for front-quarter contracts (non-annualized rate for TTM under 92 days), and 28.66% on longer-term contracts (non-annualized rate for TTM under 352 days). Moving inversely to sentiment was the net position of leveraged funds on the regulated futures market, being most short, when the mispricing was the highest, specifically for BTC Contracts.

In the end, the choice of using a retail-oriented exchange was to widen the practical utility of this dissertation to all types of investors. Those returns were available to everyone curious enough to exploit the inefficiencies created by the participants of the market in times of great exuberance, questioning, once again, how efficient it is. Previously referred returns are not the predicted by the cost-of-carry model, being much higher during the entirety of specific quarters. Such persistent mispricing is against the efficient market hypothesis.

Resumo

Esta dissertação debruça-se sobre a verificação da existência de *'mispricing'* no mercado de futuros sobre Bitcoin e Ethereum, na plataforma FTX, uma das principais e mais legitimas operadoras no mercado de derivados sobre cripto moedas.

A motivação por trás de tal estudo recai sobre a utilização desta plataforma durante o mercado caótico em que decorreu o período de analise, durante o qual, os retornos associados a estratégias de arbitragem eram persistentemente superiores ao expectável pelos modelos teóricos de avaliação de futuros. No entanto, estes retornos eram comparativamente pequenos quando comparados com estratégias simples e unidirecionais.

O processo de investigação consiste na avaliação dos contractos de futuros usando o modelo cost-of-carry, comparando posteriormente a avaliação teórica com o preço observado no mercado. O período analisado é referente a 980 observações diárias desde Julho 2019 até Março 2022, sobre os contractos baseados em dois ativos, BTC e ETH. Outras variáveis exógenas aos dados referentes a preços, como um índice de sentimento, serão usadas como variáveis explicativas.

Os resultados apontam para um *'mispricing'* persistente durante o período estudado, no entanto, significativo apenas em algumas porções da mesma, quando o índice de sentimento registava valores que implicam otimismo. Durante esses tais segmentos de tempo, o *'mispricing'* chegou a taxas periódicas de 11.39% para os contractos com expiração mais próxima da data da observação, e 28.66% considerando todos os contractos. Contrariamente ao índice de sentimento, a posição dos *levereged funds* nestes contratos era mais negativa, quanto maior fosse o *'mispricing'*, tal como esperado, no caso da BTC.

Concluindo, esta dissertação e os resultados obtidos têm relevância para todo o tipo de investidores, pois os dados foram baseados numa plataforma agnóstica à classificação dos mesmos. Os retornos podem ser obtidos por todos os que estiverem dispostos a explorar as ineficiências de mercado, especialmente durante tempos de otimismo desmedido. Os retornos mencionados são muito maiores que o previsto pelo modelo cost-of-carry, durante períodos longos. A existência persistente de contractos sobreavaliados, que permitem a exploração de lucros de arbitragem, vai contra o previsto pela teoria do mercado eficiente.

Index

List of Tables

List of Figures

1. Introduction

Cost of carry is one of the standard methods to value future contracts, adding or reducing variables depending on the underlying. The reliability of the method relies on the combination of financial operations that allow the arbitrageur to earn the carry, it is, the difference between the value of asset now, and its' value on the future. In order to execute the cash and carry strategy, an arbitrageur should borrow cash to buy the underlying asset and short sell the future contract. If the difference between the price of the contract and the spot price is higher than interest costs and asset-specific costs, the arbitrageur will make a profit. Any modifications added to this framework, must be a cost, implicit or explicit, supported during the process. Modifications include dividend adjustment for index futures (Butterworth & Holmes, 2000), and among others, transportation and storage costs for several commodities such as oil (Alizadeh & Nomikos, 2004). This arbitrage strategy, as any other strategy, should not generate predictable systematic excess returns, otherwise, it would be threat to the random walk theory of Fama (1965).

Previous literature shows examples where the returns of executing the cash-and-carry, and the reverse version of the strategy, in some assets yield positive excess returns systematically, like in oil arbitrage (Alizadeh & Nomikos, 2004), but in other, like indexes, yields are usually negative, with mispricing accounting for -0.666 p.p. for Greek Index ATHEX-20 (Fassas, 2010) and -0.02 p.p. for the Spanish IBEX-35 (Verdasco, 2017). In the first case it is logical to expect that not all market participants have the skills or knowledge to rent an oil tanker to transport Brent across the Atlantic, to deliver it on the Gulf Coast, load it on a pipeline to Cushing, Oklahoma, in order to perform the carry trade. However, with financial assets, the process is easier, and that could be one of the reasons why the cost-of-carry profits on financial assets are not on the magnitude of real assets. It is impulsive to correlate the difficulty of executing a trade with returns. However, one novel asset class, cryptocurrencies, seem to do not follow that logic. Cryptocurrencies are by far the asset class that requires less capital to trade, both on the underlying spot and future contracts. One would expect a frictionless market to be more efficient, however, mispricing in BTC and ETH futures oscillate between 17.62% discounts to 28.66% premium during the considered period, for longer-term contracts. Both rates are absolute with maturities ranging between 352 days and 98 days, thus not directly comparable.

This dissertation uses the traditional methods to measure mispricing, namely the method used by Butterworth & Holmes (2000), used as reference for much of the cost-of-carry literature. Going further and following what appears to make sense in the cryptocurrency market, exuberance seems to relate to higher mispricing values, phenomenon observed previously on other markets, as demonstrated by Stambaugh et al. (2012). Applying two more variables, underlying spot price and net position of leveraged funds, the final objective of the dissertation is based on understanding if there is mispricing during the studied period (from July 2019 until March 2022), and if there are variables capable of explaining it.

Results are not the same for BTC and ETH contracts. For the first set of contracts, a negative relationship with the net position of leveraged traders was expected and obtained on the regressions used. A positive relationship between contract mispricing is observed on the remaining variables, sentiment and price. For ETH all variables were found to have a positive relationship with mispricing, with some reservations made on the robustness testing.

The second chapter reviews the relevant literature on the efficient market, the arbitrage in several asset classes, and the counterparty risk. Next, the third chapter describes the data used in the dissertation, starting by the methods of data collection, followed by an analysis to the explainable variable, the explanatory variables, as well as adjacent variables that are important for the practical consistency of the strategy in real life. The fourth chapter explains how those variables will be used to justify mispricing, by providing equations for the purpose. The product will be the empirical results presented in the fifth chapter. There, statistical results are interpreted and transposed to financial terms, analysing if results make sense, and getting answers to our hypothesis. In order to test the robustness of the results, several methods are employed in chapter six, comparing conclusions across time and with higher granularity. The last chapter provides the product of the previous ones, by summing up the results, exposing the limitations of the dissertation and finally, proposes new related research themes.

2. Literature Review

Arbitrage could only exist if noise traders participate in the market, and differences in beliefs and consequently information are the only reason to trade (Black, 1986). In the presence of investors with different access to information, and different interpretations of it, prices will move, and if the so-called noise traders persist, in this case, by using future contracts to express their views, deviations from the fair value could occur. One could use a future contract to hedge its portfolio or just as a speculation vehicle. In both cases, shall traders know the theoretical price of the contract, deviations should be limited to a certain extent, as in theory, there other ways to hedge or speculate, that would be more attractive when considering the implicit costs of such contracts. But it is not dogmatic that every trader knows the theory behind contract pricing, perhaps not even having an opinion about the variables that are believed to define its fair value.

In the following subchapter, implications of such verification on the concept of market efficiency are exposed.

2.1. Arbitrage and Market Efficiency

In order to create coherent hypothesis on the existence of arbitrage opportunities in cryptocurrency futures markets, some background on the efficient market hypothesis is needed. One of the main features of efficient markets is the competition between participants to buy securities below their intrinsic value, and to sell them at a higher price than their intrinsic value. As information can be interpreted in different ways, investors will have heterogeneous expectations regarding the value of a specific security. In the end, the market price will be the equilibrium point between all the disagreements, and the most reliable estimate of intrinsic value. According to the random walk theory, discrepancies between the observed price and intrinsic value could exist randomly, but never persist systematically. In the last case, arbitrageurs would intervene in any systematic difference, leading to a convergence to intrinsic value. An efficient market considers instantaneous adjustment regarding new information, meaning that price changes across time should be independent and only react to new information. If the contrary happens, and series of prices are not independent, past prices could be used to predict future prices, adding predictability to the market. The independence rule will also mean that market timing cannot generate any excess returns, and by consequence, a strategy of buy-and-hold should, on average, perform equally to other market-timing strategies before any fees. Technical analysis, that relies on past prices to identify patterns on charts, should be irrelevant under the independence assumption. What it does not imply is that it is not possible for someone to beat the market by being better at exploiting big differences between the intrinsic values and observed prices. In the presence of a great number of excellent investors, their expectations on the intrinsic value of a security will be reflected in the price in accordance with their actions, leading to the "correction" of the price. The advantage to arbitrageurs should come from the correct interpretation of new information or insights, that when incorporated, will reflect the new efficient representation of fundamental value. Considering that only this subset of investors is able to exploit discrepancies, the value added by fundamental analysis to other investors is expected to be zero, meaning that a random selection of stocks should, on average, yield the same returns of stocks selected by non-sophisticated investors (E. F. Fama, 1965).

In addition to the EMH, it is important to consider that there could be some limits to arbitrage preventing the sophisticated investors of acting as they should to correct the price. The concept of arbitrage is not singular. Arbitrage could be referred as the action to correct prices by acting correctly on all the available information. However, it is often defined as the act of purchasing or selling a security in two different markets with the objective of profiting from differences in prices. In the first case, arbitrage will ensure fundamental efficiency, while in the second case, it will ensure relative efficiency. It is also important to identify the arbitrageurs and analyse how do they operate. Considering the activity as a sophisticated one, implying that only a small number of market participants engage on such strategies, it should be implicit that those professionals operate with external funds, otherwise it would not impact the market. Investors utility function will depend on risk and returns, and their allocation choice would be based, at least to a certain extent, to the past performance of a specific arbitrageur. However, even assuming no fundamental risk in the long-term, a strategy could yield negative returns in the short-term. The same reason that will lead an arbitrageur to identify a discrepancy between the observed price and fundamental value, will be responsible for its persistence in the future. If the discrepancy fails to produce returns in the short-term, or produces even negative returns temporarily, some investors could ask for their money, creating the need for the arbitrageur to close its positions, by taking the contrary position and worsening the gap between observed and theoretical fundamental value. In the end, specialized performance-based arbitrage will have limited impact in price correction, at least in high volatility environments (Shleifer & Vishny, 1997).

Accordingly, an important limit to arbitrage would be one of the reasons linked to the short-term unpredictability, denominated as noise trader risk. The risk arises from the interaction of unsophisticated investors who falsely think that they have superior information in the stock market. Those traders will buy based on arguments that will not impact the fundamentals of a security or are already included in the current market price because they are obsolete. Because the information they are incorporating in the price is not fundamentally significant, noise traders will drive the price further away from its fundamental value. As the mispricing becomes larger and larger, the opportunity for arbitrageurs will, in theory, increase as well. However, it works as a paradox, because on one side the expected return for arbitrageurs, is increasing, but on the other side, they do not know if the mispricing could continue to increase. In the end, these sophisticated investors could turn their attention to exploit noise traders themselves instead of acting in order to correct mispricing (de Long et al., 1990). As the value of the arbitrage portfolio could change even in the presence of no fundamental risk, arbitrageurs will have to bet on the time of the correction. This is denominated as synchronization risk, as the profits of betting on the elimination of mispricing will be dependent on other arbitrageurs acting on the same opportunity, implying that noise traders' volume is unable to absorb arbitrageur volume. One cause for desynchronization is the different opinions regarding the phase of price correction when they learn about the opportunity. When someone happens to learn about a gap, the arbitrageur himself could choose to wait in order to maximize the utility of its trade, avoiding holding costs until they believe that the other arbitrageurs will act on the mispricing. However, it is naïve to think that everyone will know about the mispricing in a given moment, and even if it does happen, that everyone will act on it. Eventually, when incentives are high enough, and the absorption power of behavioural traders is not enough to break the convergence to fundamental value, the price should correct, and rightfully reflect the information (Abreu & Brunnermeier, 2002).

Such limitations could explain in part the existence of market anomalies. Several of them could be found more easily during periods of positive sentiment, in part because of the limited availability and possibilities to correct overpriced securities. Considering such situation, Stambaugh et al. (2012) tested the hypothesis by analysing the returns of long-short strategies that would explore those anomalies following periods of high sentiment and low sentiment. After optimistic periods, the overall long-short position yielded excess returns of 1.23%, compared to 0.31% excess returns after pessimistic periods.

It is expected for securities with the same fundamentals to trade at the same price, or in the case of not-proportional claims, a price that implies the same fundamentals, as the value of a security is the actual value of future cash-flows. Sophisticated investors, in the presence of an apparent mispricing, bet on the convergence of the so-called twin-shares. However, this convergence might not be profitable at all, as it was found that mispricing between similar securities is systematic. Some reasons such as taxation try to explain the differences between shares representing the same company, but in different countries (Rosenthal, 1990). This difference between twin shares should be considered when testing the Law of One Price. According to the Law, identical securities should have identical prices, otherwise there will be arbitrage opportunities to exploit. Theoretically, the Law is enforced by the incentive alignment provided by the existence of profits to the ones willing to bet on it. Apart from the example referred above, other anomalies could jeopardize the acceptance of the theory. For example, closed-end funds often trade at prices that are not equal to the value of the assets owned. In this case, we should bear in mind that owning all the underlying is not a perfect substitute of owning them in packaged way through a closed-end fund, as fees will have a small impact on the performance of both methods. Nevertheless, extreme premiums higher than 100% were observed in several cases, such as Taiwan Funds and German Funds, sometimes in the absence of the main excuse of the existence of capital movement restrictions. Such differences could never be justified by small differences in fees. Another example refers to the difference between the price of ADRs and their underlying, which represent the same claims on equity but trade at different prices. In the case of Infosys, that was listed in Bombay and in Nasdaq through an ADR, the premium reached 136%, and no arbitrage circuit was feasible as it was not possible for Americans to buy the underlying on its original exchange (Lamont & Thaler, 2003).

2.2. Index Fund Arbitrage

Several arbitrage strategies were tested in a variety of markets. Starting on index arbitrage, one sophisticated investor could identify mispricing between the current value of the index and its future value. The model usually used to price index futures is denominated as cost-of-carry. The arbitrage circuit start with borrowing at the risk-free rate, funds then used to buy the underlying security. By entering in a short position on the future contract, the arbitrageur will receive, in the end of the contract, the value at which it was sold in exchange for the underlying security (if physically settled). The profit of the strategy equals to the funds received net of interest costs. The name of the strategy is cash-and-carry, and it is designed to be executed when the future is overvalued relative to its' underlying. In the case of future contract undervaluation, the strategy used by arbitrageurs is called reverse cash-and carry.

There are several factors that seem to be associated with the efficiency of the market, when analysed through the lens of mispricing. According to Miller et al. (1994), mispricing is more of statistical elusion than an exploitable arbitrage opportunity, as one the main reasons for this mispricing on the basis between index futures and its' spot price was the trading frequency differences in some of the constituents of the index. If prices are not updated continuously on constituents as they are on the futures market, arbitrage opportunities might be just theoretical, because prices are not executable in practice. Execution risk is also considered in other studies, such as Bühler & Kempf (1995), where the authors adjust for that using a risk premium, that on top of transaction costs, will be compared with the mispricing between DAX and futures contracts on the index. In this paper, the cost-of-carry model was not effective in explaining the future contract price, as it systematically was undervalued. In addition, mispricing seems to have a positive relationship with the time to maturity, leading them to conclude that arbitrage activity is higher and faster in short-dated futures. Finally, the majority of arbitrage opportunities included taking short positions on the spot market (because of contract undervaluation), that could be subject to some restrictions and increase the holding costs.

Other less developed markets show the same pattern of undervaluation regarding the futures market. An analysis regarding arbitrage opportunities in the Greek Index ATHEX-20 was performed by comparing the mispricing with transaction costs. The future contracts were underpriced 83% of the time. However, as mentioned before, the mispricing is not a synonym of existence of arbitrage opportunities, and barriers to efficiency such as liquidity might be the reason for the possible illusion of mispricing (Fassas, 2010). Different conclusions are observed for other markets, such as in Spain. During the period between 2007 and 2015, the mean underpricing between the implied price by the cost-of-carry and the observed price for the IBEX-35 future contracts was -0.02%. The value was small enough to validate the use of the model for index future pricing. Furthermore, it was found that the model has worst compliance in the presence of volatility (Verdasco, 2017).

In more liquid markets, volatility is still one driver for the supply of mispricing. One study analysed the mispricing between the cash index and its' future contract on the most liquid market in the world, the S&P 500. In order to create a more practical approach to arbitrage, the author used an ETF as the underlying asset. By doing this, the complexity of replicating the whole index of approximately 500 different securities, ones more liquid than others, could be replaced by other issues such as differences to NAV. Nevertheless, by using the ETF, fewer transactions are needed, and its' liquidity is more suitable for execute the cash-and-carry strategy. Before fees, mispricing was found to be 1 index point, decreasing to 0.4 index points after transactions costs (Richie et al., 2008).

2.3. Commodity Arbitrage

Unlike index arbitrage, cash-and-carry in commodities is somewhat more complex and can include a set of non-financial transactions. For example, the cost of transporting crude from one continent to another is significant and requires hiring several services and intermediaries. It is expected that instruments that have more specialized and costly arbitrage circuits have more pronounced mispricing (Pontiff, 1996). This is not surprising, as mispricing is expected to be corrected by arbitrageurs, and hiring an oil tanker is not in the skill set neither within the budget of small or not oil-oriented traders.

The most active commodity in the world, crude oil, is not exempted from arbitrageurs. In fact, as this commodity is pumped all around the world, it creates a new source of possible mispricing, of geographical nature. Because oil is not equal in every part of the world, several names were attributed to each blend of oil. When someone takes a short position in a WTI future contract, that same party needs to deliver the contracted size in order to settle its' position, if it is not closed prior to expiration date. The complexity starts on the dynamics of delivering foreign blends. According to the gravity of each type of oil, a discount or premium will be attributed. For example, UK Brent had a $0.3 per barrel discount against WTI, to reflect its' lower quality, difference that need to be included in the cost of carry model for Brent. Another oil specific charge is transportation costs, as every contract has a specific delivery point, in the case of WTI it is Cushing, Oklahoma. It was found that freight rates were not connected with WTI spreads, implying that when oil prices change, and the freight rates stay relatively flat, arbitrage opportunities emerge. Once again, volatility weakens the compliance of the model (Alizadeh & Nomikos, 2004).

The dynamics of storage for metals are a good illustration of how the industrial use interacts with pure financial arbitrage. According to the cost of carry model, buyers will not buy cooper to hold it if the forward curve structure is in backwardation. Such market structure implies that futures with higher time to maturity would be priced below near-term future contracts. The argument comes from finding an explanation for holding a material that is expected to fall in price in the future, is available for that lower price in the future, however some companies choose to store it. The main explanation for this situation is indeed the existence of a convenience yield for companies that need copper on a regular basis for their operational activities. For example, buying small quantities of copper just to consume it on the arrival, would imply transportations cost for each shipment of the commodity. By storing it, even with a backwardation market structure, it could be justified if the operational savings are higher than the expected loss from the fall in price of the metal. So it is expected that maintaining a low levels of inventories to add operational flexibility provides positive utility. According to Larson (1994) storage happens in times of backwardation not only because of its' convenience, as additionally, copper held in inventories protect it's owner from unplanned volatility in prices, being compared to a call option. Another study regarding other metals, aluminium, zinc, lead, tin and nickel in addition to copper, analysed the relationship between volatility and inventories, and corroborated the strong relationship implied by the Theory of Storage (Geman & Smith, 2013).

2.4. Cryptocurrency Arbitrage

Research on cryptocurrency derivatives is still scarce, in part because of the age of the underlying and the age of functional Bitcoin future markets. The first regulated future contract was launched by CME on 18th December 2017. As literature on Cryptocurrency uses data from CME and CBOE to conduct efficiency studies, retail-oriented exchanges were left behind, even when they exhibit higher derivative-related volume.

The introduction of Bitcoin futures moved the currency closer to match the narrative of unit of account. The existence of future contracts enabled hedging strategies that allowed investors to not incur pricing-related risk when buying assets denominated in Bitcoin, one of the three characteristics of a currency. However, creation of a derivatives did not reduce the volatility of the underlying asset. During the period analysed, the price discovery occurred in spot markets rather than in future markets. According to Corbet et al. (2018) it goes in line with the argument that the agents operating in the futures market are in part noise traders.

A more recent study obtained different conclusions regarding causality on a more complete and up to date sample. Akyildirim et al. (2020) verified that there were catalysts caused by regulatory changes or illegal activity related with cryptocurrency that affected the activity in both spot and future markets, as engaging with the market in rocky times could cause permanent reputational damage. The same paper determines that price discovery happens in the spot market and is transmitted to the futures market in the first part of the sample, but as time passes, it starts to show the opposite causality, with price discovery happening in future markets, and being transmitted to the spot market. It happens, in part, because of changes in market structure, as institutions and more sophisticated investors join the market. It was also discovered that information will flow from CBOE to CME, with the justification of differences between both contracts. Such conclusion could be a start to explain any shifts of this causality when adding exchanges like FTX and Binance, as future contracts have small differences when compared between themselves and when compared with the traditional ones.

The use of cost of carry model to price cryptocurrency futures is at least questionable. The framework of the model adds specific variables to adapt to each type of underlying. For example, as explained above in commodity arbitrage, storage and transportation are included in the model, variables that have direct causality in the profit of the arbitrage strategy. In the case of cryptocurrencies, some variable we don't consider could explain what is undetected today by the model. Hu et al. (2020) questions if it is appropriate to use the universal variable of the model, the risk-free rate, as compounding factor, that if tested accordingly could render the cost of carry model obsolete to price cryptocurrency futures.

Arbitrage in the field of cryptocurrencies does not only include cash and carry. The most primal definition for arbitrage states two transactions in different markets for the same asset, where they trade at different prices. By buying cheap in one market and selling it for a higher price on the second market, a riskless profit is available for grabs. Makarov & Schoar (2018) constructed an index to measure the price premium of some exchanges when compared to the price of BTC on the United States. On average, in 2017, the index points to 1.05, meaning that it would be possible to buy $1 of BTC on the cheapest exchange an sell it immediately on the most expensive exchange for $1.05. Note that this is an average value, and during some periods, the index exhibited much higher values, such as an average of 15% for the premium between the United States and South Korea, verified between December 2017 and

January 2018. In the same periods, persistent premiums of 40% were observed. When compared to other regions, the average premium between the BTC price in the US and in Japan was 10%, while averaging a lower value of 3% when measured against Europe.

2.5. Counterparty Risk

Contrary to what is expected in an ecosystem born with the desire to eliminate the middleman, and therefore, the counterparty risk, most people will interact with centralized entities in order to trade on cryptocurrency markets. These entities have the risk of bankruptcy. The counterparty risk arises from the possibility of one of the parties in the contract being unable to meet its' obligations at a certain point in time. Consequently, derivatives on the same underlying could have different prices, as the premium paid for a contract with higher chance of default should be lower when compared with a contract secured by a more stable counterparty. Furthermore, a seller correlated with the underlying asset will have higher default risk in the presence of adverse occasions (Jarrow & Yu, 2001).

When a centralized exchange acts as intermediate between parties, for example, by providing a compensation chamber that ensures appropriate liquidation mechanisms. In this case, the counterparty risk is the exchange, and evaluating the credit risk of that exchange is a crucial step for cautious investors. According to Fantazzini & Calabrese (2021), there are several factors that are important when predicting the likelihood of bankruptcy for an exchange, being the most important whether the development team is public or not. A public exchange goes in line with M. Fama et al. (2019), that returns on this asset class depend on financial conventions and the recognition of it by the market participants. Further important variables include a CER cyber security grade, as well as the number of assets listed. The latter could mean higher fees from trading, and consequently, better funded exchanges.

The lack of transparency when using OTC markets is one of the causes for this excessive risk-taking, as such bilateral agreement have lower reporting standards. Such opacity could lead to excessive risk-taking, as increased position sizes will not lead higher perceived risk or collateral requirements. Then the market is forced to guess the default probability based on limited information, leading to inefficient risk-sharing. One solution to decrease the externality created by counterparty risk is to centralize the clearing process, as doing so will create incentives for the monitoring of excessive risk by the clearing house (Acharya & Bisin, 2014). Because of that, futures exchanges acting as clearing house could require collateral and impose position limits, as the downside of failing to ensure the ability of one party to fulfil its' obligations will be internalized by the clearing house. Investors and traders should then be vigilant on the strictness of requirements of the centralized clearing entity, as this central body could be unable to fulfil its' obligations as well in the case of having to absorb a great default.

However, using a Central Clearing Party usually requires collateral, and using multiple CCPs requires even more collateral, which means lower capital efficiency and reduced netting efficiency. This inefficiency is more accentuated in the case of CCPs for one single class of derivatives. In the absence of the possibility for a counterparty to use just one CCP, interoperability between multiple clearing houses would be a great solution to reduce counterparty risk at the minimum efficiency trade-off (Duffie & Zhu, 2011). In the cryptocurrency sphere, every exchange has its own clearing mechanism, which could imply capital inefficiency. Furthermore, in the presence of significant transfer costs when a transfer needs to be realized on the blockchain, lack of centralization might create some friction to efficiency.

3. Data

3.1. Data Collection and Sources

In order to provide the description of the data related with future contract pricing on FTX, an internal platform provided by the exchange API[1] was used to fetch the price of each contract on a daily basis, the price level of the index that acts as underlying for settlement purposes, as well as the spot price of the underlying. That procedure was used for two underlying assets, Bitcoin and Ethereum. For the first underlying, the sample features contracts with quarterly expirations, starting with the one expiring September 2019 and finishing with the contract expiring in March 2022, being the first day of observations 20th July 2019 and the last 25th March 2022. Data on Ethereum future contracts has the same structure considering the expiration dates and sample period. For Spot prices, sample have the same structure, with the minor difference on the starting date, which is one day after the original one, 21st July 2019 for BTC and 14th September 2019 for ETH. Since all contracts are financially settled, the underlying price is defined by an index or reference rate. FTX uses a custom-made index that accounts for the price of the underlying in several exchanges. Index data is available for all observations on the sample. FTX indexes were retrieved by using the previously referred platforms' API.

The only compounding factor to the model used to price the contracts, the risk-free rate on the 6M U.S. Bond, was retrieved from Refinitiv, considering the daily closing implied yield. The choice of the 6M U.S. Bond matches or at least approximate to the real length of the arbitrage duration. Therefore, by using such rate as benchmark, the theorical value of the contract will reflect the true opportunity cost of performing such strategy during the analysed period. When pricing contracts on the weekend, the interest rate considered is the last available, usually the closing rate on the precedent Friday. The sample period starts on 20th July 2019, and ends on the 25th March 2022.

One important variable for this study in an indicator published weekly by the CFTC (Commodity Futures Trading Commission) denominated COTs (Commitment of Traders)[2]. These reports specify the positions of several types of investors on a variety of regulated future contracts. To this study, the segment of data pointing to the long position and short position of *leveraged funds* will be used as an explanatory variable. Since the granularity of the

---

[1] Retrieved from the following endpoint: https://ftx.com/api.

[2] Data retrieved from: https://www.cftc.gov/MarketReports/CommitmentsofTraders/index.htm.

data does not match with the daily observations of prices, adjustments such as using weekly averages for other variables will be used. For Bitcoin contracts, the CFTC started to issue reports on 10th April 2018, whereas for Ethereum contracts, it started on 6th April 2021.

Another important variable is the Multifactorial Sentiment Index, which is a proxy to measure sentiment regarding Bitcoin. It considers volatility, volume, social media, surveys, dominance, and trends to measure how optimistic or pessimistic investors are. The sample consists of daily observations starting on 1st February 2018, with the last observation on 25th March 2022. The index was retrieved from Alternative.me.

3.2. Descriptive Statistics

This chapter features the description of all collected data on both exchanges as well as some external data that will be used on the methodology. Mispricing on FTX series will be computed accordingly with Butterworth & Holmes (2000):

$$M_{t,T} = \frac{F_{t,T} - F^*_{t,T}}{S_t} \tag{3.1}$$

The first variable $F_{t,T}$ is observable on the exchange, containing the price of a contract with expiration on T on the moment t. By using this formula, the mispricing will be measured against the Index price, represented by S_t. The theoretical price of the future price is computed according to the equation 3.2, the standard formula to price contracts based on the on the cost of carry framework.

$$F^*{}_{t,T} = S_t * e^{r_{t,T} * (T-t)} \tag{3.2}$$

Table 1 exhibits the correlation between the variables that will be analysed along the chapter. As starting point, such statistic provides a hint in how the variables interact between themselves. BTC Mispricing, the main variable correlates positively with BTC Price (0.406), has positive correlation with the sentiment index of 0.548 and a negative correlation of -0.339 with the Net Position of Leveraged Funds. It is expected that mispricing moves higher when sentiment is increases, probably the same relationship expected for BTC Price, and the net position of leveraged funds decreases. On Table 1, the represented mispricing refers to the one computed for BTC contracts.

Table 1 - Correlation Table for BTC Contracts

	Mispricing	BTC Price	Sentiment	OI LF
Mispricing	1.000	0.406	0.548	-0.339
BTC Price	0.406	1.000	0.250	-0.778
Sentiment	0.548	0.250	1.000	-0.382
NP LF	-0.339	-0.778	-0.382	1.000

When it comes to analysing the correlations of comparable variables but regarding to ETH, only sentiment shows a positive sign as in BTC. ETH Mispricing has a correlation of 0.550 with Sentiment, like previously observed correlation. Not in line with the BTC correlation is the negative correlation of -0.152 of the underlying prices and the mispricing, as well as a positive correlation of 0.281 between mispricing and net position of leveraged funds. The correlation matrix for ETH is presented below in Table 2.

Table 2 - Correlation Table for ETH Contracts

	Mispricing	ETH Price	Sentiment	OI LF
Mispricing	1.000	-0.152	0.550	0.281
ETH Price	-0.152	1.000	0.347	-0.683
Sentiment	0.550	0.347	1.000	-0.370
NP LF	0.281	-0.683	-0.370	1.000

3.2.1. FTX Contracts

Contracts listed on this exchange trade every day without interruption. Prices were collected once a day at 00:00 UTC. Table 3 and Table 5 show mispricing on the FTX exchange for front-quarter BTC contracts and ETH contracts respectively, meaning that values reflect the conditions for the period of the contract. The period analysed by each contract is specified in the Table 30 in the annex. Table 4 and Table 6 will exhibit the same stats, however without any restrictions on observations, leading to a different representation of the data, as some contracts were listed simultaneously with different maturities, and possibly higher premiums/discounts in contracts with higher TTM (Time to Maturity). In essence, values are not comparable across contracts due to differences in sample characteristics.

Table 3 - Summary statistics relating to the mispricing series of BTC Contracts on FTX (restricted to the front-quarter contracts).

Contract Name	Obs	Undervalued	Overvalued	Mean	T-stat	S.D.	Skew	Kurtosis	Min.	Max.	Reversals	Level 1 Fee	Level 2 Fee	Level 3 Fee
BTC-20190927	70	6	64	0.0061	7.7541	0.0066	-0.7849	5.4138	-0.0233	0.0227	9	40	5	0
BTC-20191227	91	19	72	0.0044	7.9080	0.0053	0.5344	-0.3597	-0.0052	0.0174	21	38	5	0
BTC-20200327	91	15	76	0.0147	8.8631	0.0158	-1.1690	7.5263	-0.0663	0.0681	2	72	44	23
BTC-20200626	91	21	70	0.0023	4.7596	0.0046	-0.4348	0.3968	-0.0108	0.0127	17	26	0	0
BTC-20200925	91	4	87	0.0108	12.6461	0.0082	0.2033	-0.7334	-0.0090	0.0277	2	67	25	4
BTC-20201225	91	2	89	0.0112	12.7555	0.0083	-4.1229	28.0837	-0.0483	0.0237	1	74	30	0
BTC-20210326	91	1	90	0.0269	14.7792	0.0173	-1.2825	4.5678	-0.0573	0.0550	1	83	66	52
BTC-20210625	91	8	83	0.0309	9.5959	0.0307	0.6722	-0.6763	-0.0018	0.1139	2	56	49	47
BTC-20210924	91	0	91	0.0091	13.7021	0.0064	1.1253	3.1518	0.0001	0.0378	0	63	18	1
BTC-20211231	98	2	96	0.0154	13.5376	0.0113	0.0721	-1.5251	-0.0006	0.0357	2	69	51	29
BTC-20220325	84	20	64	0.0049	7.0795	0.0064	0.9236	-0.0956	-0.0078	0.0202	15	37	11	0
	980	**10.00%**	**90.00%**						**-0.0663**	**0.1139**	**72**	**63.78%**	**31.02%**	**15.92%**

Observations refer to the daily snapshot of mispricing in BTC contracts in FTX based on equation 3.1. Total number of observations and in the case of front-quarter series, refer to the total number of days in the sample, presented in the last row of the table. Percentages presented in the last row of the columns Undervalued and Overvalued are the portion of days that contracts were undervalued and overvalued, respectively. Last row in columns Min. and Max. represents the minimum value and maximum values of mispricing observed in that data series. The last row of the column Reversals is the total number of times that mispricing became negative from a positive value and vice-versa. The last row of columns Level 1 Fee, Level 2 Fee and Level 3 Fee, represent the portion of days that mispricing was higher than 0.5%, 1.5% and 2.5% respectively.

Table 4 - Summary statistics relating to the mispricing series of BTC Contracts on FTX (all observations).

Contract Name	Obs	Undervalued	Overvalued	Mean	T-stat	S.D.	Skew	Kurtosis	Min.	Max.	Reversals	Level 1 Fee	Level 2 Fee	Level 3 Fee
BTC-20190927	70	6	64	0.0061	7.7541	0.0066	-0.7849	5.4138	-0.0233	0.0227	9	40	5	0
BTC-20191227	124	20	104	0.0085	10.0954	0.0094	0.9418	0.1819	-0.0052	0.0366	23	69	26	11
BTC-20200327	196	24	172	0.0140	14.6298	0.0134	-0.7201	6.6755	-0.0663	0.0681	12	152	82	41
BTC-20200626	182	35	147	0.0129	7.8182	0.0223	-3.0063	26.8385	-0.1749	0.0582	22	101	74	52
BTC-20200925	204	23	181	0.0094	8.0221	0.0168	-5.3820	60.3205	-0.1666	0.0601	18	149	45	11
BTC-20201225	201	2	199	0.0196	22.6479	0.0122	-0.1607	4.2215	-0.0483	0.0520	1	184	133	49
BTC-20210326	197	1	196	0.0322	33.5821	0.0134	-2.0596	9.2323	-0.0573	0.0550	1	189	172	156
BTC-20210625	209	8	201	0.0521	25.4772	0.0296	-0.5281	-0.8038	-0.0018	0.1139	2	174	167	165
BTC-20210924	254	0	254	0.0622	18.5582	0.0534	0.3475	-1.4253	0.0001	0.1814	0	225	176	152
BTC-20211231	352	2	350	0.0754	20.6116	0.0687	0.8030	-0.7967	-0.0006	0.2628	2	323	301	267
BTC-20220325	208	20	188	0.0307	17.3190	0.0256	0.2925	-1.2775	-0.0078	0.0796	15	161	135	112
	2197	**6.42%**	**93.58%**						**-0.1749**	**0.2628**	**105**	**80.43%**	**59.90%**	**46.24%**

Observations refer to the daily snapshot of mispricing in BTC contracts in FTX based on equation 3.1. Total number of observations and in the case of the non-restricted series, refer to the total number of observations in the sample, presented in the last row of the table. Percentages presented in the last row of the columns Undervalued and Overvalued are the portion of observations that contracts were undervalued and overvalued, respectively. Last row in columns Min. and Max. represents the minimum value and maximum values of mispricing observed in that data series. The last row of the column Reversals is the total number of times that mispricing became negative from a positive value and vice-versa. The last row of columns Level 1 Fee, Level 2 Fee and Level 3 Fee, represent the portion of observations that mispricing was higher than 0.5%, 1.5% and 2.5% respectively.

Table 5 - Summary statistics relating to the mispricing series of ETH Contracts on FTX (restricted to the front-quarter contracts).

Contract Name	Obs	Undervalued	Overvalued	Mean	T-stat	S.D.	Skew	Kurtosis	Min.	Max.	Reversals	Level 1 Fee	Level 2 Fee	Level 3 Fee
ETH-20190927	70	9	61	0.0039	3.8875	0.0084	-3.9684	26.9008	-0.0511	0.0180	6	27	5	0
ETH-20191227	91	10	81	0.0049	9.1011	0.0051	0.4629	-0.8169	-0.0051	0.0167	12	36	3	0
ETH-20200327	91	13	78	0.0164	9.9660	0.0157	-0.5789	3.3583	-0.0526	0.0594	6	74	47	28
ETH-20200626	91	29	62	0.0012	1.7700	0.0063	-1.0394	1.8971	-0.0230	0.0150	19	27	1	0
ETH-20200925	91	17	74	0.0124	10.3288	0.0115	0.3431	-0.4751	-0.0149	0.0375	7	67	28	17
ETH-20201225	91	3	88	0.0087	13.7351	0.0060	-2.1960	13.0772	-0.0272	0.0223	3	71	5	0
ETH-20210326	91	1	90	0.0306	14.9980	0.0195	-0.7792	1.3079	-0.0483	0.0637	1	85	65	57
ETH-20210625	91	5	86	0.0352	10.1033	0.0333	0.5991	-0.9210	-0.0035	0.1092	4	66	54	50
ETH-20210924	91	0	91	0.0083	11.4341	0.0069	4.3731	30.2852	0.0004	0.0588	0	61	8	1
ETH-20211231	98	2	96	0.0154	13.0129	0.0117	0.1306	-1.4397	-0.0015	0.0365	2	70	49	29
ETH-20220325	84	29	55	0.0045	6.5949	0.0063	1.1502	0.1846	-0.0013	0.0207	18	31	9	0
	980	**12.04%**	**87.96%**						**-0.0526**	**0.1092**	**78**	**62.76%**	**27.96%**	**18.57%**

Observations refer to the daily snapshot of mispricing in ETH contracts in FTX based on equation 3.1. Total number of observations and in the case of front-quarter series, refer to the total number of days in the sample, presented in the last row of the table. Percentages presented in the last row of the columns Undervalued and Overvalued are the portion of days that contracts were undervalued and overvalued, respectively. Last row in columns Min. and Max. represents the minimum value and maximum values of mispricing observed in that data series. The last row of the column Reversals is the total number of times that mispricing became negative from a positive value and vice-versa. The last row of columns Level 1 Fee, Level 2 Fee and Level 3 Fee, represent the portion of days that mispricing was higher than 0.5%, 1.5% and 2.5% respectively.

Table 6 - Summary statistics relating to the mispricing series of ETH Contracts on FTX (all observations).

Contract Name	Obs	Undervalued	Overvalued	Mean	T-stat	S.D.	Skew	Kurtosis	Min.	Max.	Reversals	Level 1 Fee	Level 2 Fee	Level 3 Fee
ETH-20190927	70	9	61	0.0039	3.8875	0.0084	-3.9684	26.9008	-0.0511	0.0180	6	27	5	0
ETH-20191227	111	13	98	0.0069	9.3813	0.0078	0.9601	0.6984	-0.0066	0.0335	15	52	17	4
ETH-20200327	110	15	95	0.0144	10.0441	0.0150	-0.2320	3.0931	-0.0526	0.0594	8	85	47	28
ETH-20200626	113	44	69	0.0003	0.2093	0.0140	-4.3351	35.3978	-0.1111	0.0298	22	34	7	5
ETH-20200925	110	17	93	0.0122	12.2318	0.0105	0.4145	-0.0022	-0.0149	0.0375	7	86	31	17
ETH-20201225	106	4	102	0.0085	14.7578	0.0059	-1.9451	11.5821	-0.0272	0.0223	5	81	6	0
ETH-20210326	119	1	118	0.0317	19.6215	0.0176	-0.9250	2.1566	-0.0483	0.0637	1	113	93	82
ETH-20210625	170	5	165	0.0553	20.7335	0.0348	-0.2889	-1.1748	-0.0035	0.1119	4	145	133	129
ETH-20210924	254	0	254	0.0656	18.4965	0.0565	0.3301	-1.3571	0.0004	0.2079	0	224	165	152
ETH-20211231	352	2	350	0.0797	19.7832	0.0755	0.8276	-0.7283	-0.0015	0.2982	2	324	298	264
ETH-20220325	208	29	179	0.0326	17.1839	0.0273	0.2405	-1.3699	-0.0013	0.0818	18	155	133	119
	1723	**8.07%**	**91.93%**						**-0.1111**	**0.2982**	**88**	**76.96%**	**54.27%**	**46.43%**

Observations refer to the daily snapshot of mispricing in ETH contracts in FTX based on equation 3.1. Total number of observations and in the case of the non-restricted series, refer to the total number of observations in the sample, presented in the last row of the table. Percentages presented in the last row of the columns Undervalued and Overvalued are the portion of observations that contracts were undervalued and overvalued, respectively. Last row in columns Min. and Max. represents the minimum value and maximum values of mispricing observed in that data series. The last row of the column Reversals is the total number of times that mispricing became negative from a positive value and vice-versa. The last row of columns Level 1 Fee, Level 2 Fee and Level 3 Fee, represent the portion of observations that mispricing was higher than 0.5%, 1.5% and 2.5% respectively.

The last three columns of both tables represent the number of observations that feature a mispricing higher than a certain level of fees, implicit and explicit. There are three levels, the first corresponding to fees of 0.5%, the second to 1.5% and the third 2.5%. Differences between the index price and the actual spot price, which may be interpreted as a hidden fee, were analysed in the following sub-chapter, with a maximum spot premium of 19 basis points of spot premium for Bitcoin and 49 basis points for Ethereum. Bid-Ask spreads on BTC and Ethereum markets are an implicit cost and should be accounted for. According to Brauneis et al. (2021), and considering the third quartile and the most expensive exchange on the study to be conservative, the accumulated spread of a round-trip trade is 15.7 basis points for BTC and 35.24 basis points for Ethereum. Transaction Fees are seven basis points per trade for the least advantaged trader[3] meaning that the four-operation process needed to gather the mispricing may account to four times that values, at 28 basis points. Additionally, some investors may be unable to own Bitcoin or Ethereum outside of a regulated custodian. A custodian fee of 50 bps per year is the rate practiced by Coinbase Custody[4]. According to the average fees on the crypto exchanges, possible custody costs for institutional investors, and differences between the spot price and the index price, providing three levels of hurdle rate provides the practical flexibility for different investors with different conditions regarding fees, liquidity, and safety, to judge if mispricing is worth exploring or not. According to the observed rates above, the investor with worse conditions will face fees of 112.7 basis points, not considering adjusting the custodian fee for the average time of a contract, still considering annual fees. All in all, the Level 2 hurdle rate of 1.5% should cover the great majority of investors.

Apart from usual statistics, both tables include a variable denominated as *Reversals*. That column counts the number of times that premium became a discount and vice-versa during the life of the contract. Such indicator could indicate the opportunity to gather the profit of mispricing convergence before maturity. On Table 3, the contracts with the higher number of reversals had expiration dates on December 2019 and June 2020. On those contracts, the mispricing converged to zero from a premium, or discount, 21 and 17 times respectively. Additionally, it is curious to note that during the lifetime of the contracts, mispricing never went above 2.5%, as the maximum premium for both contracts was 1.74% (December 2019)

[3] Retrieved from: https://help.ftx.com/hc/en-us/articles/360024479432-Fees .
[4] Retrieved from: https://www.coinbase.com/prime/custody .

and the maximum discount was 1.08% (June 2020). Even though the t-statistic supports the non-zero mispricing hypothesis, considering these values, we could assert that the price of the futures was tracking its fair value relatively close, providing little opportunity to make arbitrage worthwhile after considering all the involved costs. Regarding Table 5, the quarters with most reversal are the same, however with a slightly higher number of reversals for the contract expiring in March 2022. Considering the previously referred contracts, the maximum premium on aggregate was 1.67% and the minimum (the higher observed discount) was -2.30%. Such narrow amplitudes between the maximum and the minimum when compared to other contracts, hint for good tracking properties during those quarters.

On Table 3, there are four different contracts that exhibit meaningful premiums, this means, frequently above the 2.5% threshold, that even the marginal investor with suboptimal conditions could use to arbitrage. Those contracts are the ones expiring in March 2020 (23 days above level 3 fee), March 2021 (the highest number of days, 52), June 2021 (47 days) and December 2021 (29 days). It is important to notice that every premium above fee levels would be harvested, in theory, always in less than 98 days, so providing quarterly absolute rates, implying annualized rates being much higher. On Table 4, the effect on longer duration contracts on mispricing is clear. As contracts with TTMs as high as 352 days (case for the first day of trading on December 2021 contracts), and because some long-term contracts were trading during the periods referred to in Table 3, maximum premiums on Table 4 are higher. The same relationship is observed between the Table 5 and Table 6, for Ethereum futures.

Comparing the extremes of front quarter series, for both underlying assets, BTC and ETH, similar values are observed, as the maximum premiums were 11.39% and 10.92% respectively, observed simultaneously on the contract expiring in June 2021. On the other side, maximum discounts occurred during the same period, on the contract expiring March 2020. During this quarter, the highest discount observed on BTC contract was 6.63%, with ETH contract being simultaneously negatively mispriced by 5.26%. Considering the sample without restrictions, allowing for the calculation of mispricing on longer-term contracts, wider intervals between the extremes are observed; the highest premium recorded in BTC contracts was 26.28%, while ETH contracts registered a premium of 29.82%, both on the contract with longer lifetime, expiring in December 2021. The highest discount is observed on the contract expiring in June 2020, as BTC futures reached a discount of 17.49% and

ETH futures were trading 11.11% below fair value. It is important to register that June 2020 contracts were trading simultaneously with the one expiring in March 2020, when the highest discount on the restricted sample was observed. The higher discount values may reflect the compound effect of such discount. The same applies for the extreme premiums, as December 2021 contracts were trading concurrently with the June 2021 contracts, possibly exhibiting the same compound effect.

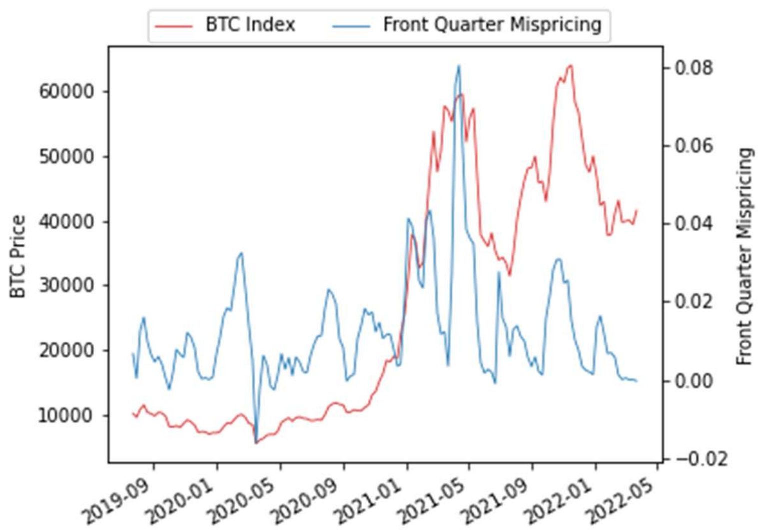
The norm for both types of contracts seems to be overpricing, the situation observed 90% of the days in BTC contracts, alongside 87.96% of the time for Ethereum contracts.

Among all contracts, only one contract has shown an observed *t* that did not allow for the null hypothesis to be rejected. Considering the ETH contract expiring in June 2020 (Table 5 and 6), and according to our sample, restricted and unrestricted, at 95% confidence level, there is no statistical evidence to support a mispricing different than zero.

3.2.2. FTX Index and Spot Price Data

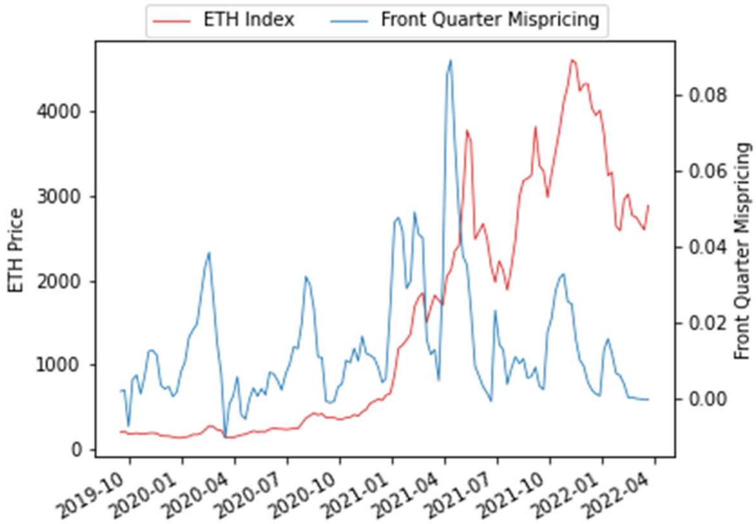
As prices are aggregate information, BTC and ETH prices do incorporate expectations on the future of those assets. It makes sense to analyse if there is any relationship between spot prices and mispricing. Figure 1 shows the evolution of front quarter mispricing and spot prices.

Figure 1 - BTC Index and Future Front Quarter Mispricing (Weekly Observations).



Bear in mind that a because of the nature of this data series, that quotes mispricing on the front quarterly contract which creates a magnet to zero mispricing when TTM is approaching zero. For ETH, the similar Figure 2 is shown below:

Figure 2 - ETH Index and Future Front Quarter Mispricing (Weekly Observations).



As the mispricing of each contract was measured against the index price, defined by the exchange as a combination of prices observed on different exchanges. This practice common, as using index rather than spot prices would difficult manipulatory practices. However, in order to execute the cash and carry strategy, one must be able to buy the underlying asset. As the index is computed as the average of trading pairs on FTX, FTXUS, Binance, OKX, Coinbase, Kraken, Bitstamp and Bittrex, getting exposure spot prices on all of these exchanges adds a layer of complexity to someone wanting to harvest any mispricing opportunity, due to inter exchange operations and opportunity cost of using the underlying asset as collateral to the futures position on FTX. It seems reasonable to increase the real-life utility of this study by measuring the FTX spot price against the index used for contract pricing.

Spot Price data samples have the distinct number of observations when considering BTC and ETH. Spot BTC price history on FTX dates back to the second available trading day on studied contracts, from 21st July 2019 until 25th March 2022 (979 days). Regarding ETH spot price history, the first observation was timestamped on 14th September 2019, being the last 25th March 2022. As used to price futures, all observations were equally recorded on 00:00 UTC. The sample is going to be split by quarters such as the period of each contract.

On the other hand, Index Price data covers all the studied period for both BTC and ETH. Both samples comprise daily observations from 20th July 2019 to 25th March 2022. The difference between the index and the spot price will be measured as the premium of FTX spot price over the index, using the formula as follows:

$$Spot_{Premium} = \frac{Spot\ Price - Index\ Price}{Index\ Price} \tag{3.3}$$

The next table describes the mispricing data for the spot premium for Bitcoin. For all contracts, the mean oscillates between -0.02% and 0.03%. The maximum premium observed was 0.19 p.p. on the Q3 2021 while 0.14 p.p. was the highest discount of spot relative to the index. These values are one of the reasons for which testing the mispricing for certain thresholds is appropriate, as it acts as an implicit fee. When the spot price is higher than the index, the mispricing could imply higher profits than could really be harvested, if we choose to buy it instead of replicating the index.

Table 7 - Descriptive Statistics for BTC Spot Premium Data, grouped by contract period.

Contract Name	Obs	Mean	Median	T-stat	S.D.	Maximum	Minimum
BTC-20190927	69	0.03%	0.02%	3.2411	0.07%	0.15%	-0.13%
BTC-20191227	91	-0.02%	-0.02%	-3.3454	0.05%	0.13%	-0.12%
BTC-20200327	91	0.00%	0.00%	-0.3223	0.05%	0.12%	-0.14%
BTC-20200626	91	0.00%	0.00%	0.0675	0.04%	0.09%	-0.12%
BTC-20200925	91	0.00%	0.00%	0.5035	0.03%	0.09%	-0.12%
BTC-20201225	91	-0.01%	-0.01%	-2.4162	0.04%	0.08%	-0.11%
BTC-20210326	91	-0.02%	-0.02%	-4.2761	0.05%	0.12%	-0.13%
BTC-20210625	91	0.00%	0.00%	-0.2288	0.05%	0.12%	-0.10%
BTC-20210924	91	0.01%	0.01%	1.5082	0.05%	0.19%	-0.12%
BTC-20211231	98	0.00%	0.00%	0.6137	0.04%	0.13%	-0.09%
BTC-20220325	84	-0.01%	-0.01%	-2.5633	0.03%	0.09%	-0.14%
						0.19%	-0.14%

Observations refer to the daily snapshot of BTC Spot Premium (difference between the index price used for settlement and spot price). The last row of Maximum and Minimum represent the maximum and minimum values, respectively, for the sample.

For the premium of Spot ETH to Index, the mean is similar to Bitcoin, -0.04% and 0.01%, the same amplitude of 5 basis points. However, the extremes exhibit higher amplitude and values, with the maximum registered premium being 49 basis points and the highest discount 37 basis points.

Table 8 - Descriptive Statistics for ETH Spot Premium Data, grouped by contract period.

Contract Name	Obs	Mean	Median	T-stat	S.D.	Maximum	Minimum
ETH-20190927	14	0.00%	0.00%	-0.0439	0.11%	0.19%	-0.20%
ETH-20191227	91	-0.04%	-0.04%	-5.8748	0.07%	0.13%	-0.21%
ETH-20200327	91	-0.02%	-0.02%	-2.2932	0.09%	0.16%	-0.37%
ETH-20200626	91	0.00%	-0.01%	-0.3558	0.06%	0.19%	-0.21%
ETH-20200925	91	0.00%	0.00%	0.4413	0.06%	0.17%	-0.15%
ETH-20201225	91	-0.01%	0.00%	-1.2384	0.05%	0.12%	-0.14%
ETH-20210326	91	-0.02%	-0.02%	-2.3698	0.07%	0.17%	-0.33%
ETH-20210625	91	0.01%	0.00%	1.3161	0.09%	0.49%	-0.18%
ETH-20210924	91	0.01%	0.01%	1.4916	0.05%	0.20%	-0.15%
ETH-20211231	98	0.00%	0.00%	-0.4895	0.07%	0.19%	-0.28%
ETH-20220325	84	0.00%	0.00%	-0.4760	0.05%	0.13%	-0.15%
						0.49%	-0.37%

Observations refer to the daily snapshot of ETH Spot Premium (difference between the index price used for settlement and spot price). The last row of Maximum and Minimum represent the maximum and minimum values, respectively, for the sample.

A histogram of the data is present on annex, Figure 5 for BTC and Figure 6 for ETH.

3.2.3. Sentiment Data

Sentiment is measured on a scale of 100, with values defined accordingly with the previously referred parameters. To maintain consistency on the data description process, the sample will be divided the same as contracts period. The following Table 9 describes sentiment data.

Table 9 - Descriptive Statistics for Sentiment Data, grouped by contract period.

Contract Name	Obs	Mean	Median	S.D.	Maximum	Minimum
BTC/ETH-20190927	70	36.13	38.5	14.15	66	5
BTC/ETH-20191227	91	34.11	33	10.22	56	15
BTC/ETH-20200327	91	40.22	42	16.46	65	8
BTC/ETH-20200626	91	34.49	40	14.83	56	8
BTC/ETH-20200925	91	57.79	49	17.12	84	38
BTC/ETH-20201225	91	75.66	86	18.51	95	40
BTC/ETH-20210326	91	80.19	79	12.53	95	38
BTC/ETH-20210625	91	44.65	37	24.65	79	10
BTC/ETH-20210924	91	46.86	48	22.35	79	10
BTC/ETH-20211231	98	50.54	49	22.15	84	16
BTC/ETH-20220325	84	27.92	24	10.73	54	10

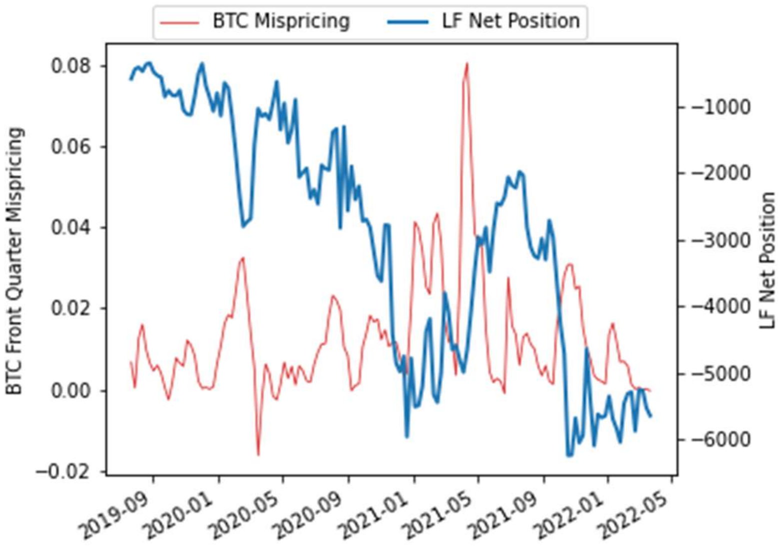
Observations refer to the daily snapshot of Sentiment Index, ranges between 1 and 100, being 1 pessimistic and 100. Periods are the same of the contracts.

The number of observations matches the ones on Table 3 and 5, maximizing the number of observations that could be used as explanatory variable for mispricing in the next chapter. The period with the highest average sentiment, registering 80.19 was the contract expiring in March 2021, while the previous period (expiring in December 2020) registered the highest median of 86. On the opposite, the lowest average and median was registered in the last observed period (expiring March 2022). The highest sentiment periods match with the periods where most observations surpassed the Level 3 Fee threshold, while the last period, the more pessimistic one, register 0 observations above the 250bps threshold.

3.2.4. Commitment of Traders

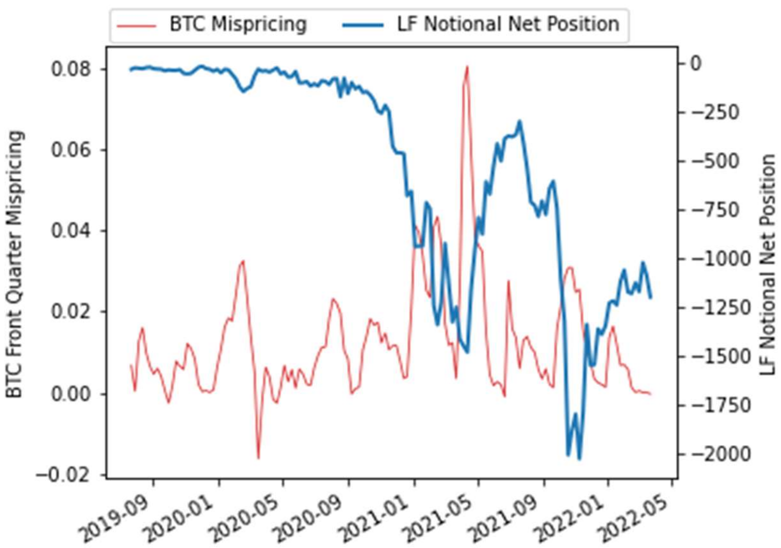
The portion of this dataset corresponding to the long and short positions of leveraged funds will be used as a proxy for the positioning of arbitrageurs in relation to studied contracts. The following Figure 3 shows the evolution of this indicator during the time of studied contracts. To contextualize historically, a simultaneous representation of front-quarter mispricing will be in the chart, on the left axis. The negative correlation referred on the beginning of the chapter is noticed the representation of the data.

Figure 3 - Evolution of Net Position of Leveraged Funds (measured in number of contracts) and Front Quarter Mispricing for BTC Contracts.



There is another caveat on this type of data. The number of contracts does not consider the notional value of the position of this type of traders. Since Bitcoin value was $11 759.49 on the first day represented, reached $66 921.6 on November 2021, and $42 374.67 on the last week of the observation period, it is plausible to analyse the notional net position. Since there is no information on which contracts the positions were taken, adjusting to the price of the underlying and contract size serves as a proxy. The adjusted data is shown the Figure 4 below:

Figure 4 - Evolution of Net Position of Leveraged Funds (measured in Million USD) and Front Quarter Mispricing for BTC Contracts.



In the end of both datasets, the behaviour seems different as in the first figure, one might conclude that exposure of those funds is near all-time highs as mispricing is exhibiting premiums near zero. In the second one, the global minimum for notional net position coincides with a local maximum for mispricing. For ETH, identical graphs are found in the annex as Figures 7 and 8.

4. Methodology

On aggregate, the methodology of the dissertation consists in using the factors that are believed to be correlated to mispricing by the literature on the topic. All the following equations will be estimated using the least squares method (OLS).

4.1. Underlying Spot Price

To test the relationship of Spot Price for both assets with the mispricing, the following equation will used. According to the previous literature review, and considering the positive correlation between both variables, the coefficient is expected to be positive.

$$M_t = \beta_1 + \beta_2 S_t + e_t \tag{4.1}$$

The dependent variable M_t, represents the mispricing (discount or premium), of the future contract in question. The variable S_t represents the price of underlying in moment t. As observations have different Time to Maturity, the following equation aims at controlling for those differences, by implementing the variable TTM_t.

$$M_t = \beta_1 + \beta_2 S_t + \beta_3 TTM_t + e_t \tag{4.2}$$

4.2. Sentiment

Consistent with the conclusion reached by Stambaugh et al. (2012) that anomalies are higher after period of optimism, the following equation 4.3 aims to test the relationship between mispricing, which can be considered an anomaly, and the sentiment observed in the cryptocurrency market, by using as proxy the previously referred sentiment index.

$$M_t = \beta_1 + \beta_2 Sent_t + e_t \tag{4.3}$$

The explanatory variable $Sent_t$, refers to the to the sentiment index observed during observation *t*. Because of the dynamics of futures, more specifically the fact that at settlement mispricing vanishes as the price of the future is supposed to be equal to spot price (in this case index price), a variable TTM is going to be present on the following equation to test if, along with sentiment, the time remaining to settlement affects this value.

$$M_t = \beta_1 + \beta_2 Sent_t + \beta_3 TTM_t + e_t \tag{4.4}$$

4.3. Commitment of Traders

In order to test the relationship between net position of leveraged traders on the mispricing of future contracts, the following equation will be used:

$$M_t = \beta_1 + \beta_2 NP_t + e_t \tag{4.5}$$

The variable that regards the net position of leveraged funds on BTC and ETH contracts was abbreviated to NP_t. As previously presented, the proxy for the notional value of this variable might be more appropriated, so the following equation is going to be used as well:

$$M_t = \beta_1 + \beta_2 NP_{t}\_Notional + e_t \tag{4.6}$$

The only difference is that $NP\_Notional_t$ reflects the total exposure in monetary units, in billions of us dollars and not in number of contracts. Finally, the following two equations add time to maturity to the previous ones:

$$M_t = \beta_1 + \beta_2 NP_t + \beta_3 TTM_t + e_t \tag{4.7}$$

$$M_t = \beta_1 + \beta_2 NP\_Notional_t + \beta_3 TTM_t + e_t \tag{4.8}$$

This last variable is used to control differences in mispricing that are not necessarily related with any other variable other than time, controlling the effect of observations with different TTM.

4.4. Summary

In order to maximize the explanatory power of a given regression, and because empirical results show that variables feature high levels of individual significance, the following to equations will be used to build a more complete model.

$$M_t = \beta_1 + \beta_2 S_t + \beta_3 Sent_t + \beta_4 NP\_Notional_t + e_t \tag{4.9}$$

$$M_t = \beta_1 + \beta_2 S_t + \beta_3 Sent_t + \beta_4 NP\_Notional_t + \beta_5 TTM_t + e_t \tag{4.10}$$

5. Empirical Results

In this chapter, an empirical analysis of the data will be performed. Timeseries used will have weekly granularity to build different models with those variables. Variables with daily granularity will be inserted as a weekly average of their respective daily observations.

5.1. Underlying Spot Price

The following table represents the regression results for equations 4.1 and 4.2. Bitcoin price is measured in thousands of US dollars, in order to make the reading of the results easier to read.

Table 10 - Regression results summary for BTC Front Month contracts, underlying price as explanatory variable.

BTC Mispricing	4.1	4.2
Constant	0.004614	-0.006234
	(0.001895)	(0.002345)
BTC Price	0.000302***	0.000305***
	(0.0000578)	(0.0000506)
TTM		0.000243***
		(0.0000370)
R-squared	0.165	0.364
Adjusted R-squared	0.159	0.355
Observations	140	

Standard errors are reported in parentheses. *, **, *** indicates significance at the 90%, 95%, and 99% level, respectively. Coefficients represent the absolute impact on Mispricing. Price is measured in thousands USD, and TTM in days. All regressions represented achieve global significance at a confidence level of 99%.

The results presented imply a positive relationship between *BTC price* and *Mispricing,* as each time that BTC quote increases by a thousand US Dollars, on average, the mispricing is expected to increase by 3.02 basis points. On the time adjusted equation (B), such coefficient is 2.05 basis points, while the time value of a day is 2.43 basis points, that is, the amount by which mispricing is expected to decrease as the contract approaches maturity. The variable is individually significant on both equation settings, explaining as much as 36.4% of mispricing variance on the time-adjusted one.

Table 11 - Regression results summary for ETH Front Month contracts, underlying price as explanatory variable.

ETH Mispricing	4.1	4.2
Constant	0.011028	-0.0000783
	(0.002072)	(0.002996)
ETH Price	0.001735*	0.002010**
	(0.001005)	(0.000930)
TTM		0.000240***
		(0.0000499)
R-squared	0.022	0.171
Adjusted R-squared	0.015	0.158
Observations	132	

Standard errors are reported in parentheses. *, **, *** indicates significance at the 90%, 95%, and 99% level, respectively. Coefficients represent the absolute impact on Mispricing. Price is measured in thousands USD, and TTM in days. Regressions 4.1 and 4.2 achieve global significance at the confidence level of 90% and 99% respectively.

The results when applying equations to ETH based contracts are not clear as for the BTC ones. The sign on the coefficients remains the same as the one estimated for BTC contracts, however, showing lower levels of significance on both equations. On the first equation, it is estimated that for each price increase of one thousand in ETH Price, mispricing will increase by 17.35 basis points. The regression feature very low levels for the R-squared, being able to explain only 2.2% of the dependent variable variance, not achieving global significance on the 95% significance level, only on 90%. The time-adjusted equation features higher levels of individual significance for the Ethereum Price, estimating an increase of 20.1 basis points per each thousand dollars price increase in ETH, being significant at 95% confidence level.

5.2. Sentiment

The following table represents the regression results for equations 4.3 and 4.4.

Table 12 - Regression results summary for BTC Front Quarter contracts, Sentiment as explanatory variable.

BTC Mispricing	4.3	4.4
Constant	-0.003667	-0.013770
	(0.002349)	(0.002481)
Sentiment	0.000337***	0.000332***
	(0.0000438)	(0.0000377)
TTM		0.000234***
		(0.0000333)
R-squared	0.301	0.486
Adjusted R-squared	0.295	0.478
Observations	140	

Standard errors are reported in parentheses. *, **, *** indicates significance at the 90%, 95%, and 99% level, respectively. Coefficients represent the absolute impact on Mispricing. Sentiment is measured in a scale from 1 to 100, and TTM in days. All regressions represented achieve global significance at a confidence level of 99%.

On equation 4.3 *Sentiment* is used as an explanatory variable for the mispricing level, and the results could be interpreted as the value of each index point in mispricing terms, meaning that on average for each point increase in the sentiment index, the level of mispricing increases 0.0337 p.p. or 3.37 basis points. It is worth noting the predictive power of *Sentiment*, that is capable of explain a relatively high proportion of the variance of mispricing on its own. Among univariable models, Sentiment explains on its own, 30.1% of the target variable, when compared to 16.5% for BTC Price and even lower values for the Leveraged Funds Net Position related variables.

Table 13 - Regression results summary for ETH Front Quarter contracts, Sentiment as explanatory variable.

ETH Mispricing	4.3	4.4
Constant	-0.005529	-0.015586
	(0.002815)	(0.003103)
Sentiment	0.000389***	0.000386***
	(0.0000517)	(0.0000466)
TTM		0.000230***
		(0.0000410)
R-squared	0.303	0.440
Adjusted R-squared	0.298	0.431
Observations	132	

Standard errors are reported in parentheses. *, **, *** indicates significance at the 90%, 95%, and 99% level, respectively. Coefficients represent the absolute impact on Mispricing. Sentiment is measured in a scale from 1 to 100, and TTM in days. All regressions represented achieve global significance at a confidence level of 99%.

The results for the explanatory power of sentiment on ETH mispricing are similar to BTC ones. According to the first equation, it is estimated that for each point increase in the sentiment index, mispricing will increase by 3.89 basis points. The univariable regression is able to explain as much as 30.3% of ETH Mispricing variance. On the time-adjusted version, the regression can explain 44% of the dependent variable variance, estimating the value of each sentiment index point at 3.86 basis points of mispricing. The time value of a day measured in mispricing is estimated to be 2.3 basis points, very similar to the time value observed in BTC based contracts.

5.3. Commitment of Traders

For the third and last variable, both contract net position a nominal net position was tested to explain mispricing, first individually, and later inside a time-adjusted equation.

Table 14 - Regression results summary for BTC Front Quarter contracts, Net Position of LF as explanatory variable.

BTC Mispricing	4.5	4.6	4.7	4.8
Constant	0.004856	0.007196	-0.006301	-0.003219
	(0.002165)	(0.001540)	(0.002584)	(0.002136)
Net Position	-0.002636***		-0.002734***	
	(0.000622)		(0.000548)	
Notional Net Position		-0.010259***		-0.010103***
		(0.002012)		(0.001776)
TTM			0.000245***	0.000237***
			(0.0000383)	(0.0000375)
R-squared	0.115	0.159	0.319	0.349
Adjusted R-squared	0.109	0.152	0.309	0.339
Observations	140			

Standard errors are reported in parentheses. *, **, *** indicates significance at the 90%, 95%, and 99% level, respectively. Coefficients represent the absolute impact on Mispricing. Net Position is measured in number of contracts, Notional Net Position is measured in Billion USD, and TTM in days. All equations represented achieve global significance at a confidence level of 99%.

The negative coefficients for the two versions of the net position of leveraged funds on BTC futures go in line with the theory, as according to the formulated hypothesis, funds looking to collect the carry, would short the future and buy the underlying asset, the leading their net position further into negative territory. For each thousand contracts the funds sell short, the mispricing is expected to rise by 26.36 basis points, while for each Billion USD decrease in leveraged funds net position, mispricing is expected to follow with an increase of 102.59 basis points. In the most time-adjusted relevant setting, based on the regression result for equation 4.8, the impact of a notional net position decrease is similar, at 101.03 basis points, with the effect of a day expected to change the mispricing by 2.37 basis points, in line with previous estimated coefficients. In the end, the variable is statistically significant on both versions for a confidence level of 99%.

Table 15 - Regression results summary for ETH Front Quarter contracts, Net Position of LF as explanatory variable.

ETH Mispricing	4.5	4.6	4.7	4.8
Constant	0.030588	0.021319	0.010955	0.001517
	(0.007791)	(0.005723)	(0.007483)	(0.006312)
Net Position	0.008794**		0.008487**	
	(0.004284)		(0.003507)	
Notional Net Position		0.020233		0.017763
		(0.017795)		(0.014787)
TTM			0.000431***	0.000431***
			(0.0000860)	(0.0000898)
R-squared	0.079	0.026	0.396	0.342
Adjusted R-squared	0.06	0.006	0.371	0.314
Observations	51			

Standard errors are reported in parentheses. *, **, *** indicates significance at the 90%, 95%, and 99% level, respectively. Coefficients represent the absolute impact on Mispricing. Net Position is measured in number of contracts, Notional Net Position is measured in Billion USD, and TTM in days. Regressions 4.5, 4.7 and 4.8 achieve global significance at a confidence level of 95%, 99% and 99% respectively. Regression 4.6 does not achieve global significance.

When analysing the impact of the same variables on ETH contracts, results are not what was expected. All regressions establish a positive impact of net position of leveraged funds on mispricing, precisely the contrary that was observed on BTCs case, and what was expected by the literature. The lower number of observation and the fact that COTs data for ETH futures only became available 06/04/2021 when CME futures for that asset were inaugurated, might be reflected in these results. According to the regression 4.5 and 4.7, for each thousand contracts funds go long, the mispricing will increase by 87.94 basis points and 84.87 basis points respectively. The variable is significant at the confidence level of 95% in both equations. On the notional variable, regression 4.6 show the same positive sign, and imply a 202.33 basis point increase in mispricing for each addition of one Billion USD to Leveraged traders net position, while the time-adjusted equation 4.8, shows 177.63 b.p. increase in mispricing for each million dollars of additional net position. In both equations, the variable did not surpass the minimum confidence level of 90%.

5.4. Summary

The first equation 4.9 is not adjusted for time to maturity while the last one takes such variable in consideration. The notional net position was chosen in the place of the absolute net position (the variable containing the net position in number of contracts), because of the higher R-Square attributed to the notional when comparing previous regression results, at least for BTC.

Table 16 - Regression results summary for BTC Front Quarter contracts, all explanatory variables.

BTC Mispricing	4.9	4.10
Constant	-0.007386	-0.018231
	(0.002645)	(0.002618)
BTC Price	0.000232*	0.000291**
	(0.000137)	(0.000115)
Sentiment	0.000294***	0.000290***
	(0.0000434)	(0.0000363)
Notional Net Position	0.000733	0.002739
	(0.004796)	(0.004017)
TTM		0.000238***
		(0.0000308)
R-squared	0.378	0.569
Adjusted R-squared	0.364	0.556
Observations	140	

Standard errors are reported in parentheses. *, **, *** indicates significance at the 90%, 95%, and 99% level, respectively. Coefficients represent the absolute impact on Mispricing. Price is measured in thousands USD, Sentiment is measured in a scale from 1 to 100, Notional Net Position is measured in Billion USD and TTM in days. All regressions represented achieve global significance at a confidence level of 99%.

As observed during the chapter, adding TTM to the regressions always improve R-Squared, in line with what's expected in theory (Bühler & Kempf, 1995). Regression results for equations 4.9 and 4.10, show that Notional Net Position is not individually significant among the other variables. Furthermore, as coefficients are expected to be negative, and achieve individual significance when that's the case, we may conclude that such variable is more useful when considered alone. Looking further into equation 4.9, the variable that regards BTC price barely passes the minimum confidence level of 90%. On the other hand, Sentiment keeps its statistical significance in both equations, as well as TTM. In the end, equation 4.10 is the one responsible to explain the higher number of variations in the model, with an R-Square of 0.569, the highest value among all previous regressions regarding BTC. As mentioned previously, the high explanatory power of Sentiment could be observed by comparing the univariable R-Squares of 0.301 and 0.486 of equations 4.3 and 4.4 with the multivariable R-Squares of 0.378 and 0.569 of equations 4.9 and 4.10.

Table 17 - Regression results summary for ETH Front Quarter contracts, all explanatory variables.

ETH Mispricing	4.9	4.10
Constant	0.002115	-0.026930
	(0.011888)	(0.010604)
ETH Price	0.000357	0.007708*
	(0.005390)	(0.004396)
Sentiment	0.000708***	0.000612***
	(0.000107)	(0.0000849)
Notional Net Position	0.063072**	0.085259***
	(0.026543)	(0.021040)
TTM		0.000372***
		(0.0000662)
R-squared	0.496	0.701
Adjusted R-squared	0.464	0.675
Observations	51	

Standard errors are reported in parentheses. *, **, *** indicates significance at the 90%, 95%, and 99% level, respectively. Coefficients represent the absolute impact on Mispricing. Price is measured in thousands USD, Sentiment is measured in a scale from 1 to 100, Notional Net Position is measured in Billion USD and TTM in days. All regressions represented achieve global significance at a confidence level of 99%.

Regarding ETH Mispricing, the variables preserving individual significance are not the same as observed on BTC Mispricing. The price of ETH did not achieve the threshold for individual significance in equation 4.9, being significant at 90% confidence level in equation 4.10. Sentiment has shown to be robust and highly significant along the study, and that's not different considering these equations. On the other hand, Notional Net Position of leveraged funds is highly significant in equation 4.10, not expected considering the previous results.

Even though it is individually significant, the sign of the coefficient goes against what's expected. The highest R-Square among all regressions, for BTC and ETH, is obtained in applying regression 4.10 to ETH contracts, explaining 70.1% of the variance of ETH Mispricing.

6. Robustness Testing

6.1. Daily Data Granularity

Using a weekly level of granularity on the previous chapter came on the need to use and combine COTs (Commitment of Traders) data with the remaining variables. However, all the other variables are available as daily observations, which could add or subtract robustness to the results. Using equation, 4.1 and 4.2 to test the impact of BTC Price on such data structure, as well as equations 4.3 and 4.4 for the impact of Sentiment, the results are as follows in the tables 18 and 20, respectively. For ETH, results on those two variables are presented in the tables 19 and 21.

Table 18 - Regression results summary for BTC Front Quarter contracts, underlying price as explanatory variable (Daily Series).

BTC Mispricing	4.1	4.2
Constant	0.004523	-0.007734
	(0.000821)	(0.001005)
Bitcoin Price	0.000303***	0.000302***
	(0.0000249)	(0.0000218)
TTM		0.000277***
		(0.0000159)
R-squared	0.131	0.337
Adjusted R-squared	0.130	0.336
Observations	979	

Standard errors are reported in parentheses. *, **, *** indicates significance at the 90%, 95%, and 99% level, respectively. Coefficients represent the absolute impact on Mispricing. Price is measured in thousands USD, and TTM in days. All regressions represented achieve global significance at a confidence level of 99%.

The first group of regressions, that used BTC price as the explanatory variable, show very similar coefficients, as well as slightly lower R-Square, possibly because of noise of daily observations when comparing with weekly observations. All the variables keep their statistical significance on higher levels of granularity. The number of observations is different due a technicality in the sample, that does not contain the sample for the first day of BTC Index price at 00:00 GMT. Under the regression results, one can affirm that, with a 99% confidence level, that on average, a change of one thousand USD on BTC price, implies an increase of 3.03 basis points BTC Mispricing, or 3.02 basis points if adjusted to maturity.

Table 19 - Regression results summary for ETH Front Quarter contracts, underlying price as explanatory variable (Daily Series).

ETH Mispricing	4.1	4.2
Constant	0.011003	-0.001780
	(0.000864)	(0.001222)
ETH Price	0.001731***	0.001971***
	(0.000417)	(0.000381)
TTM		0.000279***
		(0.0000204)
R-squared	0.018	0.184
Adjusted R-squared	0.017	0.182
Observations	924	

Standard errors are reported in parentheses. *, **, *** indicates significance at the 90%, 95%, and 99% level, respectively. Coefficients represent the absolute impact on Mispricing. Price is measured in thousands USD, and TTM in days. All regressions represented achieve global significance at a confidence level of 99%.

On the Ethereum side, daily observations increase the significance of the variable subject to analysis. On the maturity adjusted equation, the R-squared surpasses the original one, explaining 18.4% of the dependent variable variance. Under daily observations, on average, with a 99% confidence level, it is expected that for each one thousand USD price change in ETH, mispricing is expected to increase by 17.31 basis points, or 19.71 basis points in the time adjusted regression. Those values are in line what was determined in the previous chapter, adding robustness to the results.

The following Table 20, tests the sentiment related regressions:

Table 20 - Regression results summary for BTC Front Quarter contracts, Sentiment as explanatory variable (Daily Series).

BTC Mispricing	4.3	4.4
Constant	-0.003192	-0.013593
	(0.001012)	(0.001076)
Sentiment	0.000327***	0.000306***
	(0.0000187)	(0.0000165)
TTM		0.000257***
		(0.0000150)
R-squared	0.237	0.413
Adjusted R-squared	0.236	0.412
Observations	980	

Standard errors are reported in parentheses. *, **, *** indicates significance at the 90%, 95%, and 99% level, respectively. Coefficients represent the absolute impact on Mispricing. Sentiment is measured in a scale from 1 to 100, and TTM in days. All regressions represented achieve global significance at a confidence level of 99%.

By analysing the table above, the higher number of observations implies a lower R-Squared, in part because of the noise that could embed in higher granularity samples. In the last chapter, a weekly average of the sentiment indicator was used, operation that could dilute the effect of noise along the sample. Nevertheless, a higher level of observations did not lead us to other conclusions in matters of statistical significance. The coefficients maintain the same it's positivity as expected.

Table 21 - Regression results summary for ETH Front Quarter contracts, Sentiment as explanatory variable (Daily Series).

ETH Mispricing	4.3	4.4
Constant	-0.0052	-0.015173
	(0.001120)	(0.001234)
Sentiment	0.000378***	0.000359***
	(0.0000207)	(0.0000189)
TTM		0.000246***
		(0.0000172)
R-squared	0.254	0.383
Adjusted R-squared	0.253	0.382
Observations	980	

Standard errors are reported in parentheses. *, **, *** indicates significance at the 90%, 95%, and 99% level, respectively. Coefficients represent the absolute impact on Mispricing. Sentiment is measured in a scale from 1 to 100, and TTM in days. All regressions represented achieve global significance at a confidence level of 99%.

Regarding ETH contracts, the individual significance in conserved on such level of granularity. With a 99% confidence level, it is expected that on average, for each time the sentiment index ticks up, the mispricing will increase by 3.78 basis points (compared with 3.89 b.p. on weekly observations), or considering the maturity adjusted regression, will increase by 3.59 basis points (compared with 3.86 b.p. on weekly observations).

Finally, regarding the interaction between variables, the following regressions will be used to test all the available variables on daily observations.

$$M_t = \beta_1 + \beta_2 S_t + \beta_3 Sent_t + e_t \tag{6.1}$$

$$M_t = \beta_1 + \beta_2 S_t + \beta_3 Sent_t + \beta_4 TTM_t + e_t \tag{6.2}$$

The following Table 22 presents the regression results on such equations.

Table 22 - Regression results summary for BTC Front Quarter contracts, Underlying Price and Sentiment as explanatory variables (Daily Series).

BTC Mispricing	6.1	6.2
Constant	-0.006911	-0.017530
	(0.001049)	(0.001074)
BTC Price	0.000216***	0.000222***
	(0.0000231)	(0.0000199)
Sentiment	0.000284***	0.000262***
	(0.0000185)	(0.0000160)
TTM		0.000260***
		(0.0000141)
R-squared	0.299	0.479
Adjusted R-squared	0.298	0.478
Observations	979	

Standard errors are reported in parentheses. *, **, *** indicates significance at the 90%, 95%, and 99% level, respectively. Coefficients represent the absolute impact on Mispricing. Price is measured in thousands USD, Sentiment is measured in a scale from 1 to 100 and TTM in days. All regressions represented achieve global significance at a confidence level of 99%.

Both regression results for equations 6.1 and 6.2 support the individual significance of the variables as well as the global significance of multivariable models. By comparing regression results on BTC, it fair to conclude that conclusions obtained on the previous chapter are robust enough to hold on a higher level of granularity, while being subjective to higher levels of noise.

Table 23 - Regression results summary for ETH Front Quarter contracts, Underlying Price and Sentiment as explanatory variables (Daily Series).

ETH Mispricing	6.1	6.2
Constant	-0.007571	-0.018415
	(0.001291)	(0.001375)
ETH Price	0.001725***	0.001947***
	(0.000361)	(0.000325)
Sentiment	0.000378***	0.000359***
	(0.0000214)	(0.0000193)
TTM		0.000258***
		(0.0000175)
R-squared	0.266	0.406
Adjusted R-squared	0.265	0.405
Observations	924	

Standard errors are reported in parentheses. *, **, *** indicates significance at the 90%, 95%, and 99% level, respectively. Coefficients represent the absolute impact on Mispricing. Price is measured in thousands USD, Sentiment is measured in a scale from 1 to 100 and TTM in days. All regressions represented achieve global significance at a confidence level of 99%.

The scenario for ETH is no different. Under daily granularity, Price and Sentiment are variables of individual significance in every regression. On the multivariable equation, the same results hold regarding significance. When compared with regression results based on weekly observations, ETH price coefficients are accepted at the highest confidence level of 99% instead of 95%. It is worth pointing out the expected consistency of the coefficient of TTM on every equation on this subchapter 6.1., ranging between 2.46 basis points and 2.79 basis points.

6.2. Periodic Samples

By looking at table 1 and 3 on the appendix, there are three quarters that stand out from the others, as during that time, the number of observations that surpassed the hurdle rate of 2.5% is significant. In order to test how variables behave in different periods we divided the sample, considering the following three sample periods: (1) from 20/07/2019 to 25/12/2020, representing a relatively calm phase were the mispricing surpassed the hurdle rate for 25 times; (2) 26/12/2020 to 25/06/2021, which could be called the golden age of mispricing, with a total 96 mispricing observations above the hurdle rate; (3) and the final portion of the sample, from 26/06/2021 until 25/03/2022, with 20 observations surpassing the hurdle rate.

The analysis will test the individual robustness of the three variables subject to the study as well as the global significance of a multivariate equation, both TTM adjusted.

The results for Sentiment indicator are present in table 24 and 25, according to equation 4.2.

Table 24 - Regression results summary for BTC Front Quarter contracts, Sentiment as explanatory variable (divided by periods).

BTC Mispricing	(1)	(2)	(3)
Constant	-0.007133	-0.008822	-0.010239
	(0.002396)	(0.008140)	(0.002012)
Sentiment	0.000224***	0.000256*	0.000194***
	(0.0000363)	(0.000136)	(0.0000335)
TTM	0.000118***	0.000457***	0.000280***
	(0.0000296)	(0.000130)	(0.0000264)
R-squared	0.399	0.569	0.796
Adjusted R-squared	0.382	0.531	0.785
Observations	75	26	39

Standard errors are reported in parentheses. *, **, *** indicates significance at the 90%, 95%, and 99% level, respectively. Coefficients represent the absolute impact on Mispricing. Sentiment is measured in a scale from 1 to 100, and TTM in days. All regressions represented achieve global significance at a confidence level of 99%.

Both variables maintain it's individual significance, only in the middle period, sentiment is only considered significant at a 90% confidence level. The global significance for each period is assured as well, with crescent Adjusted R-Squares over time. Differences on estimators represent the sensitivity of mispricing to those variables. On the period (2), for each point increase on the sentiment index, the mispricing would increase, on average, 2.56 basis points, when compared to an average increase of 1.94 b.p. in the period (3). As explained before the TTM could be interpreted as the value of time on mispricing. During the period (2), on average, the mispricing would decrease by 4.57 b.p. per day, while on the period (1), the value of a day was 1.18 basis points.

Table 25 exhibit how the same variables explain the mispricing observed in ETH contracts.

Table 25 - Regression results summary for ETH Front Quarter contracts, Sentiment as explanatory variable (divided by periods).

ETH Mispricing	(1)	(2)	(3)
Constant	-0.007081	-0.008128	-0.011396
	(0.003231)	(0.008783)	(0.002053)
Sentiment	0.000250***	0.000237	0.000225***
	(0.0000487)	(0.000147)	(0.0000341)
TTM	0.0000824***	0.000561***	0.000266***
	(0.0000386)	(0.000141)	(0.0000269)
R-squared	0.311	0.596	0.789
Adjusted R-squared	0.289	0.561	0.778
Observations	67	26	39

Standard errors are reported in parentheses. *, **, *** indicates significance at the 90%, 95%, and 99% level, respectively. Coefficients represent the absolute impact on Mispricing. Sentiment is measured in a scale from 1 to 100, and TTM in days. All regressions represented achieve global significance at a confidence level of 99%.

Under table 25, TTM is significant at the 99% confidence level, in every period, as it was during all previous regressions. Interestingly, the variable sentiment loses individual significance in the middle period, as happened on the equivalent regression for BTC. Across periods, the sentiment indicator seems to be less volatile than TTM, indicator that maintained relatively closed coefficient when compared across different regressions applied to the total sample.

The results for BTC Price are present in Table 26.

Table 26 - Regression results summary for BTC Front Quarter contracts, Underlying Price as explanatory variable (divided by periods).

BTC Mispricing	(1)	(2)	(3)
Constant	-0.003495	-0.04239	-0.025559
	(0.002259)	(0.011655)	(0.003721)
Bitcoin Price	0.000934***	0.000958***	0.000502***
	(0.0000289)	(0.000231)	(0.0000738)
TTM	0.000121***	0.000596***	0.000287***
	(0.0000351)	(0.0000911)	(0.0000243)
R-squared	0.197	0.716	0.828
Adjusted R-squared	0.175	0.691	0.818
Observations	75	26	39

Standard errors are reported in parentheses. *, **, *** indicates significance at the 90%, 95%, and 99% level, respectively. Coefficients represent the absolute impact on Mispricing. Price is measured in thousands USD, and TTM in days. All regressions represented achieve global significance at a confidence level of 99%.

The robustness of BTC Price as explanatory variable holds significant for all periods. It is interesting to note that during the period (1) and (2), the coefficient of for BTC Price was 0.000934 and 0.000958 respectively, while in period (3), it was 0.000502, almost half of observed on the other two periods. It means that on average, when BTC Price increased by $1000, the mispricing would increase 9.34 b.p. and 9.58 b.p. during (1) and (2), while increasing 5.02 b.p. in period (3). The value of time when measured on mispricing varied between all periods, nonetheless, keeping its statistical significance at 99% confidence level.

Table 27 - Regression results summary for ETH Front Quarter contracts, Underlying Price as explanatory variable (divided by periods).

ETH Mispricing	(1)	(2)	(3)
Constant	-0.0000454	-0.010418	-0.022435
	(0.003769)	(0.013073)	(0.004177)
Ethereum Price	0.020113**	0.004988	0.005860***
	(0.009428)	(0.004662)	(0.001088)
TTM	0.0000716	0.000724***	0.000296***
	(0.0000454)	(0.000132)	(0.0000306)
R-squared	0.082	0.571	0.743
Adjusted R-squared	0.053	0.534	0.729
Observations	67	26	39

Standard errors are reported in parentheses. *, **, *** indicates significance at the 90%, 95%, and 99% level, respectively. Coefficients represent the absolute impact on Mispricing. Price is measured in thousands USD, and TTM in days. All regressions represented achieve global significance at a confidence level of 99%.

In chapter 5, when using equations 4.1 and 4.2, the coefficients attributed to the price of the underlying asset were only significant on a 90% confidence level and 95% level respectively. In line with the previous variable, in period 2, the underlying price is not statistically significant at the minimum confidence level of 90%. For the first time, among all previous regressions, TTM is not statistically significant at the minimum level of 90% in period 1. The different number of observations between periods limit the direct comparison of different regressions, however, it is possible to observe consistency on the signs of the coefficients.

The results for Notional Net Position of Leveraged Funds are presented in Table 28.

Table 28 - Regression results summary for BTC Front Quarter contracts, Notional Net Position as explanatory variable (divided by periods).

BTC Mispricing	(1)	(2)	(3)
Constant	-0.001875	-0.022425	-0.009586
	(0.002125)	(0.007638)	(0.002581)
Notional Net Position	-0.017129**	-0.031592***	-0.006847***
	(0.007624)	(0.007764)	(0.001767)
TTM	0.000110***	0.000457***	0.000282***
	(0.0000360)	(0.0000967)	(0.0000308)
R-squared	0.140	0.711	0.723
Adjusted R-squared	0.116	0.686	0.707
Observations	75	26	39

Standard errors are reported in parentheses. *, **, *** indicates significance at the 90%, 95%, and 99% level, respectively. Coefficients represent the absolute impact on Mispricing. Notional Net Position is measured in Billion USD, and TTM in days. All equations represented achieve global significance at a confidence level of 99%.

Finally, the Notional Net Position of Levered funds shows that previous conclusions about the statistical significance of the variable are robust, even though the coefficients imply some differences on the elasticity of mispricing to changes in Notional Net Position. On average, each time that Leveraged Fund decreased their net position on Bitcoin futures by $1Bn, the mispricing would increase by 3.16 p.p. in the period of exuberance (2), with more modest values of 1.71 p.p. for period (1) and 0.68 p.p. for period (3).

Table 29 - Regression results summary for ETH Front Quarter contracts, Notional Net Position as explanatory variable (divided by periods).

ETH Mispricing	(2)	(3)
Constant	0.007534	-0.012186
	(0.012255)	(0.002110)
Notional Net Position	0.155044**	-0.031120***
	(0.068493)	(0.004640)
TTM	0.001121***	0.000267***
	(0.000110)	(0.0000266)
R-squared	0.952	0.794
Adjusted R-squared	0.942	0.783
Observations	12	39

Standard errors are reported in parentheses. *, **, *** indicates significance at the 90%, 95%, and 99% level, respectively. Coefficients represent the absolute impact on Mispricing. Notional Net Position is measured in Billion USD, and TTM in days. All equations represented achieve global significance at a confidence level of 99%.

Previous regressions regarding the net position of leveraged funds and its relationship with ETH were not in line with the ones observed with BTC, as the sign of the coefficient was positive. In the table 29, it is possible to observe two different signs in two different periods. Being those two statistically significant (at 95% confidence level for period two and 99% confidence level for period 3), caution in interpreting results is needed, mainly because of the lower number observations for the net position of levered traders.

To sum up, all variables seem to be statistically significant in the periods mentioned above, implying robustness of the results obtained on the previous chapter regarding BTC contracts. Considering the results obtained for ETH contracts, one can doubt the coefficients estimate for the period number two as neither sentiment nor underlying price achieved individual significance for the confidence level of 99%.

7. Conclusion and Suggestions for Future Research

There are a quite few cycles in the economy, each one responding to certain variables, with some variables being shared across multiple asset cycles. This thesis adds solid foundations to explain the cryptocurrency cycle more profoundly. The main variable of this study, the mispricing of future contracts based on the cost of carry framework, is for some periods, undoubtedly different from zero, the expected value to imply a non-arbitrage condition. The evidence holds way stronger for the upside than for the downside, as when negative, contracts trade with very little discount, of course non considering some punctual outliers; while premiums are the norm for some contracts, especially in times of exuberance.

When compared to the base studies used for this dissertation, mainly on cost-of-carry for indexes, there two great differences. The first one is that indexes future contracts were chronically undervalued, in the case of German DAX (Bühler & Kempf, 1995), the Greek ATHEX (Fassas, 2010), and the Spanish IBEX (Verdasco, 2017), comparing with approximately 88% and 90% overvaluation for both contracts. Even in studies where that does not happen, for example considering the FTSE 250, we see maximum discounts at 1.24% and maximum premiums of 0.948% (Butterworth & Holmes, 2000). These values are way lower than mispricing values observed in this dissertation.

In accordance with the empirical results, sentiment is indeed on the main factors causing the future contracts to be mispriced. There is an interesting result on the period 2, on testing the robustness of this indicator, as for both contracts, sentiment loses its individual significance. Such particularity could be subject to further research. When participants in the market were optimistic, contracts deviated the most from what was expected from the cost-and-carry framework. One might argue that the model is not adequate for crypto assets, and model risk is a very high threat for every novelty in finance. However, the cost-of-carry model is not only theoretical, as it is used as the base of the arbitrage strategy cash-and-carry, the transformation of variables into a series of financial operations.

Then, another variable that could help traders and investors to understand where they are on the crypto market cycle, is the net position of leveraged funds as reported by the CFTC on the weekly released datapoint Commitment of Traders. Such funds are most short when the sentiment is very high and consequently, when the futures observable prices imply a consistent and meaningful premium over the theoretical price. The value added by this indicator could hint for their position in futures, as to execute the cash-and-carry strategy, the one that would be short on the future contract while longing the underlying asset. Such logic is observed on BTC contracts as well as verified in the last period of analysis for ETH. The opposite unexpected relationship on the empirical results and in the first period of analysis for ETH in the following chapter might be justified by the initial development of the market for that asset. Such situation and its justifications could be subject to future research.

The ad hoc analysis of differences between index price and spot price reassures those differences between both parameters are not a justification for the dismissal of strategies based on the profiting from mispricing, as both values follow each other very closely.

In the end, everyday non-institutional traders had access to excess returns topping 11.39% for front-quarter cash-and-carry plays, and 29.82% for longer arbitrage setups, during the studied period. The only barrier between them and those returns was knowledge itself.

The data reflects just one point in time for each day. More granularity on data could show special situations such has opportunities where the mispricing between the theoretical price and current price is higher than presented, as well as situations where the mispricing was much and, by consequence, a good moment to realize profits before expiration. Another consequence of low granularity data is reflected on Spot Premium, that in certain moments, can reach way higher values than the ones reflected on Table 7 and 8, and those only measure the indicator for on moment in time daily.

This study ignores liquidity, meaning that at a large scale and in environments of sudden and unexpected low volume, the price impact could render the arbitrage strategy less profitable than expected.

Some exchanges offer futures with different specifications, with FTX accepting BTC and ETH as collateral, while others only accept the quote currency as margins. Such differences have influence on the capital efficiency of each contract. The more efficient a contract is, the less will be paid in non-core trading costs, as for example, financing costs of borrowed money to maintain margin, or the nearly null chance of liquidation when using the nominal contract exposure of the underlying asset as collateral. It would be interesting to test if more efficient contracts trade at different prices when compared with less capital efficient contracts as suggested by section 5 of the literature review.

The counterparty of the contract is important as well, even when a compensation chamber stands in the middle to prevent a peer-to-peer default. However, there are many different exchanges, and perceived risk in certain times of distress might dislocate identical contract quotations just because of the platform it is being traded in. Testing the correlation between perceived default of one platform and its prices for identical contract would certainly add value to the literature.

Bibliography

Abreu, D., & Brunnermeier, M. K. (2002). Synchronization risk and delayed arbitrage. *Journal of Financial Economics*, *66*(2–3), 341–360. https://doi.org/10.1016/S0304-405X(02)00227-1

Acharya, V., & Bisin, A. (2014). Counterparty risk externality: Centralized versus over-the-counter markets. *Journal of Economic Theory*, *149*(1), 153–182. https://doi.org/10.1016/j.jet.2013.07.001

Akyildirim, E., Corbet, S., Katsiampa, P., Kellard, N., & Sensoy, A. (2020). The development of Bitcoin futures: Exploring the interactions between cryptocurrency derivatives. *Finance Research Letters*, *34*. https://doi.org/10.1016/j.frl.2019.07.007

Alizadeh, A. H., & Nomikos, N. K. (2004). Cost of carry, causality and arbitrage between oil futures and tanker freight markets. *Transportation Research Part E: Logistics and Transportation Review*, *40*(4), 297–316. https://doi.org/10.1016/j.tre.2004.02.002

Black, F. (1986). Noise. *Journal of Finance*, *41*(3), 529–543. https://EconPapers.repec.org/RePEc:bla:jfinan:v:41:y:1986:i:3:p:529-43

Brauneis, A., Mestel, R., Riordan, R., & Theissen, E. (2021). How to measure the liquidity of cryptocurrency markets? *Journal of Banking and Finance*, *124*. https://doi.org/10.1016/j.jbankfin.2020.106041

Bühler, W., & Kempf, A. (1995). Dax index futures: Mispricing and arbitrage in German markets. *Journal of Futures Markets*, *15*(7), 833–859. https://doi.org/10.1002/fut.3990150706

Butterworth, D., & Holmes, P. (2000). Mispricing in stock index futures contracts: evidence for the FTSE 100 and FTSE mid 250 contracts. *Applied Economics Letters*, *7*(12), 795–801. https://doi.org/10.1080/135048500444822

Corbet, S., Lucey, B., Peat, M., & Vigne, S. (2018). Bitcoin Futures—What use are they? *Economics Letters*, *172*, 23–27. https://doi.org/10.1016/j.econlet.2018.07.031

de Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990). Noise Trader Risk in Financial Markets. *Journal of Political Economy*, *98*(4), 703–738. https://doi.org/10.1086/261703

Duffie, D., & Zhu, H. (2011). Does a Central Clearing Counterparty Reduce Counterparty Risk? *Review of Asset Pricing Studies*, *1*(1), 74–95. https://doi.org/10.1093/rapstu/rar001

Fama, E. F. (1965). Random Walks in Stock Market Prices. *Financial Analysts Journal*, *21*(5), 55–59. http://www.jstor.org/stable/4469865

Fama, M., Fumagalli, A., & Lucarelli, S. (2019). Cryptocurrencies, Monetary Policy, and New Forms of Monetary Sovereignty. *International Journal of Political Economy*, *48*(2), 174–194. https://doi.org/10.1080/08911916.2019.1624318

Fantazzini, D., & Calabrese, R. (2021). Crypto Exchanges and Credit Risk: Modeling and Forecasting the Probability of Closure. *Journal of Risk and Financial Management*, *14*(11), 516. https://doi.org/10.3390/jrfm14110516

Fassas, A. (2010). Mispricing in Stock Index Futures Markets – the Case of Greece. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1873949

Geman, H., & Smith, W. O. (2013). Theory of storage, inventory and volatility in the LME base metals. *Resources Policy*, *38*(1), 18–28. https://doi.org/10.1016/j.resourpol.2012.06.014

Hu, Y., Hou, Y. G., & Oxley, L. (2020). What role do futures markets play in Bitcoin pricing? Causality, cointegration and price discovery from a time-varying perspective? *International Review of Financial Analysis*, *72*. https://doi.org/10.1016/j.irfa.2020.101569

Jarrow, R. A., & Yu, F. (2001). Counterparty Risk and the Pricing of Defaultable Securities. *The Journal of Finance*, *56*(5), 1765–1799. https://doi.org/10.1111/0022-1082.00389

Lamont, O. A., & Thaler, R. H. (2003). Anomalies: The Law of One Price in Financial Markets. *Journal of Economic Perspectives*, *17*(4), 191–202. https://doi.org/10.1257/089533003772034952

Larson, D. F. (1994). *Copper and the negative price of storage*.

Makarov, I., & Schoar, A. (2018). Trading and Arbitrage in Cryptocurrency Markets. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3171204

Miller, M. H., Muthuswamy, J., & Whaley, R. E. (1994). American Finance Association Mean Reversion of Standard &Poor's 500 Index Basis Changes: Arbitrage-Induced or Statistical Illusion? In *Source: The Journal of Finance* (Vol. 49, Issue 2).

Pontiff, J. (1996). Costly Arbitrage: Evidence from Closed-End Funds. *The Quarterly Journal of Economics*, *111*(4), 1135–1151. https://doi.org/10.2307/2946710

Richie, N., Daigler, R. T., & Gleason, K. C. (2008). The limits to stock index arbitrage: Examining S&P 500 futures and SPDRS. *Journal of Futures Markets*, *28*(12), 1182–1205. https://doi.org/10.1002/fut.20365

Rosenthal, L. (1990). The seemingly anomalous price behavior of Royal Dutch/Shell and Unilever N.V./PLC. *Journal of Financial Economics*, *26*(1), 123–141. https://doi.org/10.1016/0304-405X(90)90015-R

Sanchez Verdasco, J. (2017). Stock Market Efficiency under the Cost of Carry Model. Evidence from the Spanish Market. *SSRN Electronic Journal*, 1–21. https://doi.org/10.2139/ssrn.2824670

Shleifer, A., & Vishny, R. W. (1997). The limits of arbitrage. *Journal of Finance*, *52*(1), 35–55. https://doi.org/10.1111/j.1540-6261.1997.tb03807.x

Stambaugh, R. F., Yu, J., & Yuan, Y. (2012). The short of it: Investor sentiment and anomalies. *Journal of Financial Economics*, *104*(2), 288–302. https://doi.org/10.1016/j.jfineco.2011.12.001

Appendix

Table 30 - Period specification for front-quarter contract analysis.

Contract	Start Date	End Date
BTC/ETH-20190927	20/07/2019	27/09/2019
BTC/ETH-20191227	28/09/2019	27/12/2019
BTC/ETH-20200327	28/12/2019	27/03/2020
BTC/ETH-20200626	28/03/2020	26/06/2020
BTC/ETH-20200925	27/06/2020	25/09/2020
BTC/ETH-20201225	26/09/2020	25/12/2020
BTC/ETH-20210326	26/12/2020	26/03/2021
BTC/ETH-20210625	27/03/2021	25/06/2021
BTC/ETH-20210924	26/06/2021	24/09/2021
BTC/ETH-20211231	25/09/2021	31/12/2021
BTC/ETH-20220325	01/01/2022	25/03/2022

Figure 5- Daily Distribution of BTC Spot Premium over Index Price.

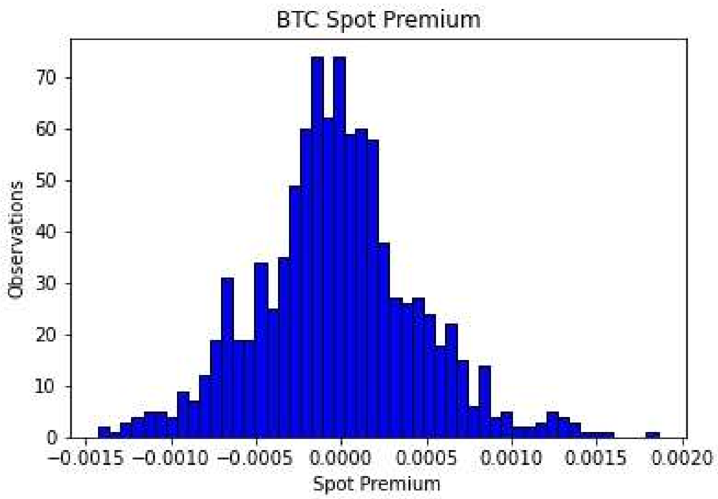


Figure 6 - Daily Distribution of ETH Spot Premium over Index Price.

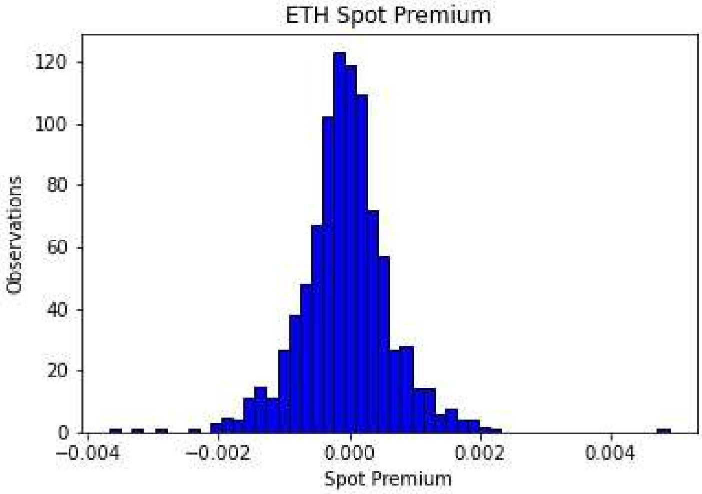


Figure 7 - Evolution of Net Position of Leveraged Funds (measured in number of contracts) and Front Quarter Mispricing for ETH Contracts.

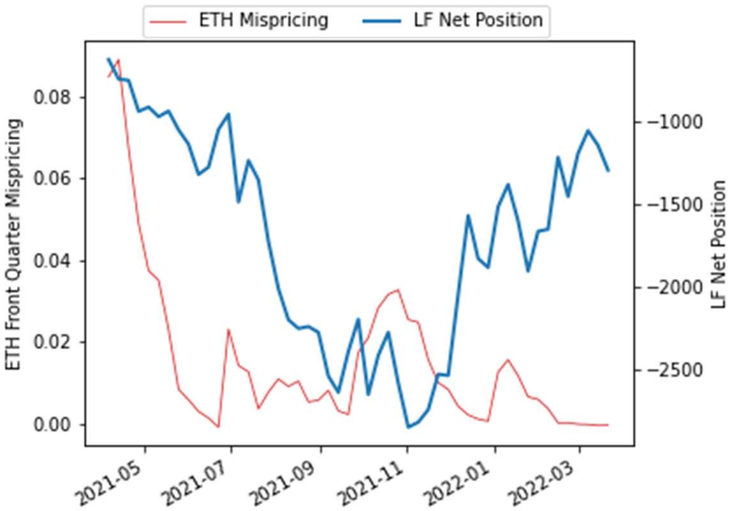


Figure 8 - Evolution of Net Position of Leveraged Funds (measured in Million USD) and Front Quarter Mispricing for ETH Contracts.

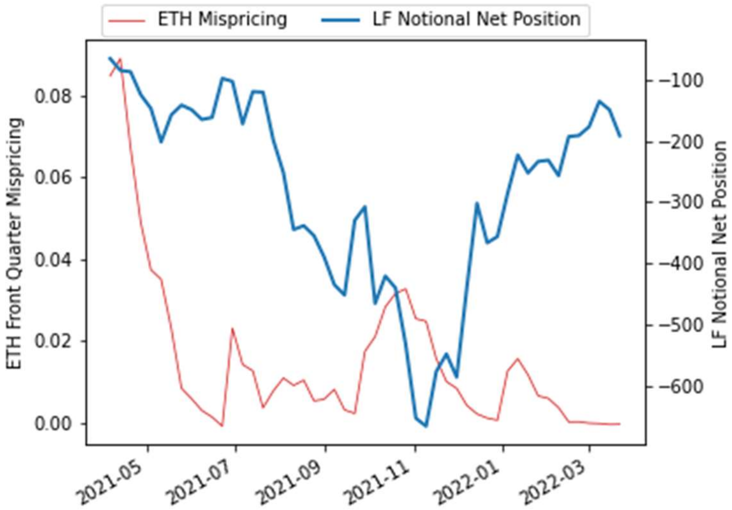


Figure 9 - Distribution of Mispricing regarding the period represented by BTC-20190927 (Restricted to Front Quarter Observations).

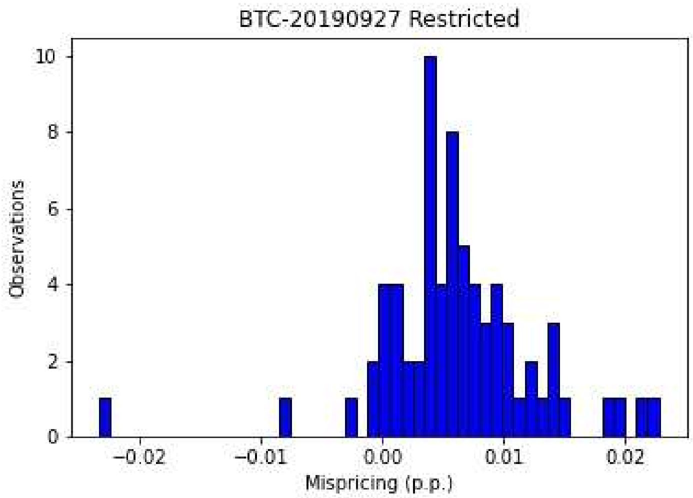


Figure 10 - Distribution of Mispricing regarding the period represented by BTC-20191227 (Restricted to Front Quarter Observations).

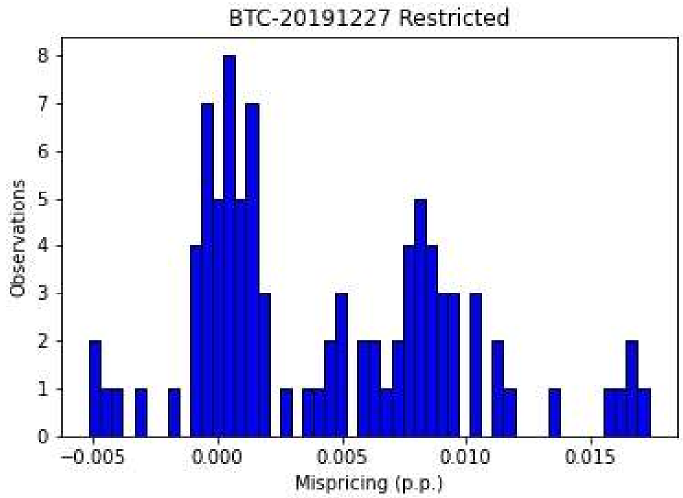


Figure 11 - Distribution of Mispricing regarding the period represented by BTC-20200327 (Restricted to Front Quarter Observations).

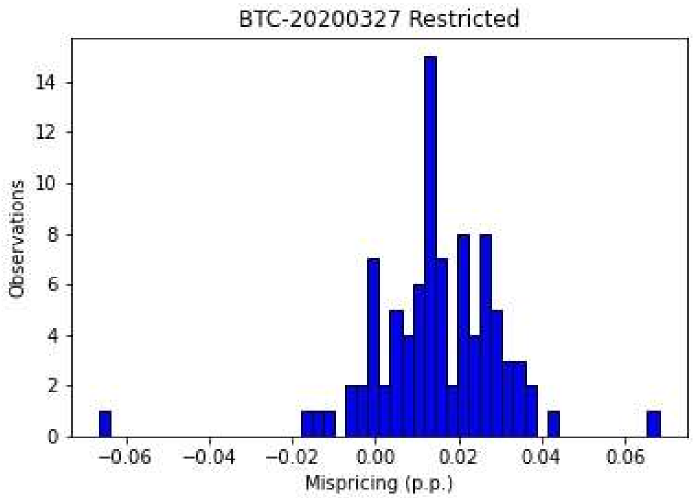


Figure 12 - Distribution of Mispricing regarding the period represented by BTC-20200626 (Restricted to Front Quarter Observations).

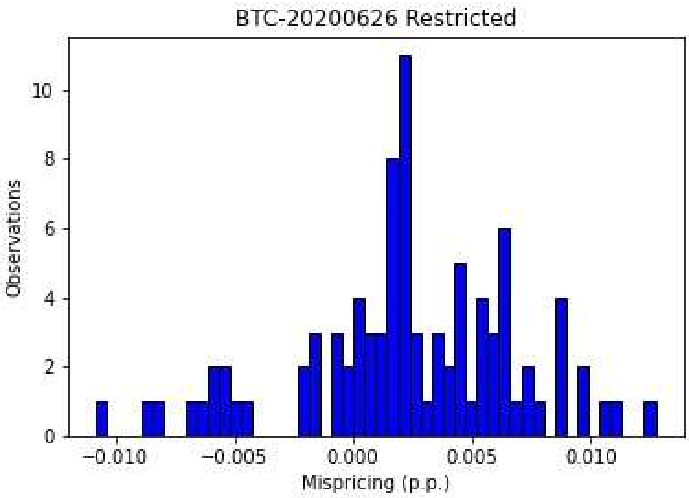


Figure 13 - Distribution of Mispricing regarding the period represented by BTC-20200925 (Restricted to Front Quarter Observations).

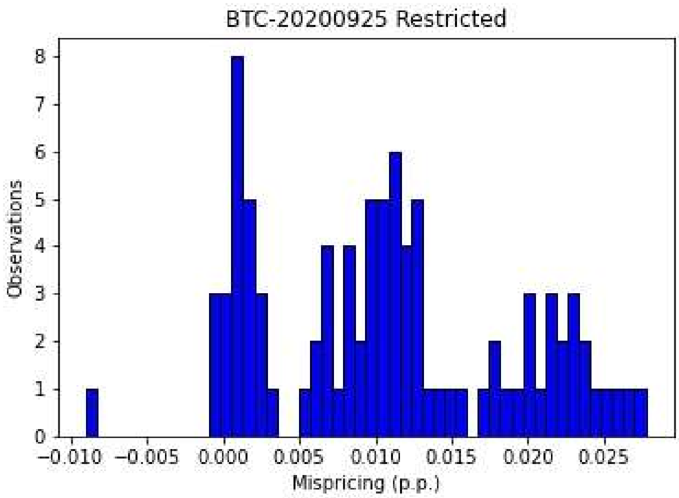


Figure 14 - Distribution of Mispricing regarding the period represented by BTC-20201225 (Restricted to Front Quarter Observations).

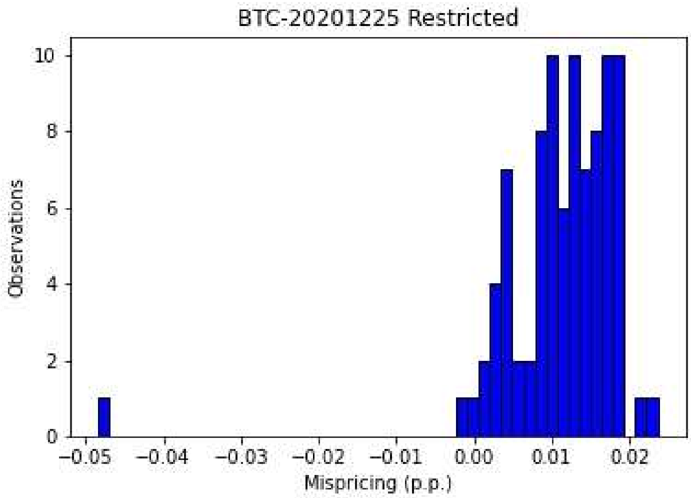


Figure 15 - Distribution of Mispricing regarding the period represented by BTC-20210326 (Restricted to Front Quarter Observations).

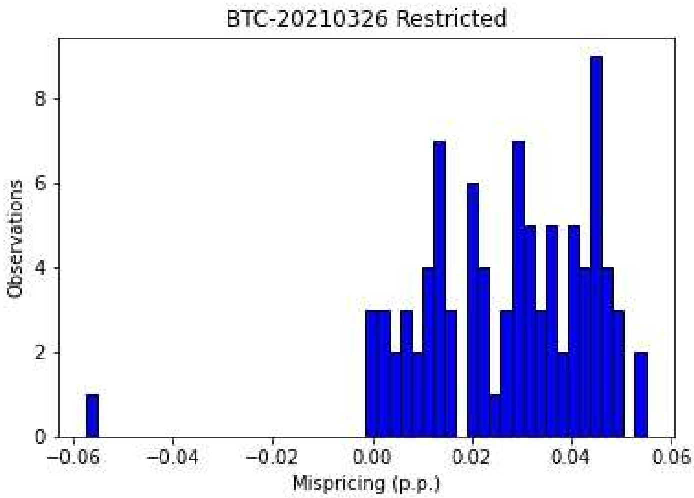


Figure 16 - Distribution of Mispricing regarding the period represented by BTC-20210625 (Restricted to Front Quarter Observations).

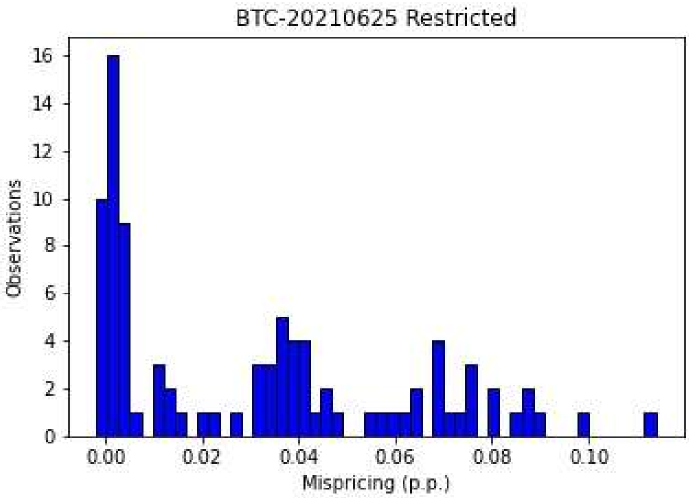


Figure 17 - Distribution of Mispricing regarding the period represented by BTC-20210924 (Restricted to Front Quarter Observations).

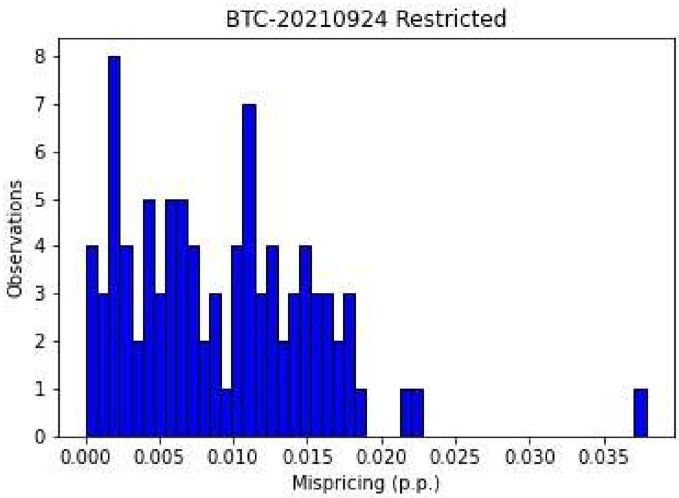


Figure 18 - Distribution of Mispricing regarding the period represented by BTC-20211231 (Restricted to Front Quarter Observations).

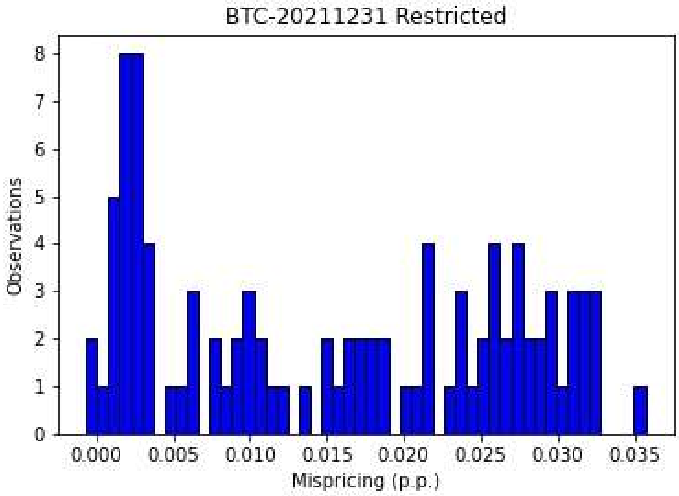


Figure 19 - Distribution of Mispricing regarding the period represented by BTC-20220325 (Restricted to Front Quarter Observations).

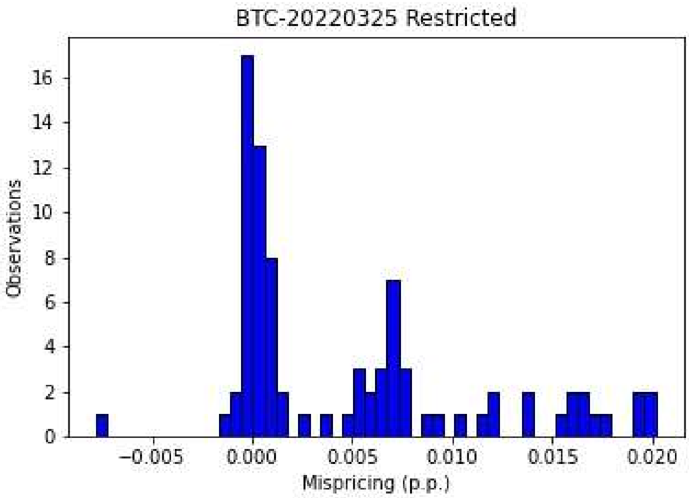


Figure 20 - Distribution of Mispricing regarding the period represented by BTC-20190927 (All Observations).

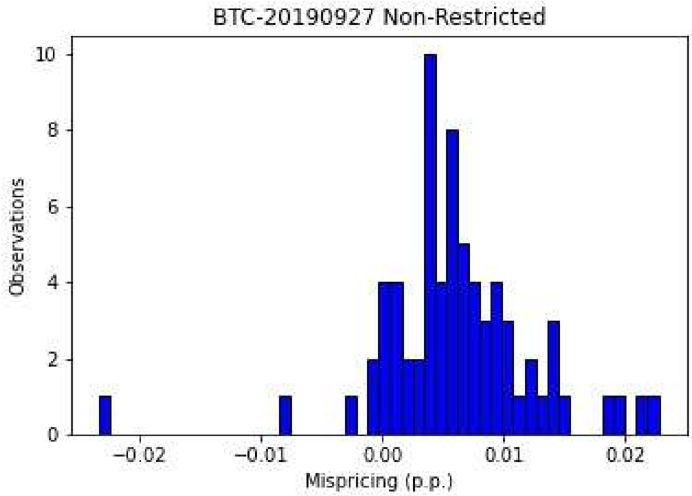


Figure 21 - Distribution of Mispricing regarding the period represented by BTC-20191227 (All Observations).

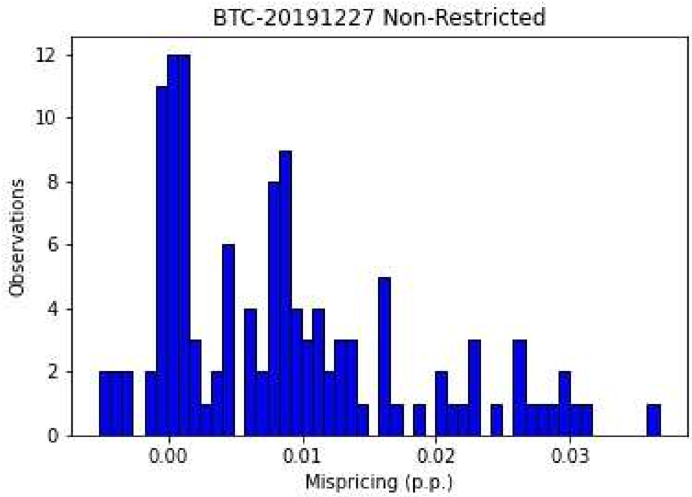


Figure 22 - Distribution of Mispricing regarding the period represented by BTC-20200327 (All Observations).

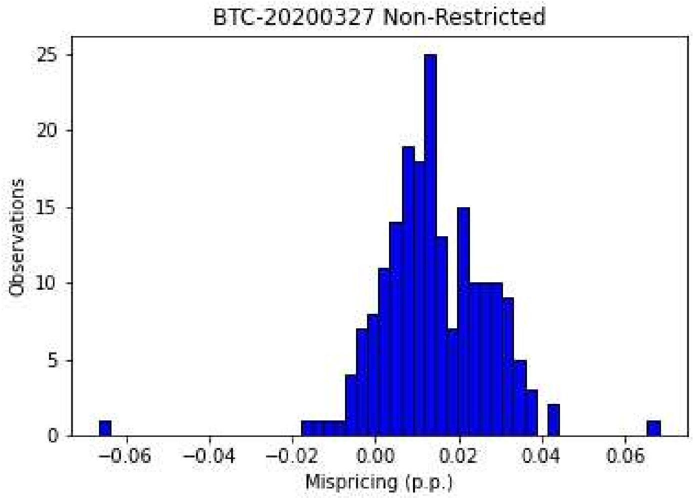


Figure 23 - Distribution of Mispricing regarding the period represented by BTC-20200626 (All Observations).

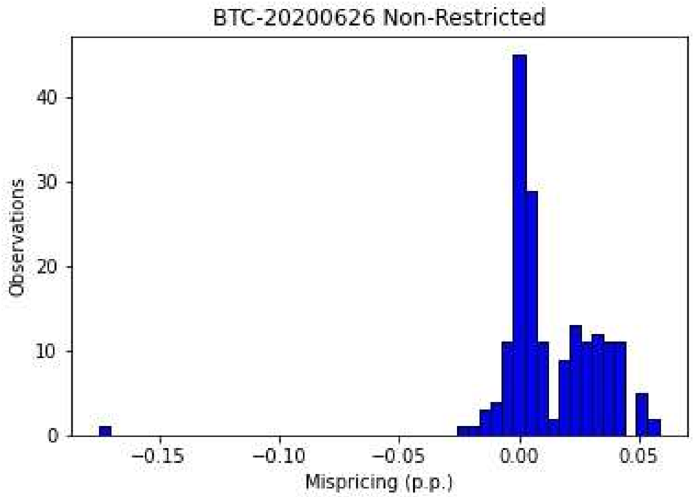


Figure 24 - Distribution of Mispricing regarding the period represented by BTC-20200925 (All Observations).

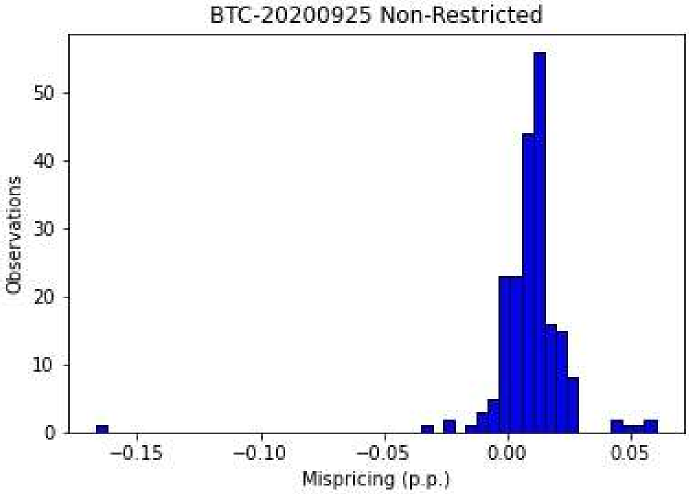


Figure 25 - Distribution of Mispricing regarding the period represented by BTC-20201225 (All Observations).

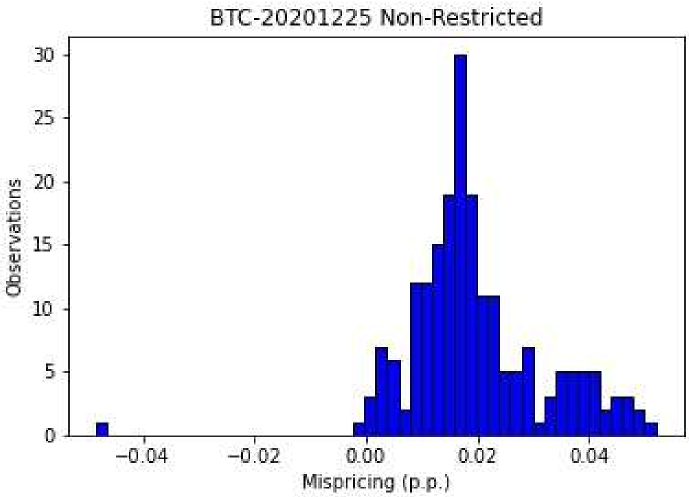


Figure 26 - Distribution of Mispricing regarding the period represented by BTC-20210326 (All Observations).

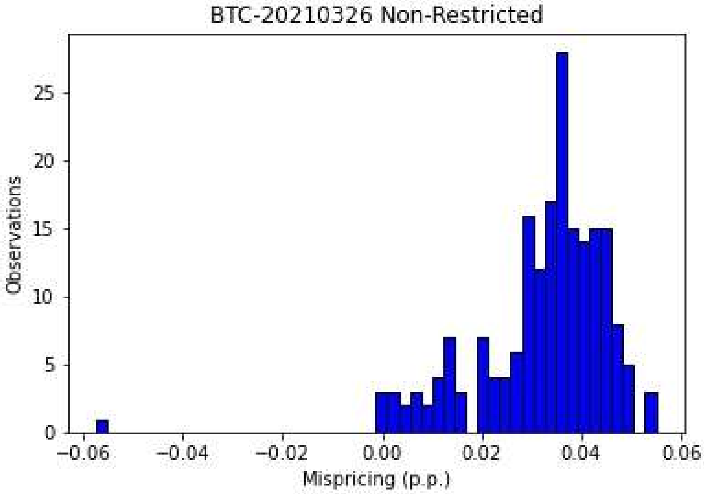


Figure 27 - Distribution of Mispricing regarding the period represented by BTC-20210625 (All Observations).

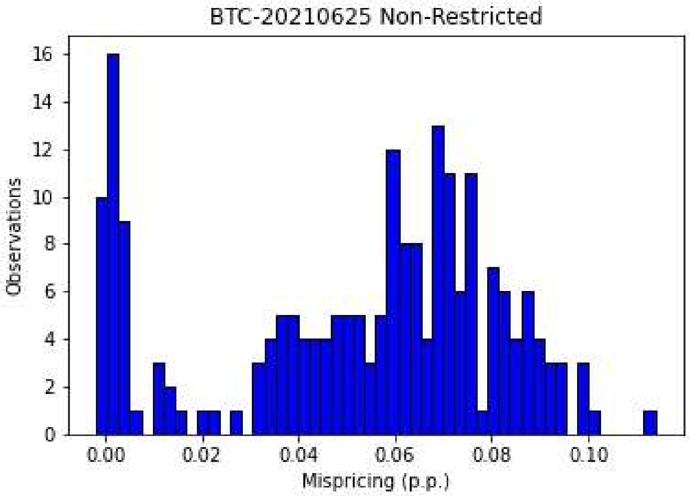


Figure 28 - Distribution of Mispricing regarding the period represented by BTC-20210924 (All Observations).

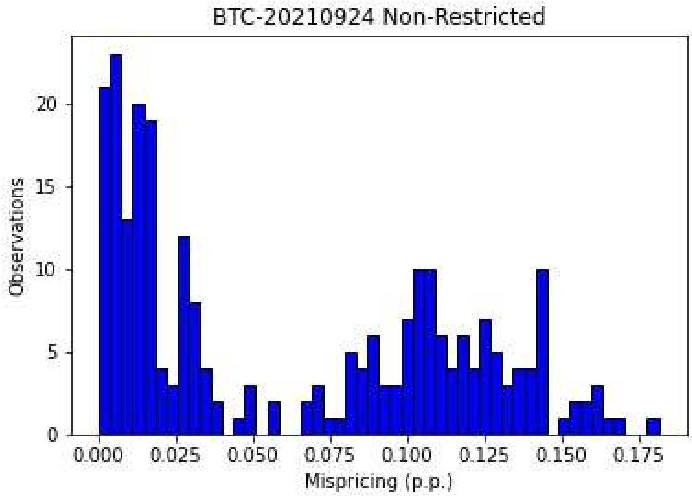


Figure 29 - Distribution of Mispricing regarding the period represented by BTC-20211231 (All Observations).

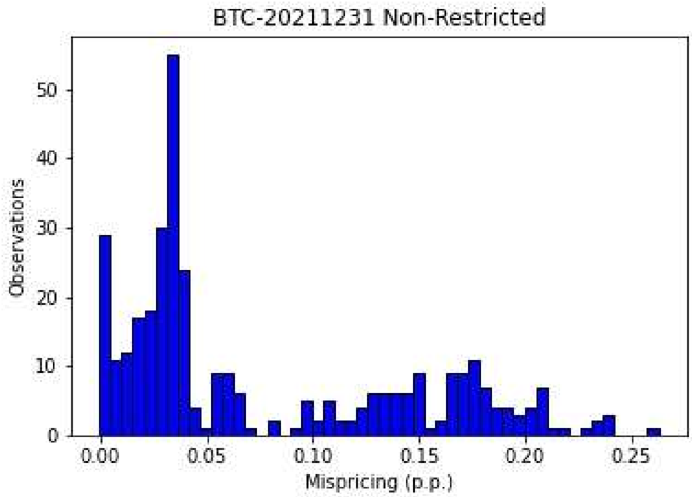


Figure 30 - Distribution of Mispricing regarding the period represented by BTC-20220325 (All Observations).

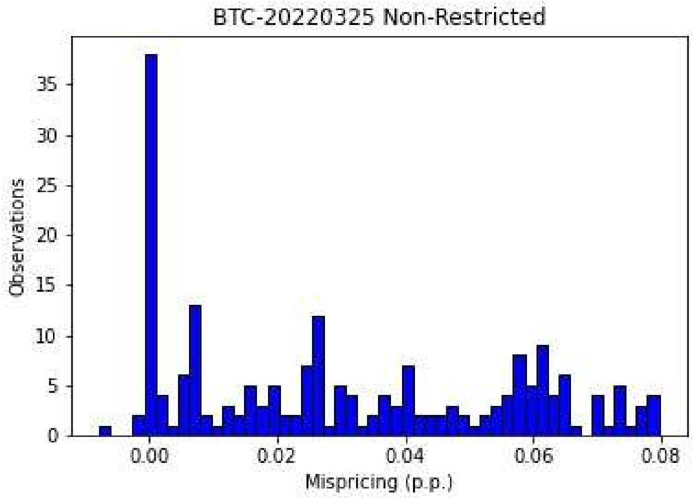


Figure 31 - Distribution of Mispricing regarding the period represented by ETH-20190927 (Restricted to Front Quarter Observations).

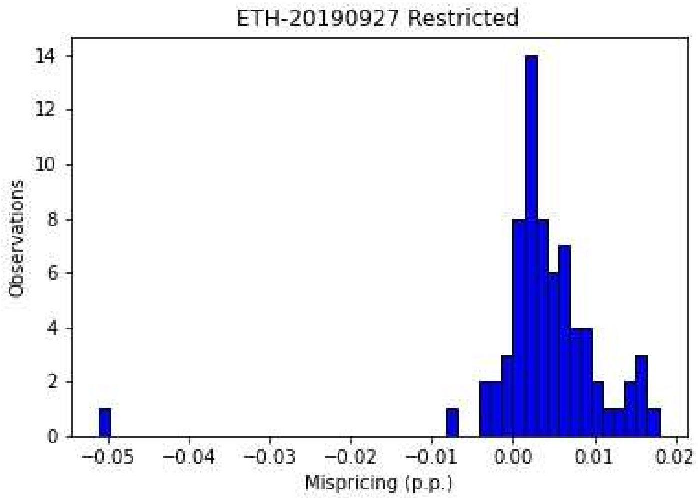


Figure 32 - Distribution of Mispricing regarding the period represented by ETH-20191227 (Restricted to Front Quarter Observations).

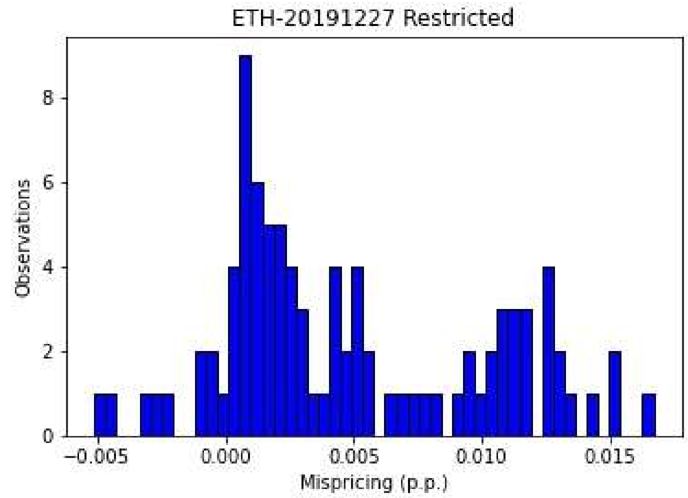


Figure 33 - Distribution of Mispricing regarding the period represented by ETH-20200327 (Restricted to Front Quarter Observations).

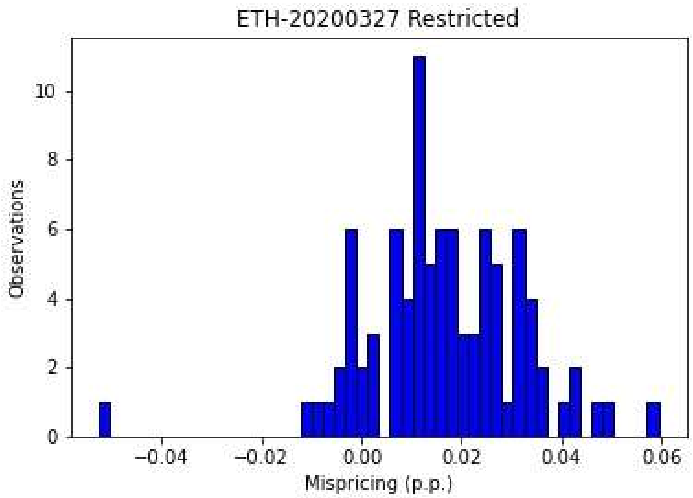


Figure 34 - Distribution of Mispricing regarding the period represented by ETH-20200626 (Restricted to Front Quarter Observations).

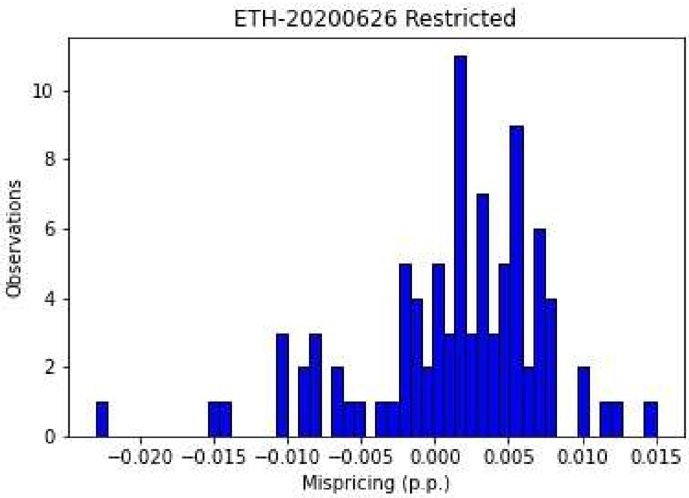


Figure 35 - Distribution of Mispricing regarding the period represented by ETH-20200925 (Restricted to Front Quarter Observations).

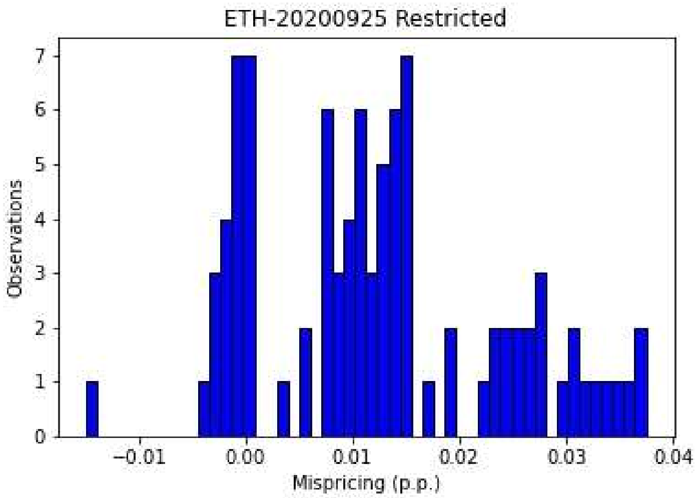


Figure 36 - Distribution of Mispricing regarding the period represented by ETH-20201225 (Restricted to Front Quarter Observations).

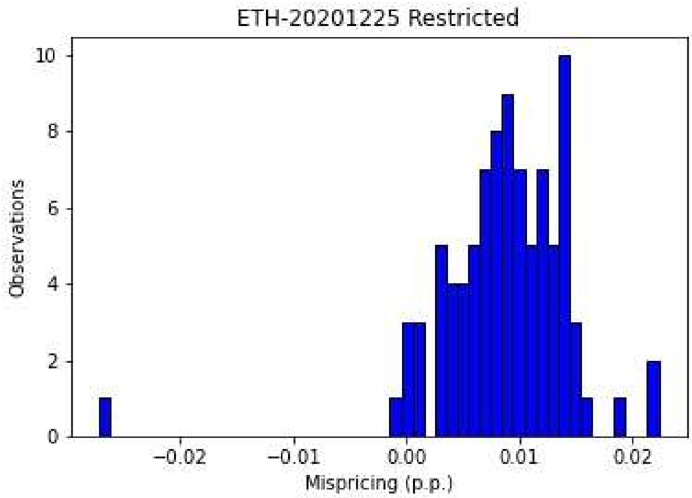


Figure 37 - Distribution of Mispricing regarding the period represented by ETH-20210326 (Restricted to Front Quarter Observations).

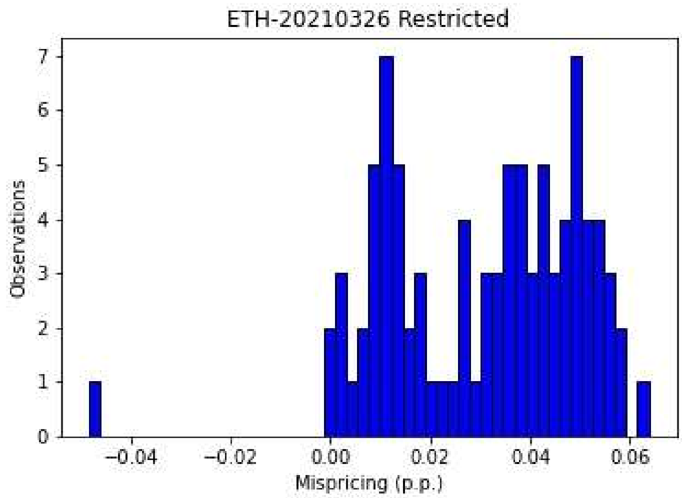


Figure 38 - Distribution of Mispricing regarding the period represented by ETH-20210625 (Restricted to Front Quarter Observations).

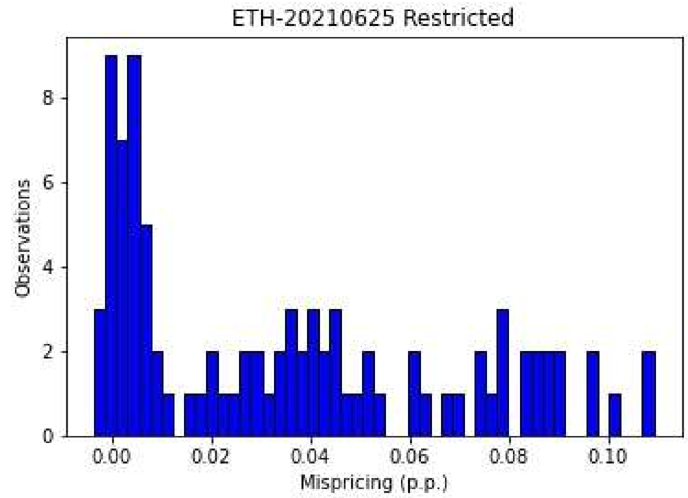


Figure 39 - Distribution of Mispricing regarding the period represented by ETH-20210924 (Restricted to Front Quarter Observations).

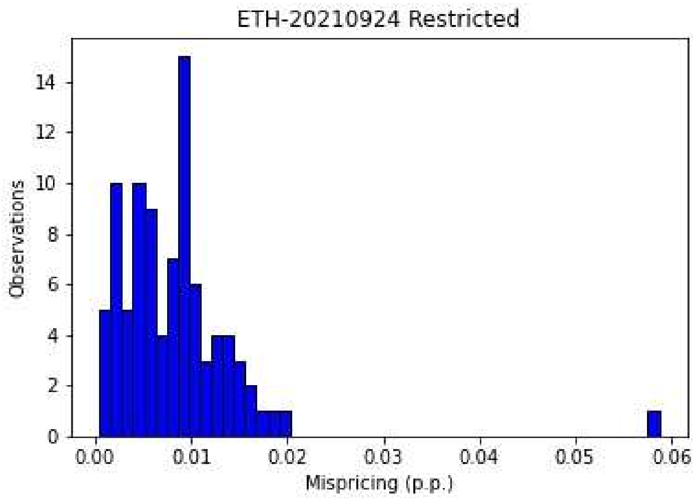


Figure 40 - Distribution of Mispricing regarding the period represented by ETH-20211231 (Restricted to Front Quarter Observations).

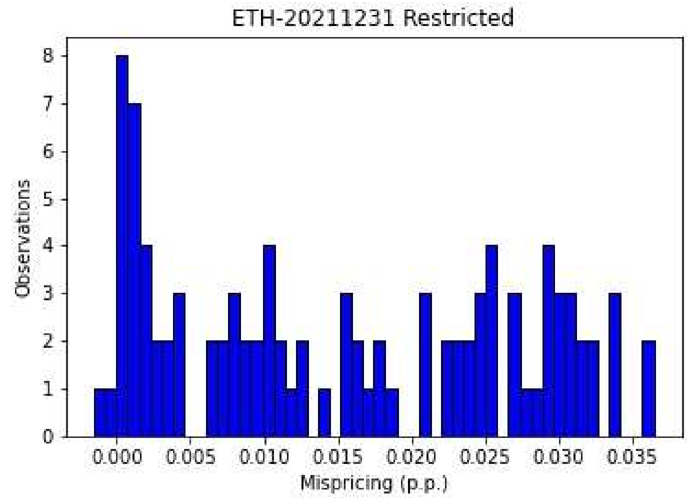


Figure 41 - Distribution of Mispricing regarding the period represented by ETH-20220325 (Restricted to Front Quarter Observations).

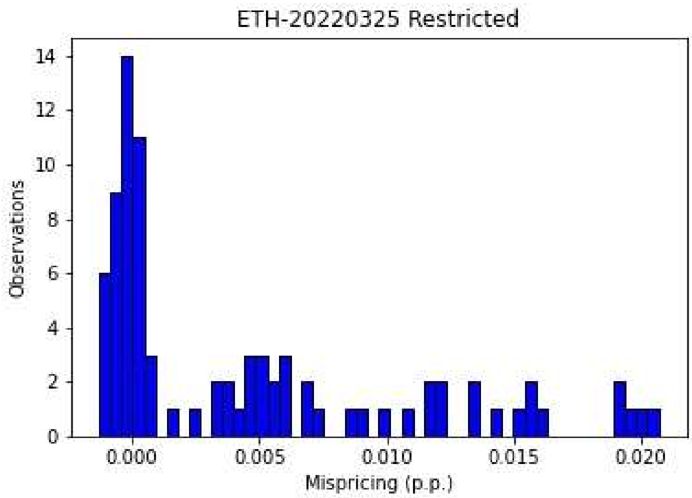


Figure 42 - Distribution of Mispricing regarding the period represented by ETH-20190927 (All Observations).

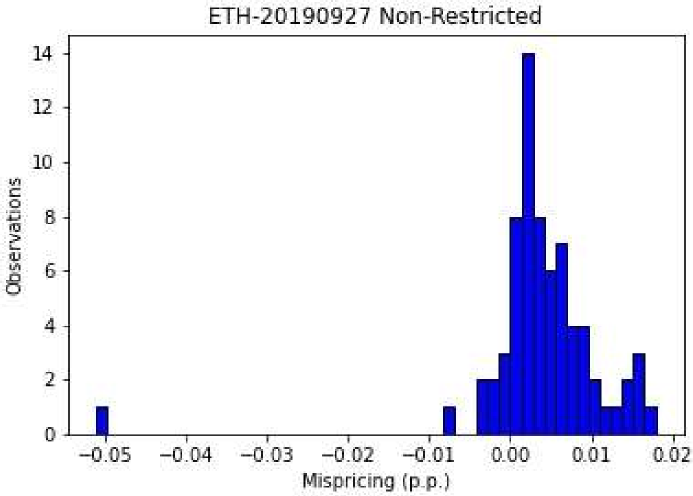


Figure 43 - Distribution of Mispricing regarding the period represented by ETH-20191227 (All Observations).

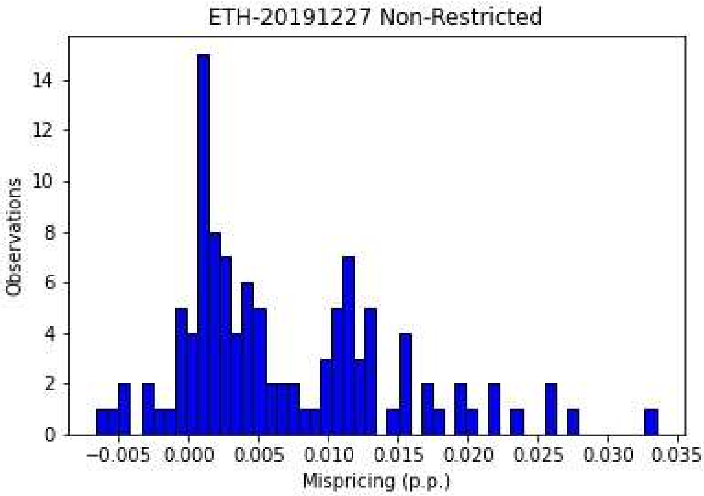


Figure 44 - Distribution of Mispricing regarding the period represented by ETH-20200327 (All Observations).

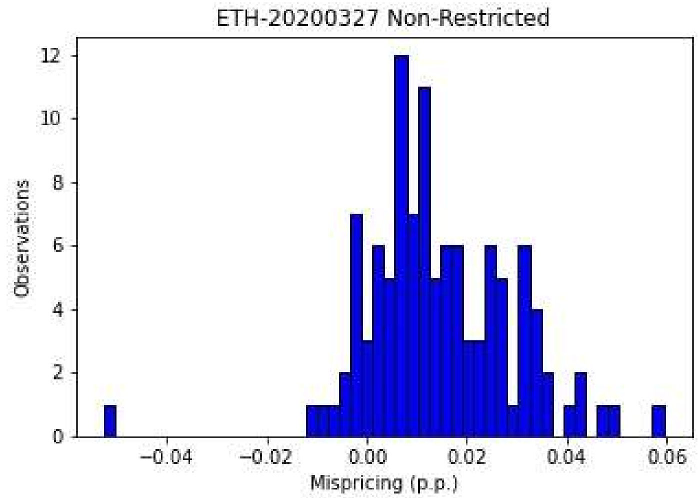


Figure 45 - Distribution of Mispricing regarding the period represented by ETH-20200626 (All Observations).

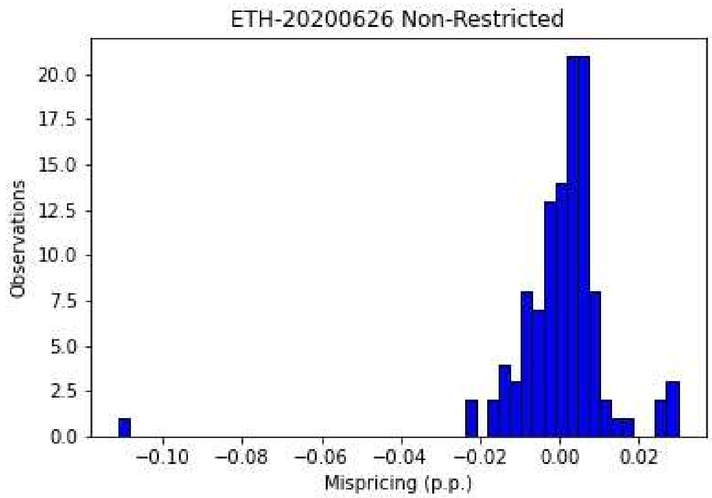


Figure 46 - Distribution of Mispricing regarding the period represented by ETH-20200925 (All Observations).

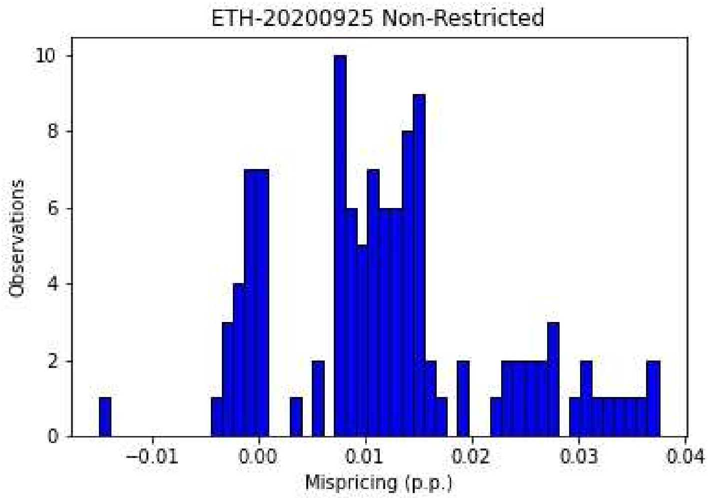


Figure 47 - Distribution of Mispricing regarding the period represented by ETH-20201225 (All Observations).

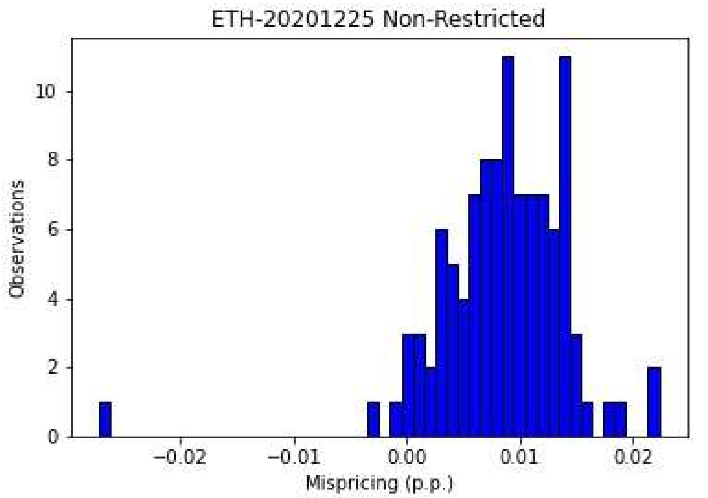


Figure 48 - Distribution of Mispricing regarding the period represented by ETH-20210326 (All Observations).

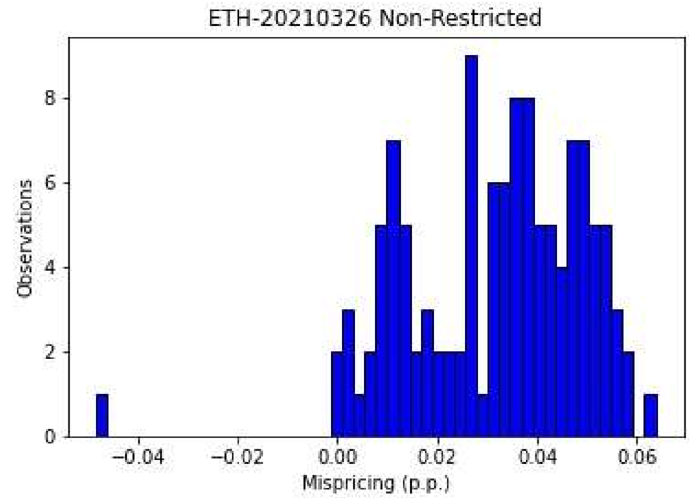


Figure 49 - Distribution of Mispricing regarding the period represented by ETH-20210625 (All Observations).

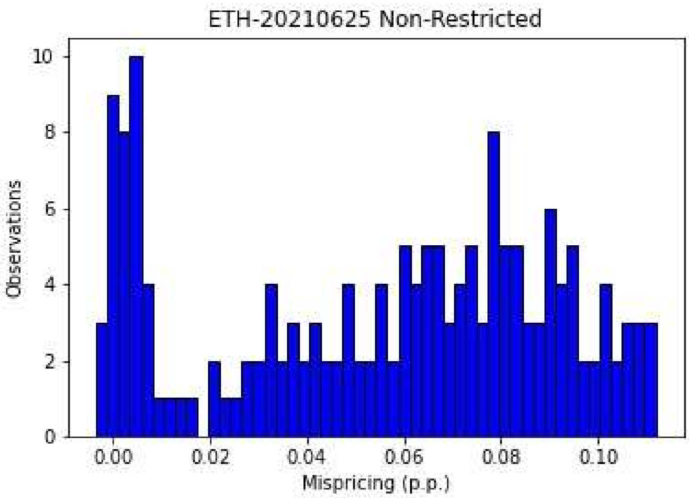


Figure 50 - Distribution of Mispricing regarding the period represented by ETH-20210924 (All Observations).

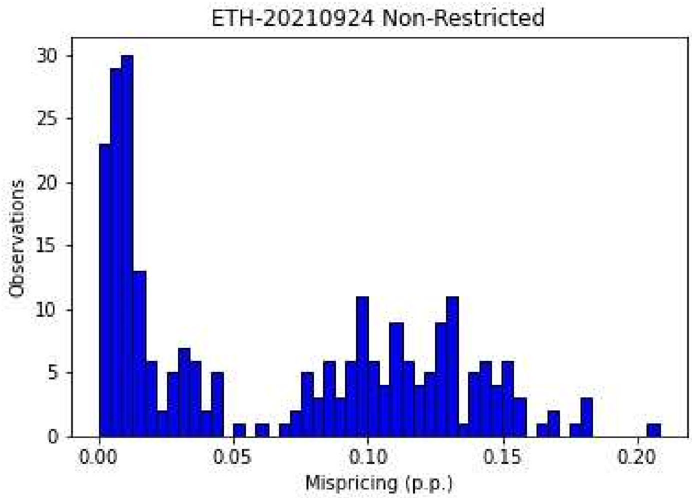


Figure 51 - Distribution of Mispricing regarding the period represented by ETH-20211231 (All Observations).

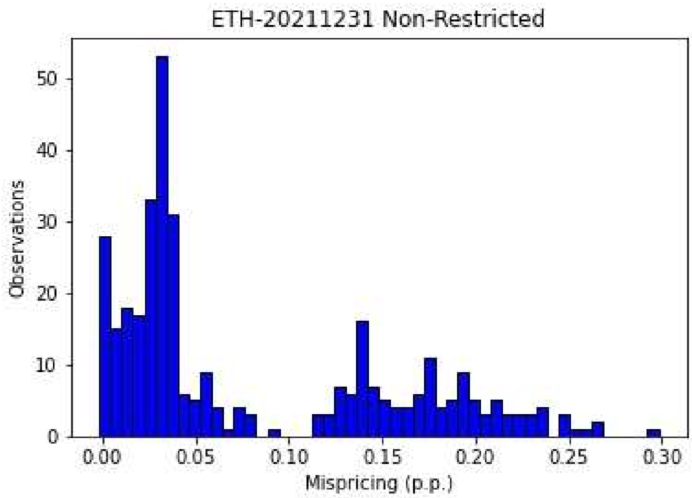


Figure 52 - Distribution of Mispricing regarding the period represented by ETH-20220325 (All Observations).