Predicting the volatility and liquidity of cryptocurrency futures contracts using its maturity

André Couto

Dissertation

Master in Modeling, Data Analysis and Decision Support Systems

Supervised by:

Carlos Francisco Alves, PhD

Pedro José Ramos Moreira de Campos, PhD

2022

Abstract

The main objective of this dissertation is to understand if the maturity of a Futures contract serves as a good predictor for the volatility and liquidity of the cryptocurrency Futures markets and their underlying Spot markets.

Firstly, a regression was performed to understand the relationship between the volatility and liquidity metrics and the days left to maturity of a contract. Following this, three distinct machine learning techniques - K-Nearest Neighbours, Support Vector Regression and Gradient Boosting Machine - were used to build prediction models for both the volatility and liquidity metrics.

The results show evidence that the volatility of Bitcoin and Ethereum's markets increases as the days to maturity approach zero. In addition, there is strong evidence showing that the liquidity of all cryptocurrencies, except Bitcoin, tends to increase as the days to maturity approach zero. However, neither the prediction models for the volatility nor the liquidity metrics of the cryptocurrency Futures and their underlying Spot markets could produce good predictions using machine learning.

Keywords: Cryptocurrencies; Volatility; Liquidity; Maturity effect; Machine Learning; OLS regression; K-Nearest Neighbours; Support Vector Regression; Gradient Boosting Machine

Resumo

O principal objetivo desta dissertação é estudar se o período de tempo até à liquidação de um contrato de futuros é um bom preditor da volatilidade e liquidez dos mercados de futuros de criptomoedas e dos respetivos mercados à vista.

Primeiro, foi elaborada uma regressão para compreender a relação entre a volatilidade e as métricas de liquidez e os dias restantes até à liquidação de um contrato. Em seguida, três técnicas distintas de aprendizagem automática - *K-Nearest Neighbours*, *Support Vector Regression* e *Gradient Boosting Machine* - foram utilizadas para construir modelos de previsão tanto para as métricas de volatilidade como para as métricas de liquidez.

Os resultados evidenciam que a volatilidade dos mercados de Bitcoin e Ethereum aumenta à medida que os dias até à maturidade se aproximam de zero. Além disso, há fortes evidências que mostram que a liquidez de todas as criptomoedas, exceto a Bitcoin, tende a aumentar à medida que os dias até à maturidade se aproximam de zero. No entanto, nem os modelos de previsão da volatilidade, nem das métricas de liquidez dos futuros de criptomoedas e dos respetivos mercados à vista foram capazes de produzir boas previsões utilizando as técnicas de aprendizagem automática.

Keywords: Criptomoedas; Volatilidade; Liquidez; Efeito da Maturidade; Aprendizagem Automática; regressão OLS ; K-Nearest Neighbours; Support Vector Regression; Gradient Boosting Machine

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Cryptocurrencies and their Spot markets have exploded in popularity since the introduction of Bitcoin by Nakamoto (2008). Moreover, their development led to financial derivatives, namely, cryptocurrency Futures contracts.

These cryptocurrency Futures markets have some differences when compared with traditional Futures markets. For example, the lack of a centralised clearing-house, the prominence of unregulated exchanges, the usage of inversely structured Futures, and the usage of perpetual swap Futures, a type of contract currently not used in any other markets (BitMEX, 2016; Wu, 2021).

Existing literature on traditional Futures markets indicates an increase in volatility as the Futures settlement date approaches, an effect known as the maturity effect (Anderson, 1985; Fama & French, 1988; Milonas, 1986; Samuelson, 1965). Evidence of increased trading volume in the Futures and its underlying markets can also be found in some markets (Serletis, 1992; Stoll & Whaley, 1987). Although cryptocurrency Futures have grown to have, on average, over USD$150 billion of daily volume (*Coingecko.com*, n.d.), such effects have not been studied yet.

This dissertation aims to study and understand the volatility and liquidity of these markets by taking into account the maturity of the nearest Futures contracts and understanding if these effects also hold for these markets and to what extent. Furthermore, this can then be used to understand if the maturity of cryptocurrency Futures contracts can predict the volatility and liquidity of their underlying Spot and Futures markets.

Two approaches will be taken. First, a regression will be estimated to understand if the maturity effect is present, similarly to earlier studies applied to traditional Futures (Galloway & Kolb, 1996; Milonas, 1986), and if a similar effect is also relevant when it comes to liquidity. Secondly, machine learning techniques will be used to build a prediction model that considers the days left to maturity of a future contract for liquidity and volatility.

The first machine learning technique used is K-Nearest Neighbours, which, although considered relatively simple, has been used with some success for financial data prediction (Alkhatib, Najadat, Hmeidi, & Shatnawi, 2013; Lora, Santos, Expósito, Ramos, & Santos, 2007; Tang, Pan, & Yao, 2018). The second machine learning technique used is Support Vector Regression, a version of the popular Support Vector Machine adapted for regression problems and has also been used extensively for the prediction of financial data (Cao & Tay, 2001; Ince & Trafalis, 2006; Kim, 2003; Okasha, 2014; Tay & Cao, 2001). The last technique used is gradient boosting machine, a machine learning algorithm that combines various weak learners to build a prediction (Bentéjac, Csörgő, & Martínez-Muñoz, 2021) which, like both other machine learning techniques, has also been used with some success to forecast financial data (Derbentsev, Matviychuk, Datsenko, Bezkorovainyi, & Azaryan, 2020; Sun, Liu, & Sima, 2020).

The main conclusions from this dissertation are that there is some evidence proving that the maturity effect is present in Bitcoin and Ethereum's markets and that there is strong evidence showing that the liquidity increases as the settlement date approximates for all cryptocurrencies - except Bitcoin which seems to remain constant. However, none of the machine learning techniques were able to produce adequate predictions for either the volatility or the liquidity of cryptocurrency Futures and their underlying Spot markets.

This dissertation is composed of six chapters: (1) *Introduction* establishes the relevance of the topic, problem and objectives; (2) *Literature review* shows previous relevant publications, as well as gives some context about the problem; (3) *Methodology* explains the predictive models that were used; (4) *Descriptive analysis* provides some insights into the data and its variables; (5) *Results* presents the outcomes from the regression and predictive models; (6) *Conclusion* includes a brief discussion about the findings, indicating some of the limitations and topics for future work.

Chapter 2

Literature review

2.1 Derivatives Markets

Derivatives are financial instruments that transfer risks from one party to another. It is named derivative as its value derives from an underlying asset, right or interest, such as a bond, a commodity, equity, or even a basket of assets (Heckinger, Mengle, Steigerwald, Ruffini, & Wells, 2013). Moreover, these underlying assets, rights or interests can also be the spread of different products or metrics, such as the CBOE Volatility Index, which measures the stock market's expectation of volatility based on S&P 500 stock market index Options (Whaley, 2009).

According to Chui (2012) the primary function of derivatives is to allow users to meet the demand for cost-effective protection against the risk associated with price movements of the underlying. Thus, derivative transactions involve transferring these risks from entities less willing or able to manage them to those more willing or able to do so. As expected, these transactions are common among numerous entities, ranging from small non-financial entities to central banks, with generally the most prevalent derivative type being currency derivatives (Bartram, Brown, & Fehle, 2009).

There are two main types of markets for derivatives, organised exchanges and over-the-counter markets, also referred to as OTC markets (Hull, 2003).

In organised exchanges, the contracts are standardised with specific delivery and settlement terms. The trading and settlement are done on an open market, with all trades publicly reported. Additionally, as the trading is usually done in an organised exchange, each counterparty to a contract does not know the identity of the other counterparty. Therefore, it cannot evaluate the correspondent counterparty risk, that is, the risk of the counterparty defaulting before the contract's expiration. Thus, most exchanges rely on clearinghouses, which act as a middleman between counterparties and require both to maintain enough collateral or margin to have their positions open, ensuring that both are well-capitalised (Lynch, 2011).

On the other hand, over-the-counter markets do not involve an exchange. Instead, the contracts are privately negotiated between the two parties, with all terms such as delivery quality, quantity, location and date being decided by both parties. Therefore, these contracts are, by definition, not standardised. Furthermore, as these contracts are privately traded, the trading and settlement are not reported publicly, making the details of these types of trading difficult to know, resulting in worse price transparency than organised exchanges. However, the flexibility of OTC markets is ideal for sophisticated market participants, such as banks and hedge funds, to trade contracts with no high order flow or special requirements. Consequently, these OTC markets also have the role of incubators for new financial products (Chui, 2012).

Margin

The margin is the collateral an investor needs to deposit and have available in their respective brokers or exchanges to cover any credit risk. For example, the investor can pose a credit risk for the broker or exchange if they borrow cash to buy a financial instrument, borrow a financial instrument to sell them short or enter into a derivative contract.

If the equity of the margin account of the investor drops below the maintenance margin level, the minimum equity the broker or exchange requires for the investor's positions to remain open, a margin call is issued, which requires the investor to either reduce their credit risk or deposit additional margin (FINRA, n.d.).

Limit Order Book

In financial markets, derivatives included, a limit order is a type of order to buy or sell a specific amount of contracts at a set price. For example, a limit buy order, also known as a bid, for ten at $2 indicates that a buyer is looking to buy ten contracts at the price of $2 per contract. On the other hand, a limit sell order, also known as an ask, for ten at the price of $2 indicates that a seller is looking to sell ten contracts at the price of $2 per contract (Tsantekidis et al., 2017).

The limit order book, also known as just an order book, can be considered an aggregation of all limit orders available at a certain time. The order book consists of two sides, the bid side containing all buy orders for each price, and the asks side, containing all sell orders for each price.

Bid-Ask Spread

Bid-ask spread is the difference between the best buy and sell orders available in the limit order book at a certain point in time, representing the execution cost of a roundtrip trade, the consecutive buy and sell of an asset for the same amount (Su & Tokmakcioglu, 2021).

As Stange and Kaserer (2009) noted, from a financial risk perspective, liquidity can be considered the ease of trade in a given market, partly represented by the size of the bid-ask spread. Similarly, Amihud, Mendelson, and Lauterbach (1997) also considers this cost of immediate execution, the bid-ask spread, as a natural measure of illiquidity. Further studies, such as the one done by Fleming (2001), also show that the bid-ask spread can serve as a reliable proxy for more complex liquidity measures for the U.S. Treasury markets.

2.1.1 Futures

A Futures contract is an agreement between two parties to buy or sell an asset at a specific time in the future for a certain price (Hull, 2003). These contracts are generally traded in organised exchanges with specified standardised features, such as quantity, location and quality. The underlying assets range from commodities, such as sugar and cotton, to financial assets, such as currencies and bonds.

Settlement

CME (n.d.) defines settlement as fulfilling the legal delivery obligations associated with the original contract. The date at which the contract settles is generally referred to as the settlement date. The settlement can be done in two ways, depending on the type of contract.

One is a physical delivery, where the amount specified of the underlying asset in the contract is delivered at a previously agreed location. This is a common type of settlement with commodities. However, very few of the Futures contracts lead to the delivery of the underlying asset as most are closed out early (Hull, 2003).

The second type is cash or financial settlement. The contracts are settled in cash on the underlying reference rate because it might be inconvenient or impossible to deliver the underlying asset. For example, this is the case of a Futures contract on the S&P 500, where delivering the underlying asset would require delivering a portfolio of 500 stocks. Therefore, when a contract is settled in cash, outstanding contracts, sometimes referred to as open interest, are declared closed and the price is set equal to the reference rate, representing the Spot price of the underlying asset.

2.1.2 The effects of Futures settlement

Maturity effect

Existing literature indicates that Futures price volatility may increase as the settlement date approaches, also referred to as the maturity effect. Samuelson (1965) provides strong empirical evidence that Futures contracts close to maturity show greater volatility than Futures contracts away from maturity. Other authors, such as Fama and French (1988) and Anderson (1985), followed in testing this maturity effect with similar results. One of these authors, Milonas (1986), attributes this behaviour to the fact that, since the price of a Futures contract at maturity must be virtually equal to the underlying Spot price, nearer contracts will tend to respond strongly to new information for the price of the future to converge to the underlying Spot price. Furthermore, Serletis (1992) and Herbert (1995) also find evidence that trading volume increases as Futures contracts approach maturity for multiple energy Futures markets.

However, although this effect has been widely studied, empirical evidence is still mixed for a majority of Futures contracts, with studies such as those done by Galloway and Kolb (1996) finding that the maturity effect plays a significant role in the volatility of Futures prices for commodities that experience seasonal demand or supply, such as agricultural and energy commodities, but not for other Futures, such as financial commodities and precious metals. Similar findings were also observed by Moosa and Bollen (2001) when studying the maturity effect on S&P 500 Futures contracts and by Daal, Farhat, and Wei (2006) when studying 6805 individual contracts separately across 61 different commodities.

Furthermore, Miller (1972) finds an inverse relationship between the volatility and time to maturity in the live beef Futures, the opposite of the maturity effect.

Expiration day effect

The settlement does not only affect Futures contracts. It can also affect the underlying Spot markets. For example, the increased trading volume associated with recurring special events, such as expiration days, is known as the expiration day effect in the literature. This is especially the case with stock index Futures contracts due to the arbitrageurs. Doing arbitrage between index Futures and the underlying cash index and the cash settlement feature of index Futures contracts, they need to reduce and close their positions in the stock market. This frequently explains the increased volume, especially during the last hour of trading on days on which the index Futures expire (Stoll & Whaley, 1987).

2.2 Cryptocurrencies

Cryptocurrency can be understood as a token, intended to be used as a medium of exchange, issued via a system that uses an often collectively-maintained digital ledger along with cryptography to some degree to replace the trust in institutions (Pernice & Scott, 2021).

The first cryptocurrency was eCash, a centralised system launched in the late 1990s and owned by a single institution. However, it already had some cryptographic features which would later be used by e-Gold until its liquidation in 2008 (Chuen, Guo, & Wang, 2017). Building on top of these features, Nakamoto (2008) introduced Bitcoin, the first digital currency that allowed to exchange value digitally without any third-party oversight by using a distributed ledger and cryptography technology. Although this cryptocurrency remains the most widely used and valuable, thousands more exist, named Altcoins, a combination of alternative and cryptocurrency (Chuen et al., 2017).

2.2.1 Cryptocurrency Markets

Market Structure

Initially, only cryptocurrencies Spot markets were available through centralised crypto exchanges. Later, derivatives appeared, namely Futures and subsequently Options and some decentralised exchanges. As of October 30, 2021, the 24-hour volume of derivatives exchanges is approximately US$170 billion, while Spot exchanges are US$151.0 billion (*Coingecko.com*, n.d.).

Price Mechanics

Ciaian and Rajcaniova (2018) suggest that the prices of Bitcoin and some altcoins are interdependent. Though more in the short-term than in the long-term, as well as that the price of both is impacted by global macroeconomics and financial developments – particularly, the price of gold, the exchange rate of USD/EUR and CNY/USD and the 10-Year Treasury Constant Maturity Rate. However, Jakub (2015) finds that global macroeconomics and financial developments do not drive the price of Bitcoin. Instead, the demand-side factors seem to have more importance.

Furthermore, Jakub (2015) also analysed the effect of public announcements on the price of Bitcoin and concluded that it reacts to publicly announced information.

2.3 Cryptocurrency Futures Markets

2.3.1 Cryptocurrency Futures exchanges

Similarly to traditional markets, organised exchanges and over-the-counter markets exist for cryptocurrencies. However, most cryptocurrency Futures volume is predominantly from unregulated organised exchanges. That is, exchanges that are not regulated or supervised by any financial regulatory body, such as the FCA, the SEC or the CFTC.

The difference relatively to traditional exchanges

As with traditional exchanges, these unregulated exchanges have standardised contracts and follow mark-to-market accounting practices. This is a practice where positions and collateral are measured using their market value, providing a more realistic appraisal of value using current market conditions and publicly reported trades.

However, unlike traditional exchanges, these unregulated exchanges do not use clearing-houses to minimise counterparty risk. Instead, they rely on liquidating the client's positions as soon as the unrealised losses are close to the value of their collateral and the remaining collateral value is below the maintenance margin, which is the minimum equity a client must hold to maintain a position open. In the case where the losses incurred are superior to the collateral value, generally, instead of letting the client's balance fall below zero and asking them to deposit additional funds, three approaches are used, occasionally combined:

1. establishing a fund, customarily called an 'insurance fund', which is mainly funded by part of the fees gathered by the platform and is used to cover any negative balances users might have (Binance, 2019; BitMEX, 2020)

2. having a liquidity provider program, where market makers or those willing to provide backstop liquidity agree to get a client's position up to a specific size in exchange for keeping their remaining balance (FTX, 2022)

3. forcibly closing profitable user's positions against the negative positions, generally referred to as 'auto-deleveraging', 'breaking open interest' or 'unwinding' (Binance, 2020; Bybit, 2021)

The importance of unregulated exchanges

Alexander, Choi, Park, and Sohn (2020) identify BitMEX's perpetual swaps, an unregulated product, as the price discovery leader, meaning that new information tends to be reflected faster on this product. Although different methodologies exist to establish this, Alexander et al. (2020) used a modified information share (MIS) (Lien & Shrestha, 2009), a modification of the more popular information share (IS) proposed by Hasbrouck (1995), as well as component shares (CS) proposed by Gonzalo and Granger (1995), both these approaches are derived from the cointegration relationship between markets.

Similarly, based on minute-by-minute transaction data, Alexander and Heck (2020) found a robust and consistent dominance of unregulated derivatives products over regulated Futures, which they attributed to three features: (1) the trading volume on the individual unregulated derivatives is much larger than on CME Futures, (2) unlike in more traditional asset classes, in crypto-asset markets, smaller players – such as miners or crypto-specialised hedge funds which have no access to large regulated exchanges, like CME, and instead trade on unregulated crypto-exchanges – can be considered informed traders, (3) the often-suspected manipulation in bitcoin markets could be another explanatory factor.

2.3.2 Inverse Futures

Unlike traditional markets, inverse Futures have a significant market share over vanilla Futures. In these contracts, the margining and the cash settlement are done in the base currency as opposed to the quote currency, meaning that with an inverse structure, a Bitcoin/USD contract would be margined and settled in Bitcoin as opposed to USD. Bragin (2015) noted that this came from the need to settle Futures contracts without having to clear any fiat currency. However, this type of contract also provides a non-linear payoff since the margining currency's price, which serves as collateral, fluctuates with the contract price, unlike in vanilla contracts, where it remains unchanged.

Thus, in a traditional vanilla linear contract, the payoff for a long position would be calculated as $(P_{close} - P_{open}) * Qty$, where P_{close} is the closing price of the position and P_{open} is the opening price of the position,

In contrast, in an inverse non-linear contract, it would be $(1/P_{close} - 1/P_{open}) * Qty$.

An example of these Futures contracts is the popular BitMEX perpetual inverse Futures XBTUSD contract, which has Bitcoin as its base currency and is quoted in USD while fully margined in Bitcoin (BitMEX, 2016; Wu, 2021).

Chapter 3

Methodology and Data

This chapter presents a brief introduction to the metrics, techniques and data that will be used in this dissertation.

Three distinct metrics will be used, two measuring liquidity, the bid-ask spread and the liquidity at 1% depth, and one measuring volatility, the normalised true range.

As shown in Table 3.1, some related studies already exist in which supervised Machine Learning techniques are applied to cryptocurrency financial time series. However, none take into account the maturity of the cryptocurrency Futures contracts.

Authors	Year	Market	Goal/Question	Methodology
Anderson	1985	Commodities futures	Study the effect of maturity and seasonality on the volatility of futures prices	• Nonparametric and parametric statistical tests
Milonas	1986	Commodities and Financials futures	Study the volatility as time to maturity nears	• OLS regression with nonstationarity from monthly variance removed
Galloway and Kolb	1996	Commodities futures	Study the maturity effect for 4111 different futures contracts	• OLS regression with t-statistics test
Peng et al.	2018	Cyptocurrencies and foreign currencies spot markets	Predict cryptocurrencies and foreign currencies' daily and hourly volatility	• GARCH • Support Vector Regression
Yao et al.	2018	Cryptocurrencies spot markets	Forecast short-term and long-term prices using deep-learning techniques	• Multilayer perceptron (MLP) • Long Short-Term Memory (LSTM)
Khaldi et al.	2019	Bitcoin spot markets	Predict Bitcoin's short-term and long-term volatility	• GARCH • Multilayer perceptron (MLP) • Recurrent neural network (RNN)
Antulov-Fantulin et al.	2020	Bitcoin spot markets	Predict Bitcoin's intraday volume	• ARMA and GARCH • Gradient Boosting • Temporal Mixture Ensemble
Derbentsev et al.	2020	Stock indexes, foreign currencies and cryptocurrencies	Forecast short-term financial time-series using supervised machien learning techniques	• Support Vector Regression • Artificial Neural Networks • Gradient Boosting • Random Forest
Kumar and Rath	2020	Cryptocurrencies spot markets	Forecast longer-term financial time-series using deep-learning techniques	• Multilayer perceptron (MLP) • Long Short-Term Memory (LSTM)
Cortez et al.	2021	Cyptocurrencies and foreign currencies spot markets	Predict short-term market liquidity	• ARMA and GARCH • K-nearest neighbor

Table 3.1: Relevant methodology of related studies

For the methodology, firstly, an ordinary least-squares (OLS) regression will be used to understand if the maturity effect is present and to understand if a similar effect is also relevant when it comes to liquidity. Following this, three different machine learning techniques will be used to build a prediction model that considers the days left to maturity of a future contract to predict liquidity and volatility.

The analysis and the predictions are all made using data composed of daily snapshots between the periods of 2020 and 2021 for both Futures and Spot markets.

3.1 Metrics used

3.1.1 Bid-Ask Spread

One of the liquidity metrics used is the average bid-ask spread, as noted by Su and Tokmakcioglu (2021). This is the difference between the best buy and sell orders available in the limit order book at a certain time, representing the execution cost of a roundtrip trade. That is, if the value of this metric for a particular observation is 0.5%, then the average difference between the highest bid price and the lowest ask price during that observation is 0.5%.

Authors such as Amihud et al. (1997) and Stange and Kaserer (2009) note that this cost of immediate execution, the bid-ask spread, can be considered a natural illiquidity measure. Furthermore, a study by Fleming (2001) showed that the bid-ask spread could serve as a reliable proxy for more complicated liquidity measures for the U.S. Treasury markets.

3.1.2 Liquidity at 1% depth

The second liquidity metric used is the average amount of liquidity available within 1% of the mid-price, the average price between the best bid and ask, in the order book during the observation.

This means that if the mid-price for an observation was US$ 1,000 and the liquidity at 1% depth was US$ 100,000 then there was US$ 100,000 worth of orders available between US$ 990 and US$ 1010.

3.1.3 Normalised True Range

The Normalised True Range is a measurement of volatility calculated using the True Range of an observation, which is the difference between the highest and lowest price. However, this is not normalised and thus cannot be compared across markets where the price might vary (Forman, 2006). The formula for the Normalised True Range used will be:

$$Var(p) = (H - L)/C \tag{3.1}$$

This is similar to the estimator used by Serletis (1992) and Herbert (1995), which normalised it by using the formula:

$$Var(p) = [ln(H) - ln(L)]^2/4ln_2 \tag{3.2}$$

where

Var(p) = estimate of the variance of daily price changes

H = highest trading price for the trading day

L = lowest trading price for the trading day

C = trading price at the close of the trading day

ln = natural logarithm

However, the normalisation used by Serletis (1992) and Herbert (1995) does not allow for the comparison of volatility between different Futures contracts as these would have different price ranges, and it only normalises the values to very small values.

3.2 Research question

The research question this dissertation aims to answer is "Can the maturity of cryptocurrency futures contracts predict the volatility and liquidity of their underlying spot and futures markets?" that is, this dissertation's goal is to understand if the maturity effect holds true for the cryptocurrency markets, if a similar effect is also relevant for the market liquidity and to use the maturity of the cryptocurrency Futures contracts to predict the volatility and liquidity of their underlying Spot and Futures markets.

3.3 Quantitative methodology

In this study, the principal methodology applied will be supervised Machine Learning techniques to financial time series. This will be applied to data containing daily observations.

3.3.1 OLS regression

The first approach will be using ordinary least-squares (OLS) regression to understand if the maturity effect is present. In other words, we want to verify if the volatility of Futures and Spot markets increases as settlement gets closer. This is in line with the methodology used in early studies such as Galloway and Kolb (1996), Milonas (1986), Serletis (1992) and Herbert (1995) and will serve as a baseline to understand if the days to maturity variable is relevant.

OLS regression is one of the most common techniques used in multivariate analysis. In this regression technique, the parameters estimated are the ones that yield the least sum of squared residuals, $\sum_{i=1}^{N} \epsilon_i^2$ with N being the number of observations in the sample and ϵ the error between the real and predicted value (Dismuke & Lindrooth, 2006).

For this study, the dependent variables will be considered as liquidity, represented by the bid-ask spread or by the liquidity at 1% depth, and the volatility, represented by the Normalised True Range. In contrast, the independent variable will be the days left to settlement. Thus the formula for this OLS regression model for volatility can be considered as:

$$Y = \beta X + c + e \tag{3.3}$$

where Y is the volatility indicator, β is the coefficient of variable X, the days left to settlement, c is the constant of the model, thus the value when X is 0, and e is the regression error with a mean of zero, constant variance and no autocorrelation. For the liquidity model, Y would be our liquidity indicator instead. However, everything else would remain unchanged.

Thus, if the maturity effect holds, we expect the coefficient β to be negative and statistically significant.

3.3.2 Supervised Machine Learning

Supervised Machine Learning techniques have been used to forecast financial time series for some time. As seen in Table 3.1, relevant studies use a range of techniques in this field, from Random Forests to deep-learning techniques such as long short-term memory (LSTM).

Studies such as the ones done by Derbentsev et al. (2020) and Sun et al. (2020) have been able to use traditional machine learning techniques such as gradient boosting machine to forecast financial time series with success. Thus, for this study, supervised machine learning techniques, such as K-Nearest Neighbours, Support Vector Regression, and Gradient Boosting Machine will be used.

All models will be trained to predict the last ten days of observations of a given Futures contract or Spot section in the test set.

Hyperparameter optimization

In order to optimise the hyperparameters of the Machine Learning algorithms, a grid search approach was taken. Grid search is one of the simplest approaches to optimising the hyperparameters of an algorithm. It performs an exhaustive search through a manually specified set of hyperparameters. (Feurer & Hutter, 2019; Hsu, Chang, Lin, et al., 2003)

K-Nearest Neighbours

The K-Nearest Neighbours algorithm (k-NN) is a technique used for both classification and regression which predicts a continuous or discrete label based on the labels of the K nearest observations. It was initially introduced Fix and Hodges (1951) and later expanded by Cover and Hart (1967) and Altman (1992).

This algorithm is a simple yet high accuracy algorithm that has proven effective in several cases. Although it is mainly used for classification problems, it can handle regression problems by computing the average of the values of its K nearest neighbours (Taunk, De, Verma, & Swetapadma, 2019). The idea of averaging the values of the K nearest neighbours is based on the assumption that observations in a near data space will have a similar continuous label (Kramer, 2013).

The steps to compute a simple k-NN regression are as follows:

1. Decide the number of nearest neighbours K to use
2. Compute the distance between the observation's features and all other observations
3. Average the values of the nearest K neighbours

Furthermore, studies such as those done by Alkhatib et al. (2013), which used this technique to predict stock prices for a sample of six major companies listed on the Jordanian stock exchange, Lora et al. (2007), which applied a weighted variation of this algorithm to the Spanish energy markets, Tang et al. (2018), which used k-NN together with PCA on the EUR/USD exchange rate and Chinese HS300 index, and others have obtained favourable results with using k-NN for financial data prediction.

Support Vector Regression

Support Vector Machine was developed by Cortes and Vapnik (1995) as a learning machine algorithm for two-group classification problems. This algorithm works by separating the observations into two spaces using a decision boundary, which can be linear or non-linear, commonly referred to as a hyperplane (Noble, 2006).

In its simplest form SVM finds the most efficient hyperplane by maximising the margin between the two classes, with the constraint that the classifier does not make any mistakes. This is commonly referred to as an SVM with a hard margin.

On the other hand, an SVM can also have a soft margin, in which case it still tries to maximise the margin between the two classes. However, it allows some training observations to be misclassified if their penalty is not too high (Pradhan, 2012; Suthaharan, 2016).

In both types of margin, a simple linear mathematical model such as $y = wx + b$ can be used to divide the data with a linear hyperplane. However, in cases where the data cannot be divided linearly, a non-linear kernel function has to be used (Suthaharan, 2016).

A version used for regression problems called Support Vector Regression was later introduced by Drucker, Burges, Kaufman, Smola, and Vapnik (1996). However, unlike with classification problems, the margin variable ϵ represents the tolerance level, that is, the objective of the Support Vector Regression is to find the hyperplane with the most observations within that certain tolerance margin. Furthermore, similarly to an SVM with a soft margin, a variable C is also used to represent the penalty of any deviation from that tolerance margin and a variable γ is also used to limit the influence that a single observation can have (Smola & Schölkopf, 2004).

Furthermore, due to this technique's popularity, it has been successfully applied to different financial time series prediction problems in the past (Cao & Tay, 2001; Ince & Trafalis, 2006; Kim, 2003; Okasha, 2014; Tay & Cao, 2001).

Gradient Boosting Machine

Gradient Boosting Machine was introduced by Friedman (2001) and is an algorithm that, like other boosting algorithms, combines multiple weak learners, such as decision trees, building a weighted sum of all functions to minimise a particular loss function.

Many variations exist, such as XGBoost (Chen & Guestrin, 2016), a decision tree ensemble based on gradient boosting designed to be highly scalable (Bentéjac et al., 2021).

This machine learning technique has also been used with some success to forecast financial time series (Derbentsev et al., 2020; Sun et al., 2020).

3.3.3 Validation

As the dataset is not a continuous time series but a time series where an integer is used to represent the time difference to a specific date, the expiry, traditional time series validation techniques need to be adapted since the same *DTM* will be present on all contracts.

For validation, 3-fold cross-validation was used with 20% of all contracts or Spot sections randomly picked to serve as out-of-sample test data and the remaining 80% used to train the model. After the model was trained using the train data, the same model was trained using the test data, excluding the last ten days of observations. That same model was then used to predict the last ten days of observations and compared against the actual values.

3.3.4 Performance Metrics

Performance metrics are vital for the evaluation of models. Therefore, the model will be evaluated using the Mean Absolute Error (MAE) and the Mean Absolute Percentage Error (MAPE).

The MAE is calculated by the sum of the prediction error, the difference between the actual output and the predicted output, then divided by the number of data points (Bickel & Doksum, 2015).

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - x_i| \tag{3.4}$$

where

y_i = predicted output

x_i = actual output

n = total number of data points

The MAPE is calculated by the prediction error relative to the actual output divided by the number of data points. Unlike MAE, it is not scale-dependent as it is a percentage, which allows for comparing accuracy across models and data sets. (De Myttenaere, Golden, Le Grand, & Rossi, 2016; Hyndman & Koehler, 2006)

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}|\frac{y_i - x_i}{x_i}| \tag{3.5}$$

where

y_i = predicted output

x_i = actual output

n = total number of data points

3.3.5 Software

The software that will be used to do the data pre-processing, building the models, performing the analysis, as well as visualisation will be Python. This programming language presents various advantages compared to other languages (Hilpisch, 2015; Srinath, 2017), such as:

- It is open-source, along with a majority of its libraries;
- It has a clear syntax, making its code highly readable while still being very close to the mathematics;
- It is highly extensible, with a vast array of powerful libraries available, ranging from libraries to apply deep learning models to libraries for game development;

For these and other reasons, Python is considered the fastest-growing programming language, especially among data scientists and analysts (Srinath, 2017).

3.4 Pre-processing

3.4.1 Spot market sections

Unlike Futures contracts, the Spot markets will always have the same name as they do not have a contract name. As the date of observation is also not provided to the models, the model has no way to identify from which point in time it is trying to predict as the *DTM* will repeat multiple times for the same Spot market.

To overcome this, a random letter was assigned to each section of Spot observations for the model to identify which point in time it is predicting as it will consider the Spot sections used to train with the same letter.

Thus, if it considered the Spot section from 01-01-2020 to 26-03-2020 to be the section A and it wanted to predict the last ten days of observations for this section, it would place more significant weight on the observations in the training set with the section A as it is closer in time.

3.4.2 Categorical variables

The data set presents categorical variables such as the contract, exchange and asset name. However, regressions and certain Machine Learning techniques, such as Gradient Boosting Machine, cannot easily cope with such variables.

One method to easily enable such techniques to use variables that would otherwise not be measured is to use dummy variables. For this data set, binary dummy variables were created for all categorical variables, with a value of 1 for observations where the categorical variable is present and 0 when it is not. This is commonly referred to as one-hot encoding (Suits, 1957).

Thus, an observation for Bitcoin, would have the *Bitcoin* variable as 1, and the *Ethereum*, *Solana* and *Dogecoin* variables as 0.

3.4.3 Missing Values

All the observations containing missing values were removed.

3.4.4 Data

The dataset comprises daily snapshots extracted from both Cryptowat.ch (https://cryptowat.ch) and CoinAPI (https://www.coinapi.io) for 2020 and 2021 for both Futures and Spot markets. Each observation is composed of the average bid-ask spread for that day and the average depth in USD within 1% of the mid-price, both liquidity metrics, and the normalised true range, a volatility metric and naturally, the days to the settlement of the nearest Futures contract.

The Futures markets used are from the exchanges FTX, BitMEX and Deribit and Spot markets used are from Bitfinex, Coinbase and FTX. The cryptocurrencies studied will be Bitcoin, Ethereum, Dogecoin and Solana.

There are 4050 observations for 47 different Futures contracts from three different unregulated exchanges, FTX, Bitmex and Deribit, and 5933 observations for the Spot markets of Bitcoin, Ethereum, Dogecoin and Solana from three different unregulated exchanges, Coinbase Pro, FTX and Bitfinex.

The first observation is from the 1st of January 2020, and the last is on the 30th of December 2021.

The Futures observations are composed of the following variables:

- Asset – the asset of the underlying Futures contract of the observation – e.g. BTC, ETH, etc.

- Exchange – the exchange on which the Futures contract of the observation is traded – e.g. bitmex, FTX, etc.
- Contract – the Futures contract of the observation – e.g. XBTH20, ETH-24SEP21, etc.
- Avg Bid-Ask Spread – the average bid-ask spread during the observation, the average difference between the best bid price and ask price, in percentage, one of the liquidity metrics – e.g. 0.05%, 0.1%, etc.
- Avg Liquidity at 1% depth – the average liquidity available at 1% of the mid-price, the average price between the best bid and ask, in USD during the observation, another liquidity metric – e.g. US$ 1 000 000, US$ 500 000, etc.
- DTM – the days to maturity, how many total days are left between the settlement date and the observation date – e.g. 91, 0, etc.
- NTR – Normalised True Range, the difference between the observation's highest traded price and the lowest traded price divided by the closing price, resulting in how much the price ranged during that observation. a liquidity metric – e.g. 3.04%, 112.84%, etc.

The Spot observations are composed of the following variables:

- Asset – the asset of the Spot market of the observation – e.g. BTC, ETH, etc.
- Exchange – the Spot exchange of the observation – e.g. bitfinex, FTX, etc.
- Avg Bid-Ask Spread – the average bid-ask spread during the observation, the average difference between the best bid price and ask price, in percentage, one of the liquidity metrics – e.g. 0.05%, 0.1%, etc.
- Avg Liquidity at 1% depth – the average liquidity available at 1% of the mid-price, the average price between the best bid and ask, in USD during the observation, another liquidity metric – e.g. US$ 1 000 000, US$ 500 000, etc.
- NTR – Normalised True Range, the difference between the observation's highest traded price and the lowest traded price divided by the closing price, resulting in how much the price ranged during that observation. a liquidity metric – e.g. 3.04%, 112.84%, etc.

The Spot markets were also split into different sections based on the days to the maturity of the closest Futures contract as explained on 3.4.1.

Chapter 4

Descriptive analysis

4.1 Description of data

Bitcoin

There are 21 different Bitcoin Futures contracts, seven from FTX, eight from Bitmex and six from Deribit. Although missing observations for FTX's contracts are relatively low, missing observations are significant for some of Bitmex's and Deribit's contracts, with Bitmex's XBTU21 contract containing over 51% of missing values.

Exchange	Contract	N	First Obs	Last Obs	Missing Obs
FTX	BTC-20200626	66	2020-04-21	2020-06-25	0
	BTC-20200925	91	2020-06-26	2020-09-24	0
	BTC-20201225	91	2020-09-25	2020-12-24	0
	BTC-20210326	91	2020-12-25	2021-03-25	0
	BTC-20210625	91	2021-03-26	2021-06-24	2
	BTC-20210924	91	2021-06-25	2021-09-23	2
	BTC-20211231	98	2021-09-24	2021-12-30	0
	SPOT	619	2020-04-21	2021-12-30	5
Bitmex	XBTH20	86	2020-01-01	2020-03-26	0
	XBTM20	91	2020-03-27	2020-06-25	0
	XBTU20	91	2020-06-26	2020-09-24	0
	XBTZ20	91	2020-09-25	2020-12-24	0
	XBTH21	91	2020-12-25	2021-03-25	5
	XBTM21	91	2021-03-26	2021-06-24	2
	XBTU21	91	2021-06-25	2021-09-23	47
	XBTZ21	98	2021-09-24	2021-12-30	7
Deribit	BTC-25SEP20	72	2020-07-15	2020-09-24	7
	BTC-25DEC20	91	2020-09-25	2020-12-24	0
	BTC-26MAR21	91	2020-12-25	2021-03-25	5
	BTC-25JUN21	91	2021-03-26	2021-06-24	8
	BTC-24SEP21	91	2021-06-25	2021-09-23	7
	BTC-31DEC21	98	2021-09-24	2021-12-30	0
Bitfinex	SPOT	730	2020-01-01	2021-12-30	2
Coinbase Pro	SPOT	730	2020-01-01	2021-12-30	12

Table 4.1: Bitcoin Observations

Ethereum

There are 13 different Ethereum Futures contracts, seven from FTX and six from Deribit. Missing observations for most contracts are reasonably low, with the highest being for Deribit's ETH-25SEP20 contract, with approximately 10% of missing values.

Exchange	Contract	N	First Obs	Last Obs	Missing Obs
FTX	ETH-20200626	66	2020-04-21	2020-06-25	0
	ETH-20200925	91	2020-06-26	2020-09-24	0
	ETH-20201225	91	2020-09-25	2020-12-24	0
	ETH-20210326	91	2020-12-25	2021-03-25	0
	ETH-20210625	91	2021-03-26	2021-06-24	2
	ETH-20210924	91	2021-06-25	2021-09-23	1
	ETH-20211231	98	2021-09-24	2021-12-30	0
	SPOT	619	2020-04-21	2021-12-30	5
Deribit	ETH-25SEP20	72	2020-07-15	2020-09-24	7
	ETH-25DEC20	91	2020-09-25	2020-12-24	0
	ETH-26MAR21	91	2020-12-25	2021-03-25	5
	ETH-25JUN21	91	2021-03-26	2021-06-24	8
	ETH-24SEP21	91	2021-06-25	2021-09-23	1
	ETH-31DEC21	98	2021-09-24	2021-12-30	0
Bitfinex	SPOT	730	2020-01-01	2021-12-30	1
Coinbase Pro	SPOT	730	2020-01-01	2021-12-30	12

Table 4.2: Ethereum Observations

Dogecoin

For Dogecoin, there are 7 Futures contracts, all on FTX, missing values are minimal.

Exchange	Contract	N	First Obs	Last Obs	Missing Obs
FTX	DOGE-20200626	65	2020-04-22	2020-06-25	0
	DOGE-20200925	91	2020-06-26	2020-09-24	0
	DOGE-20201225	91	2020-09-25	2020-12-24	0
	DOGE-20210326	91	2020-12-25	2021-03-25	0
	DOGE-20210625	91	2021-03-26	2021-06-24	2
	DOGE-20210924	91	2021-06-25	2021-09-23	2
	DOGE-20211231	98	2021-09-24	2021-12-30	1
	SPOT	352	2021-01-13	2021-12-30	4
Bitfinex	SPOT	254	2021-04-21	2021-12-30	5
Coinbase Pro	SPOT	199	2021-06-15	2021-12-30	10

Table 4.3: Dogecoin Observations

Solana

Similarly to Dogecoin, all of Solana's 6 Futures contracts are from FTX, and missing values are also minimal.

Exchange	Contract	N	First Obs	Last Obs	Missing Obs
FTX	SOL-20200925	60	2020-07-27	2020-09-24	0
	SOL-20201225	91	2020-09-25	2020-12-24	0
	SOL-20210326	91	2020-12-25	2021-03-25	0
	SOL-20210625	91	2021-03-26	2021-06-24	2
	SOL-20210924	91	2021-06-25	2021-09-23	2
	SOL-20211231	98	2021-09-24	2021-12-30	1
	SPOT	522	2020-07-27	2021-12-30	4
Bitfinex	SPOT	309	2021-02-25	2021-12-30	5
Coinbase Pro	SPOT	195	2021-06-19	2021-12-30	2

Table 4.4: Solana Observations

4.2 Liquidity metrics

Bitcoin

As seen in Tables 4.5 and 4.6, out of the three exchanges, the Bitcoin Futures of Bitmex seem to have both the lowest average bid-ask spread and the highest liquidity within a 1% depth.

However, Bitmex has the observation with the lowest liquidity at 1% depth, of just US$ 69 571, recorded on the XBTZ21 contract on 2021-12-27. Similarly, it has the second-highest max bid-ask spread of 0.391%, even though it has the best mean and minimum.

The average bid-ask spread of 0.391% was observed in the XBTH20 contract on 2020-03-13, the day with the second-highest volatility according to the normalised true range, which might account for this.

Spot markets seem to present better liquidity than the Futures markets.

	Bid-Ask Spread					
	FTX	*Bitmex*	*Deribit*	*Bitfinex (Spot)*	*Coinbase Pro(Spot)*	*FTX (Spot)*
N (observations)	615	669	507	728	718	614
Mean	0.032%	0.023%	0.046%	0.008%	0.005%	0.014%
Std.	0.020%	0.027%	0.037%	0.005%	0.009%	0.017%
25%	0.013%	0.006%	0.026%	0.001%	0.002%	0.003%
50%	0.028%	0.013%	0.042%	0.007%	0.003%	0.007%
75%	0.048%	0.032%	0.060%	0.010%	0.006%	0.020%
min	0.004%	0.002%	0.005%	0.001%	0.000%	0.002%
max	0.111%	0.391%	0.652%	0.051%	0.191%	0.107%

Table 4.5: Bitcoin Futures and Spot Bid-Ask Spread

	Liquidity at 1% depth					
	FTX	*Bitmex*	*Deribit*	*Bitfinex (Spot)*	*Coinbase Pro (Spot)*	*FTX (Spot)*
N (observations)	615	669	507	728	718	614
Mean	5.31	6.87	4.58	22.02	9.66	13.79
Std.	1.87	4.52	2.04	12.81	5.32	8.05
25%	3.91	4.85	2.90	10.61	5.41	9.43
50%	5.02	7.31	3.97	18.52	7.68	10.47
75%	6.37	8.80	6.20	32.90	13.85	12.91
min	2.13	0.07	0.68	1.32	0.65	1.95
max	12.46	19.30	9.83	60.80	28.05	43.58

Table 4.6: Bitcoin Futures and Spot Liquidity at 1% depth in Millions of USD

Ethereum

As seen in Tables 4.7 and 4.8 out of the two exchanges with Ethereum Futures, FTX seems to have both the lowest average bid-ask spread and the highest liquidity within a 1% depth.

Spot markets also seem to present better liquidity than the Futures markets.

	Bid-Ask Spread				
	FTX	*Deribit*	*Bitfinex (Spot)*	*Coinbase Pro (Spot)*	*FTX (Spot)*
N (observations)	616	513	617	607	614
Mean	0.051%	0.082%	0.018%	0.009%	0.028%
Std.	0.021%	0.090%	0.011%	0.006%	0.028%
25%	0.033%	0.056%	0.011%	0.004%	0.005%
50%	0.053%	0.074%	0.015%	0.008%	0.020%
75%	0.064%	0.094%	0.021%	0.011%	0.039%
min	0.008%	0.016%	0.005%	0.000%	0.002%
max	0.167%	1.981%	0.102%	0.045%	0.181%

Table 4.7: Ethereum Futures and Spot Bid-Ask Spread

	Liquidity at 1% depth				
	FTX	*Deribit*	*Bitfinex (Spot)*	*Coinbase Pro (Spot)*	*FTX (Spot)*
N (observations)	616	513	617	607	614
Mean	3.64	2.28	14.95	6.92	4.88
Std.	1.69	1.00	9.10	5.42	5.09
25%	2.39	1.58	6.30	1.88	2.33
50%	3.04	2.03	15.67	5.73	2.63
75%	4.72	2.79	22.41	11.30	4.82
min	1.01	0.42	2.31	1.04	0.88
max	8.35	5.54	42.59	22.29	24.47

Table 4.8: Ethereum Futures and Spot Liquidity at 1% depth in Millions of USD

Dogecoin

As seen in Tables 4.9 and 4.10 there is only a single exchange with Dogecoin Futures. It seems to present both a higher average bid-ask spread and lower liquidity available at 1% depth when compared with the Spot markets, both metrics indicate that it tends to be less liquid than its Spot markets.

	Bid-Ask Spread			
	FTX	*Bitfinex (Spot)*	*Coinbase Pro (Spot)*	*FTX (Spot)*
N (observations)	613	249	189	348
Mean	0.359%	0.112%	0.051%	0.153%
Std.	0.197%	0.065%	0.010%	0.260%
25%	0.234%	0.059%	0.044%	0.028%
50%	0.335%	0.102%	0.048%	0.059%
75%	0.418%	0.149%	0.058%	0.164%
min	0.092%	0.032%	0.033%	0.011%
max	2.748%	0.532%	0.082%	2.107%

Table 4.9: Dogecoin Futures and Spot Bid-Ask Spread

	Liquidity at 1% depth			
	FTX	*Bitfinex (Spot)*	*Coinbase Pro (Spot)*	*FTX (Spot)*
N (observations)	613	249	189	348
Mean	0.21	2.00	1.83	0.817
Std.	0.26	0.92	0.45	0.54
25%	0.02	1.43	1.54	0.16
50%	0.10	1.84	1.82	0.95
75%	0.33	2.53	2.05	1.17
min	0.00	0.05	0.58	0.00
max	1.94	4.72	3.03	2.00

Table 4.10: Dogecoin Futures and Spot Liquidity at 1% depth in Millions of USD

Solana

As seen in Tables 4.11 and 4.12, similarly to Dogecoin, there is only a single exchange with Solana Futures and, similarly to Dogecoin, most Spot markets for Solana seem to have on average better liquidity than the Futures markets.

	Bid-Ask Spread			
	FTX	*Bitfinex (Spot)*	*Coinbase Pro (Spot)*	*FTX (Spot)*
N (observations)	517	304	193	518
Mean	0.355%	0.135%	0.038%	0.411%
Std.	0.252%	0.221%	0.029%	0.828%
25%	0.147%	0.056%	0.017%	0.042%
50%	0.285%	0.091	0.024	0.125%
75%	0.528%	0.118%	0.052%	0.508%
min	0.065%	0.019%	0.001%	0.003%
max	1.199%	1.809%	0.167%	5.258%

Table 4.11: Solana Futures and Spot Bid-Ask Spread

	Liquidity at 1% depth			
	FTX	*Bitfinex (Spot)*	*Coinbase Pro (Spot)*	*FTX (Spot)*
N (observations)	517	304	193	518
Mean	0.93	2.26	2.72	0.57
Std.	1.22	1.74	1.47	0.52
25%	0.01	0.50	1.07	0.03
50%	0.15	1.84	3.20	0.48
75%	1.89	3.49	3.82	1.07
min	0.00	0.03	0.27	0.00
max	4.31	7.29	5.18	1.88

Table 4.12: Solana Futures and Spot Liquidity at 1% depth in Millions of USD

4.3 Volatility metrics

Bitcoin

On Table 4.13, all Futures exchanges have a similar average normalised true range. However, its maximum values range from 27.73% to 95.33%, which can be due to several reasons, namely, the fact that these observations are not all for the same periods, as FTX's contracts had its first observation on 2020-04-21, Bitmex on 2020-01-01 and Deribit on 2020-07-15, and the fact that different markets will also have different levels of liquidity and market behaviour which can lead to short periods of very high volatility.

	Normalised True Range					
	FTX	*Bitmex*	*Deribit*	*Bitfinex (Spot)*	*Coinbase Pro(Spot)*	*FTX (Spot)*
N	615	669	507	728	718	614
Mean	6.25%	6.42%	6.15%	5.87%	5.96%	5.91%
Std.	4.08%	5.73%	3.61%	4.60%	4.67%	3.93%
25%	3.67%	3.53%	3.80%	3.25%	3.36%	3.38%
50%	5.39%	5.27%	5.46%	4.84%	4.92%	4.99%
75%	7.63%	7.54%	7.57%	6.93%	7.05%	7.13%
min	0.92%	0.63%	0.99%	0.87%	0.95%	0.92%
max	41.45%	95.33%	27.73%	69.65%	68.47%	39.40%

Table 4.13: Bitcoin Futures Normalised True Range

Ethereum

Table 4.14 shows that both Futures exchanges also have a similar average normalised true range. However, its max values are 31.73% and 67.52%, which, similarly to the Bitcoin Futures, can be due to several reasons, namely, the fact that these observations are not all for the same periods, as FTX's contracts had their first observation on 2020-04-21 and Deribit on 2020-07-15, and the fact that different markets will also have different levels of liquidity and market behaviour which can lead to short periods of very high volatility.

	Normalised True Range				
	FTX	*Deribit*	*Bitfinex (Spot)*	*Coinbase Pro (Spot)*	*FTX (Spot)*
N	616	513	617	607	614
Mean	8.12%	7.99%	7.57%	7.62%	7.78%
Std.	5.33%	4.37%	4.98%	5.15%	5.31%
25%	4.79%	5.10%	4.47%	4.49%	4.57%
50%	6.98%	7.12%	6.44%	6.36%	6.54%
75%	9.81%	9.69%	9.29%	9.27%	9.47%
min	1.36%	1.49%	1.04%	1.58%	1.33%
max	67.52%	31.73%	57.72%	63.91%	66.04%

Table 4.14: Ethereum Futures Normalised True Range

Dogecoin

The average and max normalised true range observed on Dogecoin Futures (Table 4.15) is higher than those observed for Bitcoin and Ethereum. The maximum normalised true range observed is higher than Solana.

	Normalised True Range			
	FTX	*Bitfinex (Spot)*	*Coinbase Pro (Spot)*	*FTX (Spot)*
N	613	249	189	348
Mean	10.62%	11.46%	8.99%	13.76%
Std.	12.19%	9.76%	6.07%	14.29%
25%	3.96%	5.74%	5.27%	5.86%
50%	6.63%	8.19%	7.21%	8.97%
75%	12.04%	13.55%	10.55%	15.87%
min	0.95%	2.80%	3.03%	2.73%
max	112.84%	76.96%	39.70%	127.18%

Table 4.15: Dogecoin Futures Normalised True Range

Solana

Solana's average normalised true range is the highest of the Futures contracts, with an average of 13.29%.

	Normalised True Range			
	FTX	*Bitfinex (Spot)*	*Coinbase Pro (Spot)*	*FTX (Spot)*
N	517	304	193	518
Mean	13.29%	12.54%	10.75%	13.38%
Std.	8.08%	9.11%	5.59%	13.38%
25%	8.20%	7.33%	7.06%	8.27%
50%	11.12%	10.09%	9.52%	11.38%
75%	15.97%	14.51%	12.68%	16.23%
min	3.04%	2.00%	3.23%	3.03%
max	77.83%	79.32%	37.92%	82.12%

Table 4.16: Solana Futures Normalised True Range

Chapter 5

Results

5.1 OLS

For the ordinary least-squares (OLS) regression, as noted in Chapter 3, we want to verify if the volatility and liquidity of Futures and Spot markets increase as the DTM variable approaches 0.

If the maturity effect holds, we would expect the coefficient β to be negative and statistically significant when using as a dependent variable the Normalised True Range, our volatility metric, as the maturity effects tell us that Futures price volatility may increase as the settlement date approaches, which in our cases means that the NTR is expected to increase as the DTM approaches 0.

5.1.1 Spread

If the liquidity is higher closer to the settlement, we expect the coefficient β of equation 5.1 to be positive and statistically significant.

$$AverageSpread = \beta \times DTM + c + e \tag{5.1}$$

where DTM is the days to maturity, β is the coefficient of variable DTM, c is the constant of the model, thus the value when DTM is 0, and e is the regression error with a mean of zero, constant variance and no autocorrelation.

For Bitcoin, out of the 21 Futures contract, 17 had a coefficient β deemed statistically significant, with 8 having a positive coefficient. For the Bitcoin Spot markets it was split into 23 sections, 16 had a coefficient β deemed statistically significant, with 8 having a positive coefficient.

For Ethereum, out of the 13 Futures contract, ten had a coefficient β deemed statistically significant, with seven having a positive coefficient. For the Ethereum Spot markets it was split into 21 sections, 16 had a coefficient β deemed statistically significant, with ten having a positive coefficient.

For Dogecoin, out of the seven Futures contract, all had a coefficient β deemed statistically significant, with five having a positive coefficient. For the Dogecoin Spot markets it was split into nine sections, and eight had a coefficient β deemed statistically significant, with six having a positive coefficient.

For Solana, out of the six Futures contract, five had a coefficient β deemed statistically significant, with three having a positive coefficient. For the Solana Spot markets it was split into 12 sections, 11 had a coefficient β deemed statistically significant, with nine having a positive coefficient.

The full regression values can also be found in Appendix A.1.

As seen in Table 5.1 when the DTM is higher, the spread on Ethereum, Dogecoin and Solana seem to be higher, meaning that as it gets closer to maturity, the spread also tends to decrease. As the Bid-Ask Spread is an illiquidity measurement, as the contracts get closer to maturity, both Spot and Futures markets tend to become more liquid. For Bitcoin, the coefficients seem to be mixed. However, this can be because Bitcoin's spread already tends to be very low, as seen in Table 4.5.

		N	Statistically significant (p<0.05)	Positive coefficient
BTC	Futures	21	17 (81%)	8 (47%)
	Spot	23	16 (70%)	8 (50%)
ETH	Futures	13	10 (77%)	7 (70%)
	Spot	21	16 (76%)	10 (63%)
DOGE	Futures	7	7 (100%)	5 (71%)
	Spot	9	8 (89%)	6 (75%)
SOL	Futures	6	5 (83%)	3 (60%)
	Spot	12	11 (92%)	9 (82%)

Table 5.1: OLS Bid-Ask Spread Summary

5.1.2 Depth

If the liquidity is higher closer to the settlement, we expect the coefficient β of equation 5.2 to be negative and statistically significant.

$$AverageDepth = \beta \times DTM + c + e \tag{5.2}$$

where DTM is the days to maturity, β is the coefficient of variable DTM, c is the constant of the model, thus the value when DTM is 0, and e is the regression error with a mean of zero, constant variance and no autocorrelation.

For Bitcoin, out of the 21 Futures contract, 19 had a coefficient β deemed statistically significant, with only six having a positive coefficient. For the Bitcoin Spot markets it was split into 23 sections, 21 had a coefficient β deemed statistically significant, with 11 having a positive coefficient.

For Ethereum, out of the 13 Futures contract, all had a coefficient β deemed statistically significant, with only one having a positive coefficient. For the Ethereum Spot markets it was split into 21 sections, 14 had a coefficient β deemed statistically significant, with only three having a positive coefficient.

For Dogecoin, out of the seven Futures contract, six had a coefficient β deemed statistically significant, with only one having a positive coefficient. For the Dogecoin Spot markets it was split into nine sections, and 8 had a coefficient β deemed statistically significant, with only two having a positive coefficient.

For Solana, out of the six Futures contract, all had a coefficient β deemed statistically significant, and all had a negative coefficient. For the Solana Spot markets it was split into 12 sections, nine had a coefficient β deemed statistically significant, with three having a positive coefficient.

The full regression values can also be found in Appendix A.2.

		N	Statistically significant ($p<0.05$)	Positive coefficient
BTC	Futures	21	19 (90%)	6 (32%)
	Spot	23	21 (91%)	11 (52%)
ETH	Futures	13	13 (100%)	1 (8%)
	Spot	21	14 (67%)	3 (21%)
DOGE	Futures	7	6 (86%)	1 (17%)
	Spot	9	8 (89%)	2 (25%)
SOL	Futures	6	6 (100%)	0 (0%)
	Spot	12	9 (75%)	3 (33%)

Table 5.2: OLS Depth Summary

As seen in Table 5.2 when the days to maturity (DTM) is higher, the depth on Ethereum, Dogecoin and Solana seems to be lower, meaning that as it gets closer to maturity, the depth tends to increase. As the depth is a measurement of liquidity available, as the contracts get closer to maturity, both Spot and Futures markets tend to become more liquid. For Bitcoin, the coefficients seem to be mixed with a slight tendency to decrease when the DTM is higher, but not to the same magnitude as the other cryptocurrencies.

5.1.3 NTR

If the maturity effect holds, we expect the coefficient β of equation 5.3 to be negative and statistically significant.

$$NTR = \beta \times DTM + c + e \tag{5.3}$$

where DTM is the days to maturity, β is the coefficient of variable DTM, c is the constant of the model, thus the value when DTM is 0, and e is the regression error with a mean of zero, constant variance and no autocorrelation.

For Bitcoin, out of the 21 Futures contract, only 11 had a coefficient β deemed statistically significant, with four having a positive coefficient. For the Bitcoin Spot markets it was split into 23 sections, 17 had a coefficient β deemed statistically significant, with six having a positive coefficient.

For Ethereum, out of the 13 Futures contract, only eight had a coefficient β deemed statistically significant, with three having a positive coefficient. For the Ethereum Spot markets it was split into 21 sections, 14 had a coefficient β deemed statistically significant, with five having a positive coefficient.

For Dogecoin, out of the seven Futures contract, only four had a coefficient β deemed statistically significant, with three having a positive coefficient. For the Dogecoin Spot markets it was split into nine sections, but only one had a coefficient β deemed statistically significant and a positive coefficient.

For Solana, out of the six Futures contract, only two had a coefficient β deemed statistically significant, and only one had a positive coefficient. For the Solana Spot markets it was split into 12 sections, but only five had a coefficient β deemed statistically significant, with two having a positive coefficient.

The full regression values can also be found in Appendix A.3.

As seen in Table 5.3 for Bitcoin and Ethereum when the DTM is higher, the NTR tends to be lower, indicating that there seems to be higher volatility in the Futures and Spot markets as the Futures contracts get closer to maturity. However, this is not as clear as with the liquidity metrics

		N	Statistically significant (p<0.05)	Positive coefficient
BTC	Futures	21	11 (52%)	4 (36%)
	Spot	23	17 (74%)	6 (35%)
ETH	Futures	13	8 (62%)	3 (38%)
	Spot	21	14 (67%)	5 (36%)
DOGE	Futures	7	4 (57%)	3 (75%)
	Spot	9	1 (11%)	1 (100%)
SOL	Futures	6	2 (33%)	1 (50%)
	Spot	12	5 (42%)	2 (40%)

Table 5.3: OLS NTR Summary

as the percentage of statistically significant coefficients is lower, and the percentage of positive coefficients is also lower. For Dogecoin and Solana, the percentage of statistically significant coefficients is too low to provide a reliable answer.

5.1.4 OLS summary

As seen in both Tables 5.2 and 5.1 for Ethereum, Dogecoin and Solana, as its contracts get closer to maturity, both the liquidity of those contracts and of its Spot markets tend to increase. In the case of Bitcoin the liquidity available at 1% depth seems to have a slight tendency to increase, while the spread has no clear trend.

For the volatility, Table 5.3 indicates that for Bitcoin and Ethereum the volatility tends to increase for Spot and Futures markets as the maturity of the Futures contracts gets closer. However, no substantial evidence was found for Dogecoin and Solana as the percentage of statistically significant coefficients is too low to provide a reliable answer.

5.2 Supervised Machine Learning

For all machine learning techniques, three distinct models were made: one to predict NTR, another to predict the average bid-ask spread and the last to predict the average liquidity within a 1% depth.

Furthermore, in order to optimise the hyper-parameters of the models, a grid search approach was taken. Grid search is one of the simplest approaches to optimising the hyper-parameters of an algorithm. It performs an exhaustive search through a manually specified set of hyper-parameters. (Hsu et al., 2003)

5.2.1 K-Nearest Neighbours

Optimisation

For the k-NN, the only parameter optimised was K, the parameter that determines how many neighbours are used to compute an observation's label. The value K used ranged between 1 and 30.

The average MAPE of the 3-fold cross-validation indicates that:

- for the NTR model, the best K was 16 with an average MAPE of 60.7% and an MAE of 0.0385.
- for the bid-ask spread model, the best K was one with an average MAPE of 111.3% and an MAE of 0.00061.
- for the liquidity at 1% depth model, the best K was 19 with an average MAPE of 38.8% and an MAE of 1713429.

Results

The following graphs present some sample predictions for eight random contracts or Spot market sections, using each model's optimal combinations of hyper-parameters.

The random predictions of the NTR model (5.1) had a MAPE of 58.3%, slightly better than the average obtained during validation.

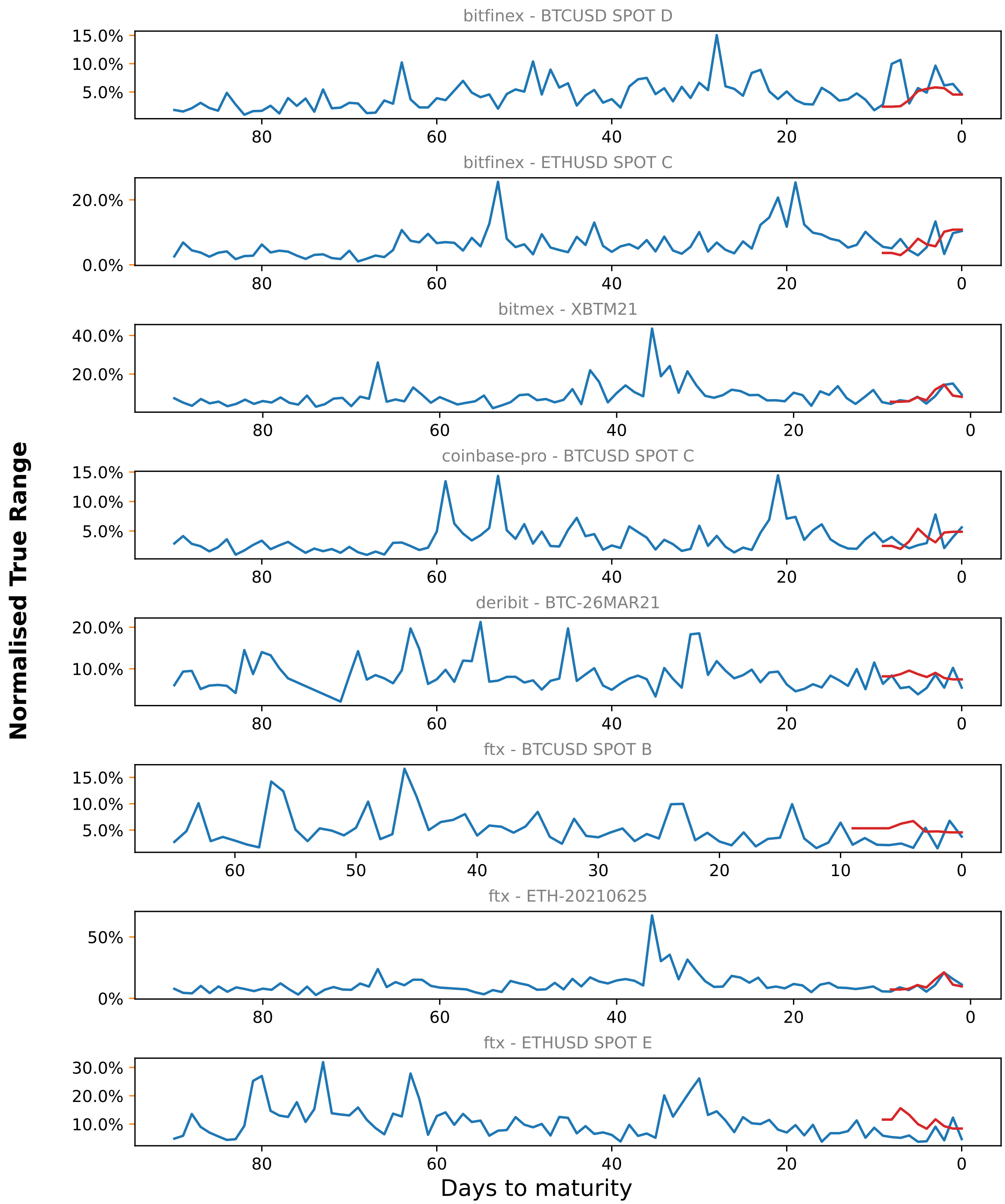

Figure 5.1: Random predictions of NTR using K-Nearest Neighbours

The random predictions of the average bid-ask spread model (5.2) had a MAPE of 75.7%, which, although better than the average obtained during validation, is still very poor.

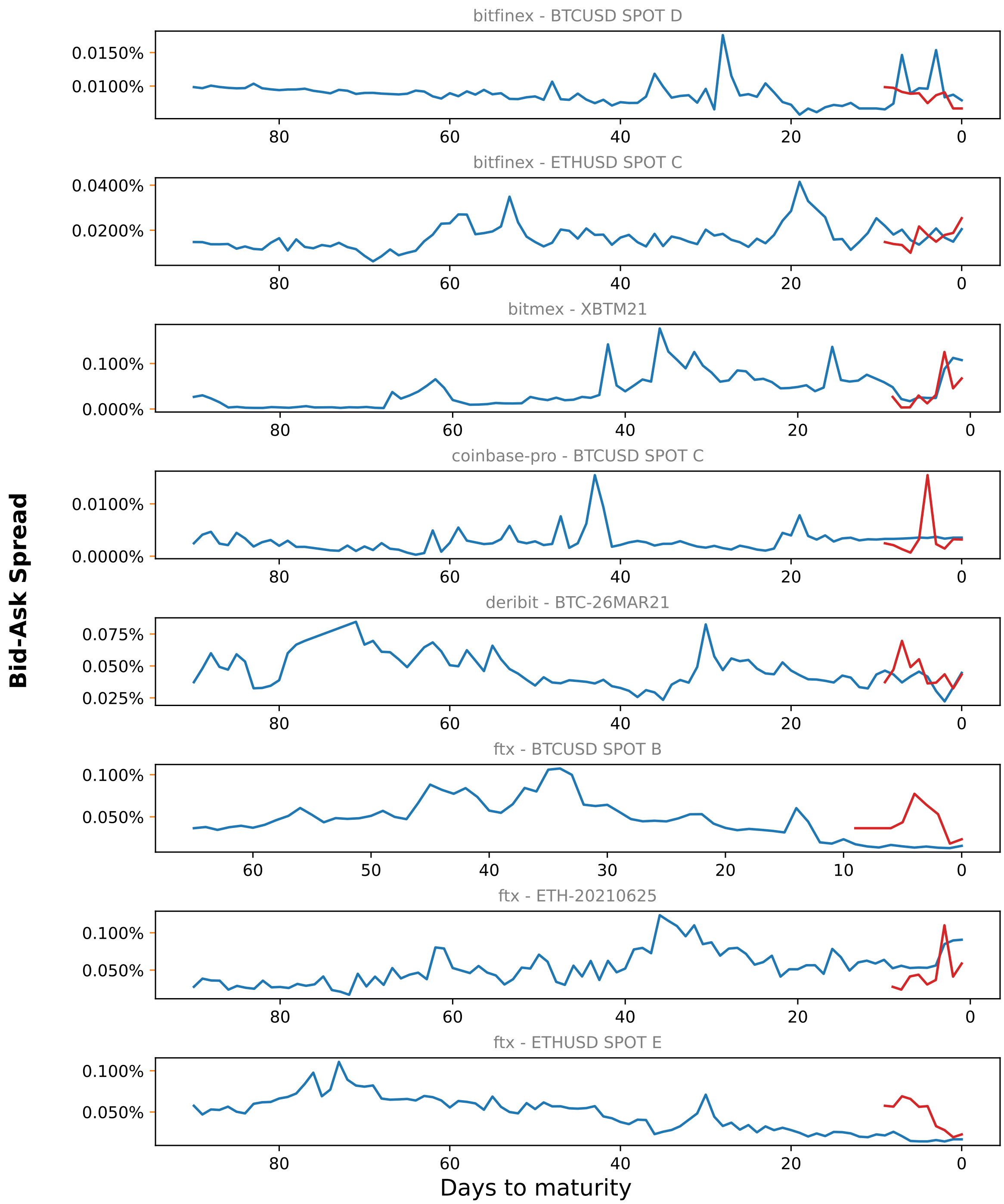

Figure 5.2: Random predictions of the average bid-ask spread using K-Nearest Neighbours

The random predictions of the model of the average liquidity at 1% depth (5.3) had a MAPE of 33.7%, slightly better than those obtained during the validation.

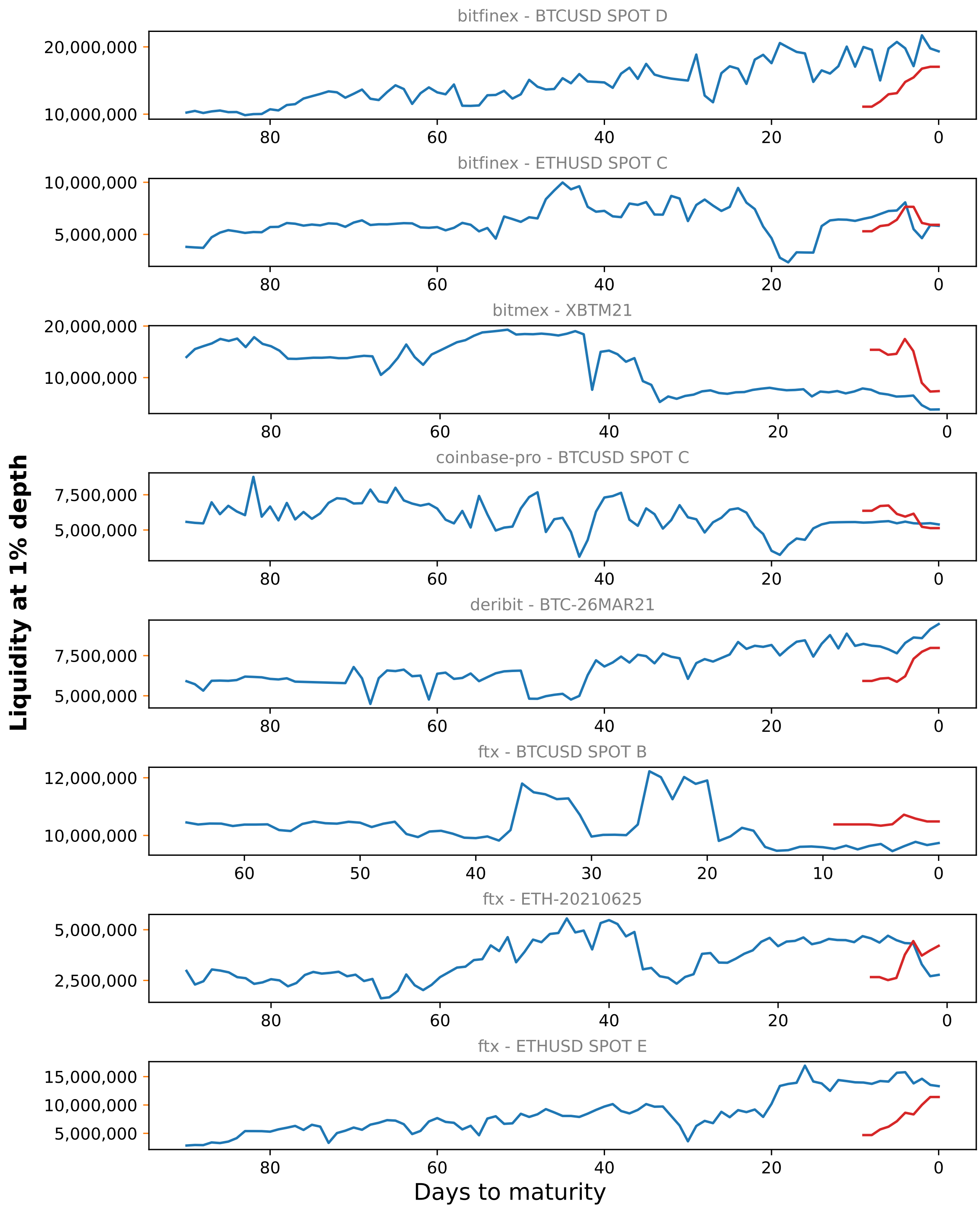

Figure 5.3: Random predictions of the average liquidity at 1% depth using K-Nearest Neighbours

5.2.2 Support Vector Regression

Optimisation

For the Support Vector Regression, three parameters were optimised, the ϵ, which is the parameter that defines within which margin the observations should be away from the hyperplane, and ranged from 0.00001 to 0.1, variable C, which defines how much the penalty will be for observations deviating away from the margin and ranged from 0.1 to 100, and variable γ which limits the influence that a single observation can have on defining the hyperplane function, and ranged from 0.0001 to 1.

However, for all models, the same non-linear kernel was used, the Radial basis function kernel, as it is one of the most widely used and yielded the best results during some brief testing.

The average MAPE of the 3-fold cross-validation indicates that:

- for the NTR model, the best parameters were a ϵ of 0.001, a C of 1 and a γ of 0.0001 with an average MAPE of 51.7% and an MAE of 0.04366.
- for the bid-ask spread model, the best parameters were a ϵ of 0.0001, a C of 10 and a γ of 0.0001 with an average MAPE of 65.9% and an MAE of 0.000232.
- for the liquidity at 1% depth model, the best parameters were a ϵ of 0.00001, a C of 0.1 and a γ of 0.1 with an average MAPE of 41.6% and an MAE of 1994478.

Results

The following graphs present some sample predictions for eight random contracts or Spot market sections, using each model's optimal combinations of hyper-parameters.

The random predictions of the NTR model (5.4) had a MAPE of 44.4%, a slight improvement over the results observed during validation.

The random predictions of the average bid-ask spread model (5.5) had a MAPE of 31.3%, a decent improvement over the results observed during validation. However, the result is still fairly poor.

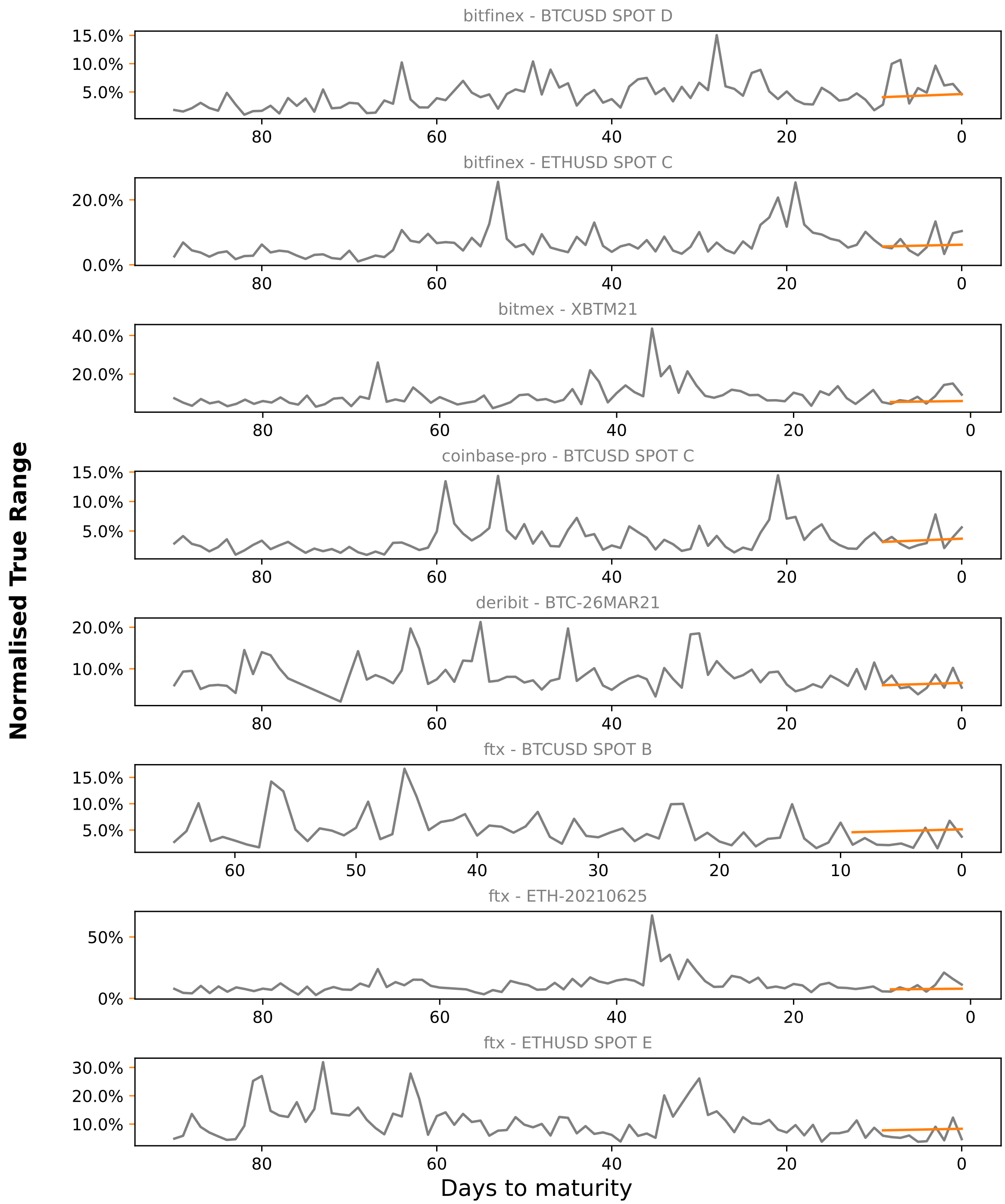

Figure 5.4: Random predictions of NTR using Support Vector Regression

The random predictions of the model of the average liquidity at 1% depth (5.6) had a MAPE of 36.5%, a slight improvement over the results observed during validation.

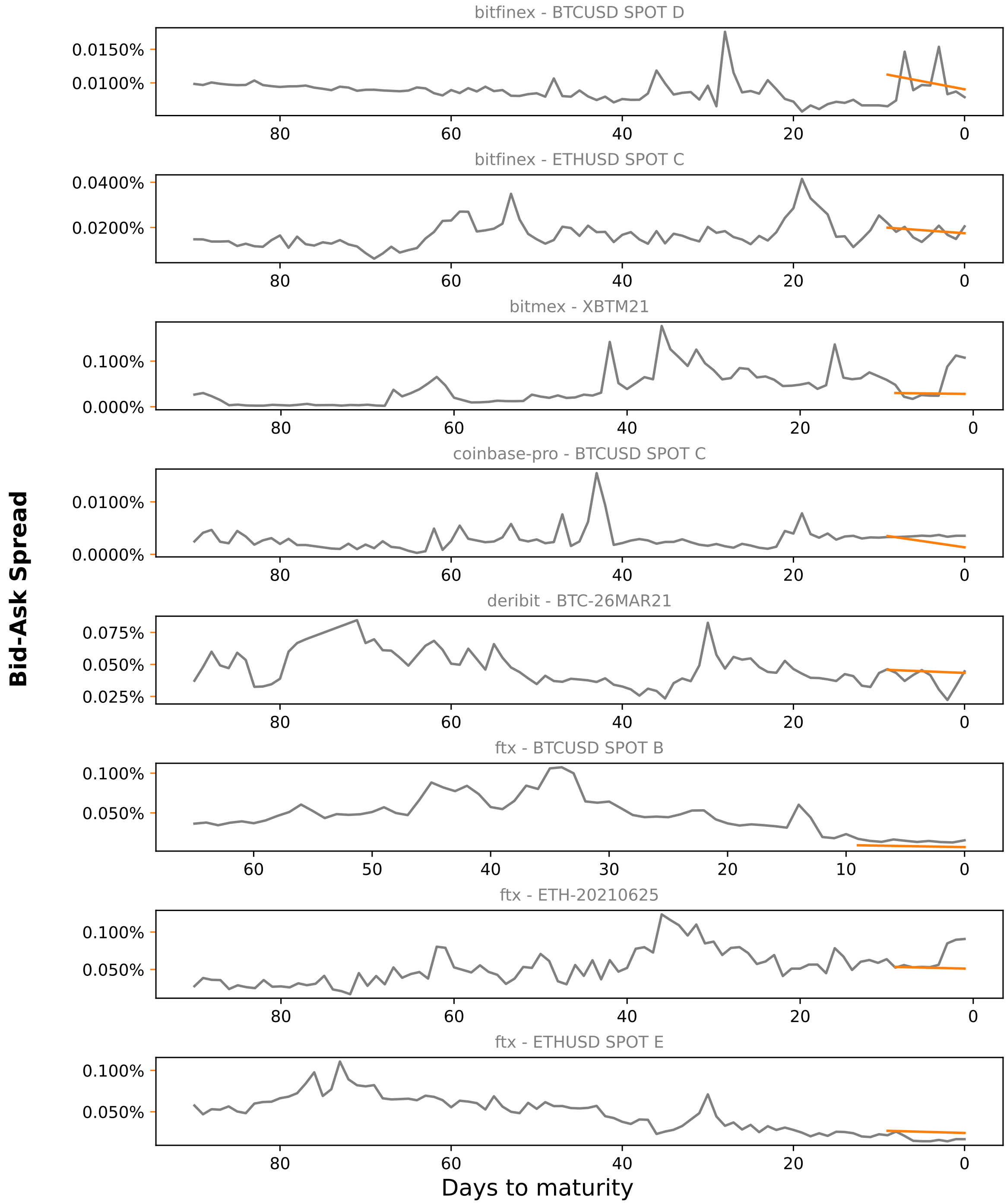

Figure 5.5: Random predictions of the average bid-ask spread using Support Vector Regression

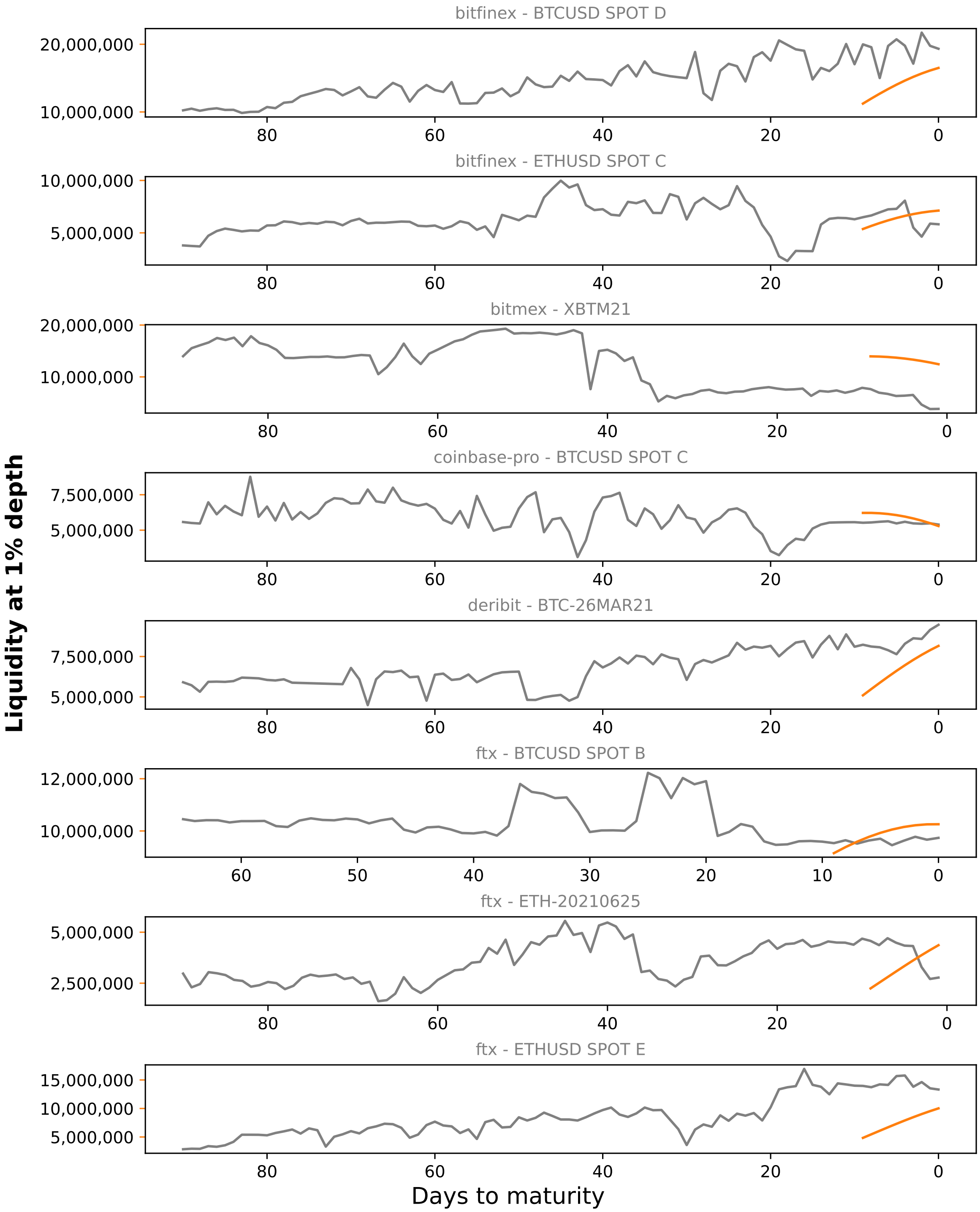

Figure 5.6: Random predictions of the average liquidity at 1% depth using Support Vector Regression

5.2.3 Gradient Boosting Machine

Optimisation

For the Gradient Boosting Machine, three parameters were optimised, the max depth, which defines the maximum depth of a tree and ranges from 3 to 9, and the learning range, which defines the step size shrinkage, which ranges from 0.05 to 0.25 and the number of estimators, which defines the number of gradient boosted trees and ranges from 100 to 400.

The average MAPE of the 3-fold cross-validation indicates that:

- for the NTR model, the best combination was a max depth of 9, a learning rate of 0.2 and 200 estimators with an average MAPE of 37.3% and an MAE of 0.0316.
- for the bid-ask spread model, the best combination was a max depth of 3, a learning rate of 0.2 and 200 estimators with an average MAPE of 56.6% and an MAE of 0.0002455.
- for the liquidity at 1% depth model, the best combination was a max depth of 9, a learning rate of 0.2 and 200 estimators with an average MAPE of 35.2% and an MAE of 1199274.

Results

The following graphs present some sample predictions for eight random contracts or Spot market sections, using each model's optimal combinations of hyperparameters.

The random predictions of the NTR model (5.7) had a MAPE of 25.9%, slightly better than the average obtained during validation.

The random predictions of the average bid-ask spread model (5.8) had a MAPE of 38.6%, slightly better than the average obtained during validation.

The random predictions of the model of the average liquidity at 1% depth (5.9) had a MAPE of 12.3%, a decent improvement over the average obtained during validation.

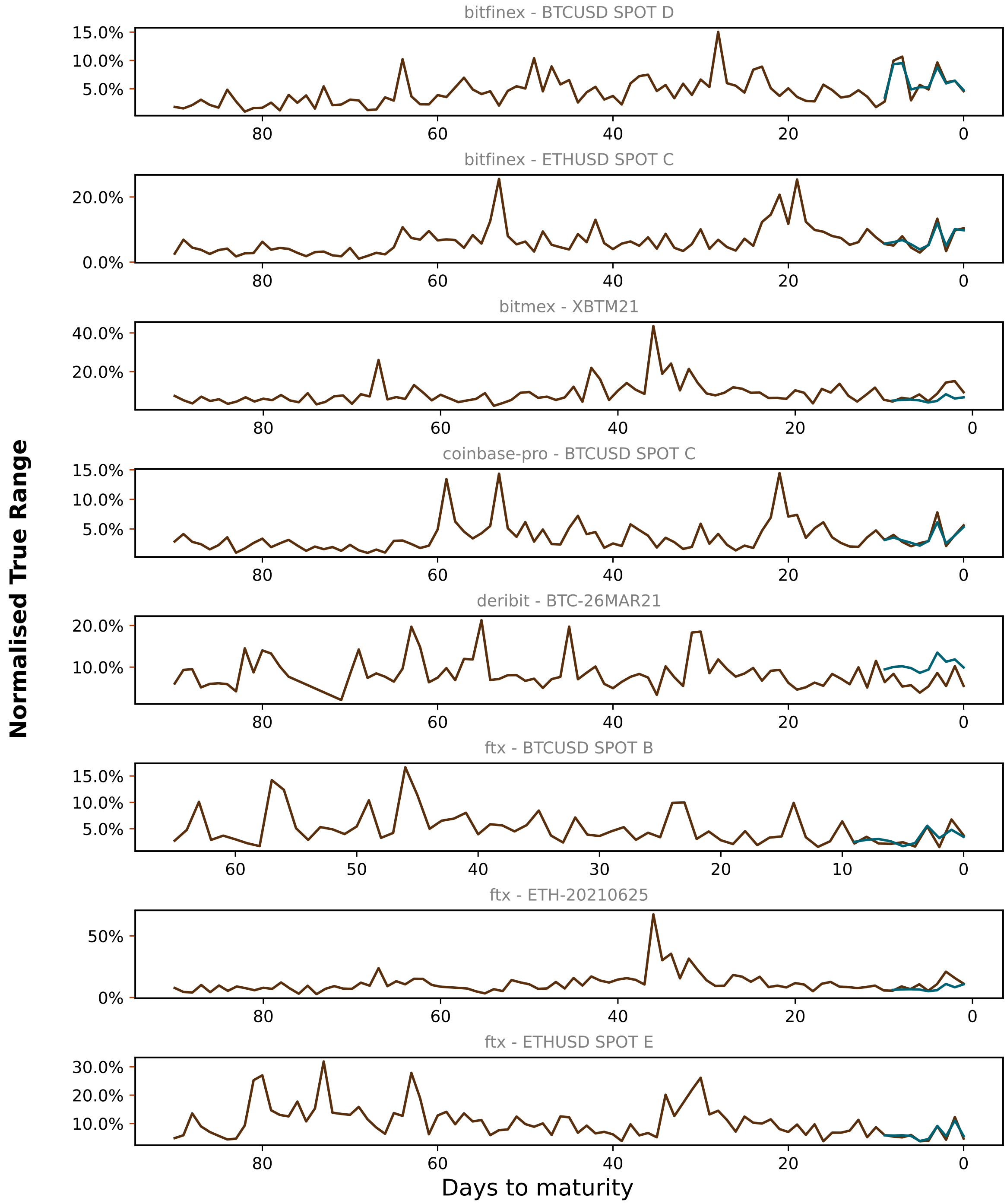

Figure 5.7: Random predictions of NTR using XGBoost model

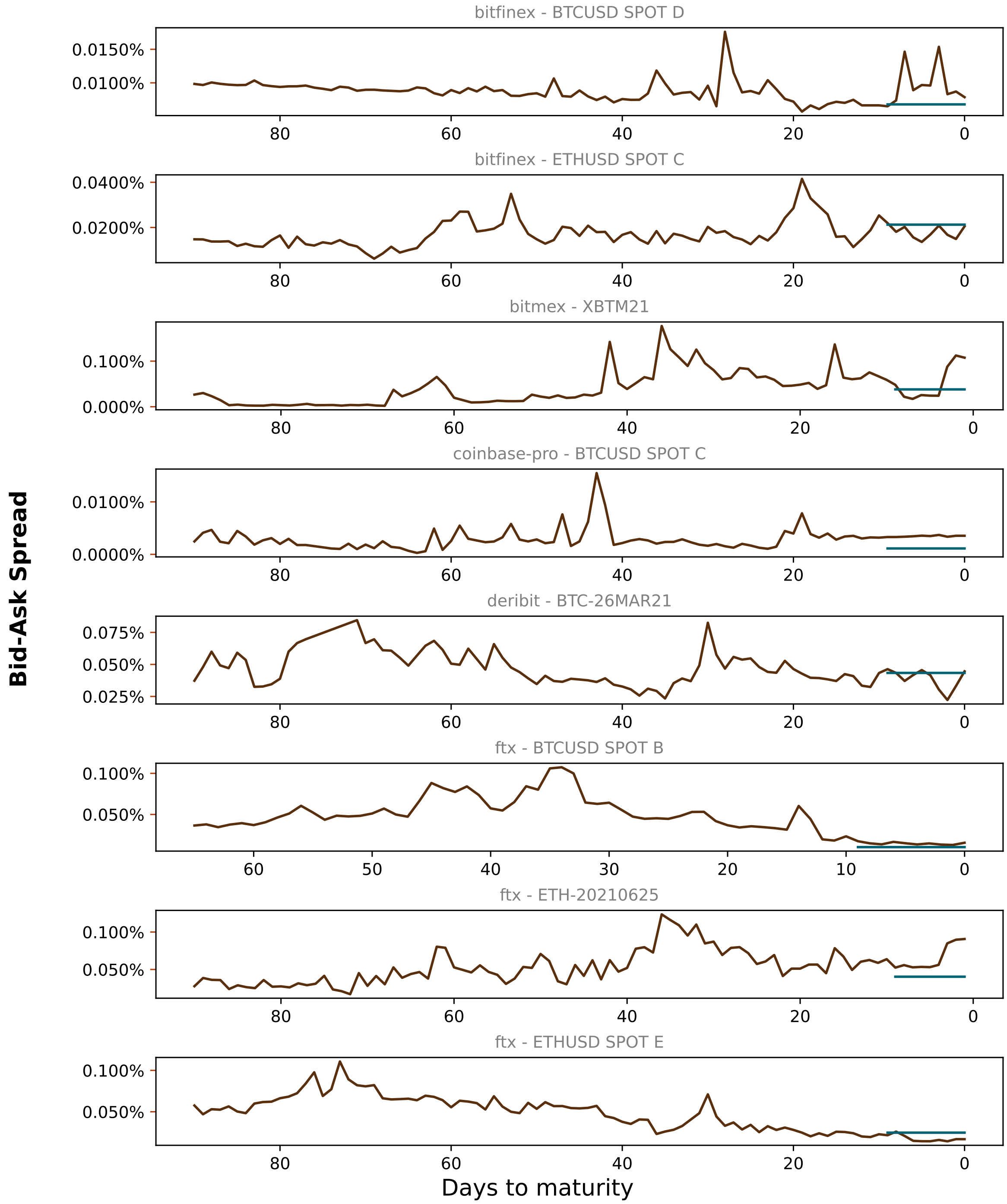

Figure 5.8: Random predictions of the average bid-ask spread using XGBoost model

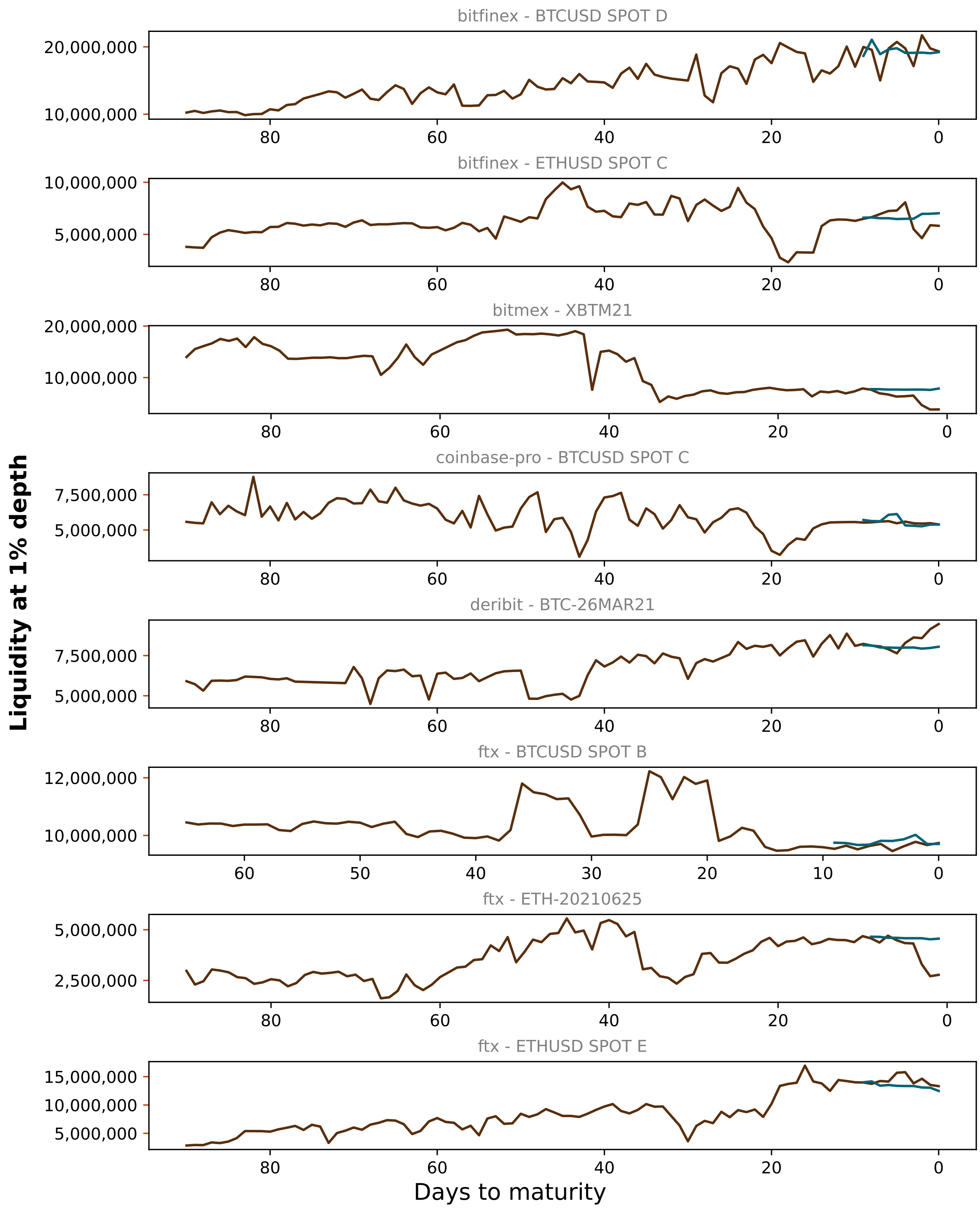

Figure 5.9: Random predictions of the average liquidity at 1% depth using XGBoost model

5.2.4 Supervised Machine Learning Summary

Machine Learning Technique	Metric	MAPE	MAE
K-Nearest Neighbours	NTR	60.7%	0.0385
	bid-ask spread	111.3%	0.00061
	liquidity at 1% depth	38.8%	1713429
Support Vector Regression	NTR	51.7%	0.04366
	bid-ask spread	65.9%	0.000232
	liquidity at 1% depth	41.6%	1994478
Gradient Boosting Machine	NTR	37.3%	0.0316
	bid-ask spread	56.6%	0.0002455
	liquidity at 1% depth	35.2%	1199274

Table 5.4: Supervised Machine Learning Summary

Surprisingly, in all cases, these techniques produced better predictions for the NTR than for the average bid-ask spread, even though the previous OLS model had more statistically significant β coefficients for the average bid-ask spread than for the NTR. Similarly, the OLS model also indicated that the average liquidity available at 1% depth had a strong trend. Again, however, predictions for these models were not robust.

However, this could be explained partially by the fact that NTR presents a lower standard deviation of 7.271 % compared to its mean of 8.426 %. At the same time, the average liquidity at 1% depth had a standard deviation of 8,176,625 US$ and a mean of 6,636,020 US $ and the average bid-ask spread had a standard deviation of 0.253 % and a mean of 0.096 %. Nonetheless, these techniques were unable to present great predictions.

The technique that yielded the best results was Gradient Boosting Machine. In some cases, as seen in Table 5.8 the model made predictions with some trend without considering any short-term fluctuations. In contrast, in the case of Tables 5.7 and 5.9 it seemed to better take into consideration possible short-term fluctuations near the settlement, and it was able to obtain an acceptable MAPE of 25.9% and a decent MAPE of 12.3%, respectively.

Both in the case of K-Nearest Neighbours and Support Vector Regression, as seen in the random predictions, most of the time, it seems to predict too many short-term fluctuations that might indicate some over-fitting, which also contributes to a less than ideal MAPE.

Chapter 6

Conclusion

6.1 Final Remarks

Although cryptocurrency and its Futures markets have exploded in popularity since the introduction of Bitcoin by Nakamoto (2008), the maturity effect on these Futures has yet to be studied. Furthermore, no study has yet tried to understand if the maturity of cryptocurrency Futures contracts can predict the volatility and liquidity of its underlying Spot and Futures markets.

The goal of this dissertation was to understand if the maturity effect is present on the Futures of the selected cryptocurrencies and if the maturity of the cryptocurrency Futures contracts could predict the volatility and liquidity of its underlying Spot and Futures markets.

The results from the OLS regression indicate that for Bitcoin and Ethereum, the volatility metric, the normalised true range (NTR), tends to be lower when the days left to maturity are higher. Thus the volatility tends to increase as the Futures approximate the settlement date. However, for Dogecoin and Solana, the percentage of statistically significant coefficients is too low to provide a reliable answer. This means that the maturity effect is present to some extent on the Futures and Spot markets of Bitcoin and Ethereum.

On the other hand, the results are more robust regarding liquidity metrics. Both liquidity metrics, the bid-ask spread and the liquidity at 1% depth indicate that for Ethereum, Solana and Dogecoin, the liquidity tends to increase as the Futures approximates the settlement date. In contrast, for Bitcoin it tends to remain roughly constant. Furthermore, the descriptive analysis indicates that, on average Spot markets tend to be more liquidity than their Futures counterparts for all coins studied in this dissertation.

All machine learning techniques used, K-Nearest Neighbours, Support Vector Regression and Gradient Boosting Machine, had the best performance metrics when predicting the liquidity at 1% depth, followed by the NTR and then by the bid-ask spread.

Surprisingly, all techniques produced better predictions for the NTR than for the average bid-ask spread, even though the OLS regression had more statistically significant β coefficients for the average bid-ask spread than for the NTR. Similarly, the OLS model also indicated that the average liquidity available at 1% depth had a strong trend. However, predictions for this metric were not robust.

Although none presented admirable predictive results, Gradient Boosting Machine was the technique that yielded the most promising results, with its best result being a MAPE of 35% for the liquidity at 1% depth prediction.

In summary, there is evidence that the maturity effect is present in Bitcoin and Ethereum's markets. Furthermore, there is strong evidence that the liquidity increases as the settlement date approximates. However, none of the machine learning techniques used in this dissertation were able to produce adequate predictions for either the volatility or the liquidity of cryptocurrency Futures and their underlying Spot markets.

6.2 Limitations and Future Research

For this dissertation, only daily observations were used to build the models. Moreover, only quarterly Futures contracts, which settle once per quarter, were considered. Therefore, there is an opportunity for future work using data with lower time frames, such as hourly data, and data from other Futures contracts, such as monthly and semi-annual contracts.

Additionally, some more complex techniques could be used to try and build a better predictive model. For instance, Deep learning techniques are consistently considered one of the best performing predictors within the machine learning field for financial time series forecasting. They are considered superior to traditional time series prediction models and other machine learning techniques since they are able to better capture noise and small features in the data, making them ideal for complex time series, such as financial time series (Sezer, Gudelek, & Ozbayoglu, 2020; Yan & Ouyang, 2018). According to a survey done by Sezer et al. (2020), the long short-term memory (LSTM) algorithm, a type of recurrent neural network (RNN), is the dominating type of deep learning model for financial time series forecasting. Thus, building a predictive model using this or a similar technique might be of interest.

Bibliography

Alexander, C., Choi, J., Park, H., & Sohn, S. (2020). Bitmex bitcoin derivatives: Price discovery, informational efficiency, and hedging effectiveness. *Journal of Futures Markets*, *40*(1), 23–43.

Alexander, C., & Heck, D. F. (2020). Price discovery in bitcoin: The impact of unregulated markets. *Journal of Financial Stability*, *50*, 100776.

Alkhatib, K., Najadat, H., Hmeidi, I., & Shatnawi, M. K. A. (2013). Stock price prediction using k-nearest neighbor (knn) algorithm. *International Journal of Business, Humanities and Technology*, *3*(3), 32–44.

Altman, N. S. (1992). An introduction to kernel and nearest-neighbor nonparametric regression. *The American Statistician*, *46*(3), 175–185.

Amihud, Y., Mendelson, H., & Lauterbach, B. (1997). Market microstructure and securities values: Evidence from the tel aviv stock exchange. *Journal of financial Economics*, *45*(3), 365–390.

Anderson, R. W. (1985). Some determinants of the volatility of futures prices. *The Journal of Futures Markets (pre-1986)*, *5*(3), 331.

Bartram, S. M., Brown, G. W., & Fehle, F. R. (2009). International evidence on financial derivatives usage. *Financial management*, *38*(1), 185–206.

Bentéjac, C., Csörgő, A., & Martínez-Muñoz, G. (2021). A comparative analysis of gradient boosting algorithms. *Artificial Intelligence Review*, *54*(3), 1937–1967.

Bickel, P. J., & Doksum, K. A. (2015). *Mathematical statistics: basic ideas and selected topics, volumes i-ii package*. Chapman and Hall/CRC.

Binance. (2019, Sep). *Introduction to insurance funds.* Retrieved from https://www.binance.com/

en/support/faq/360033525371

Binance. (2020, Jun). *What is auto-deleveraging (adl) and how does it work.* Retrieved from https://www.binance.com/en/support/faq/360033525471

BitMEX. (2016). *Xbtusd: Contract specification.* Retrieved from https://www.bitmex.com/app/contract/XBTUSD

BitMEX. (2020, Mar). *Bitmex insurance fund: Your questions answered.* Retrieved from https://blog.bitmex.com/bitmex-insurance-fund-your-questions-answered/

Bragin, A. (2015). Inverse futures in bitcoin economy. *Available at SSRN 2713755*.

Bybit. (2021, Aug). *What is auto-deleveraging (adl)?* Retrieved from https://help.bybit.com/hc/en-us/articles/900000031623-What-is-Auto-Deleveraging-ADL-

Cao, L., & Tay, F. E. (2001). Financial forecasting using support vector machines. *Neural Computing & Applications*, *10*(2), 184–192.

Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 785–794.

Chuen, D. L. K., Guo, L., & Wang, Y. (2017). Cryptocurrency: A new investment opportunity? *The Journal of Alternative Investments*, *20*(3), 16–40.

Chui, M. (2012). Derivatives markets, products and participants: an overview. *Ifc Bullefin*, *35*, 3–12.

Ciaian, P., & Rajcaniova, M. e. a. (2018). Virtual relationships: Short-and long-run evidence from bitcoin and altcoin markets. *Journal of International Financial Markets, Institutions and Money*, *52*, 173–195.

CME. (n.d.). *Get to know futures expiration and settlement.* CME Group. Retrieved from https://www.cmegroup.com/education/courses/introduction-to-futures/get-to-know-futures-expiration-and-settlement.html

Coingecko.com. (n.d.). https://www.coingecko.com/. (Accessed: 2021-10-30)

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning*, *20*(3), 273–297.

Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE transactions on information theory*, *13*(1), 21–27.

Daal, E., Farhat, J., & Wei, P. P. (2006). Does futures exhibit maturity effect? new evidence from an extensive set of us and foreign futures contracts. *Review of Financial Economics*, *15*(2), 113–128.

De Myttenaere, A., Golden, B., Le Grand, B., & Rossi, F. (2016). Mean absolute percentage error for regression models. *Neurocomputing*, *192*, 38–48.

Derbentsev, V., Matviychuk, A., Datsenko, N., Bezkorovainyi, V., & Azaryan, A. (2020). Machine learning approaches for financial time series forecasting. *Proceedings of the Selected Papers of the Special Edition of International Conference on Monitoring, Modeling & Management of Emergent Economy (M3E2-MLPEED 2020)*, 434–450.

Dismuke, C., & Lindrooth, R. (2006). Ordinary least squares. *Methods and Designs for Outcomes Research*, *93*, 93–104.

Drucker, H., Burges, C. J., Kaufman, L., Smola, A., & Vapnik, V. (1996). Support vector regression machines. *Advances in neural information processing systems*, *9*.

Fama, E. F., & French, K. R. (1988). Business cycles and the behavior of metals prices. *The Journal of Finance*, *43*(5), 1075–1093.

Feurer, M., & Hutter, F. (2019). *Automated machine learning.* Springer, Cham.

FINRA. (n.d.). *Understanding margin accounts, why brokers do what they do.* Retrieved from https://www.finra.org/investors/learn-to-invest/advanced-investing/understanding-margin-accounts

Fix, E., & Hodges, J. L. (1951). Discriminatory analysis. nonparametric discrimination: Consistency properties. *USAF School of Aviation Medicine, Randolph Field, Texas*.

Fleming, M. J. (2001). Measuring treasury market liquidity. *FRB of New York Staff Report*(133).

Forman, J. (2006). Cross-market evaluations with normalized average true range. *TECHNICAL ANALYSIS OF STOCKS AND COMMODITIES-MAGAZINE EDITION-*, *24*(6), 60.

Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of statistics*, 1189–1232.

FTX. (2022, Jan). *Backstop liquidity provider program.* Retrieved from https://help.ftx.com/hc/en-us/articles/360024479392-Backstop-Liquidity-Provider-Program

Galloway, T. M., & Kolb, R. W. (1996). Futures prices and the maturity effect. *The Journal of Futures Markets (1986-1998)*, *16*(7), 809.

Gonzalo, J., & Granger, C. (1995). Estimation of common long-memory components in cointegrated systems. *Journal of Business & Economic Statistics*, *13*(1), 27–35.

Hasbrouck, J. (1995). One security, many markets: Determining the contributions to price discovery. *The journal of Finance*, *50*(4), 1175–1199.

Heckinger, R., Mengle, D., Steigerwald, R., Ruffini, I., & Wells, K. (2013). Understanding derivatives: Markets and infrastructure. *Federal Reserve Bank of Chicago*.

Herbert, J. H. (1995). Trading volume, maturity and natural gas futures price volatility. *Energy Economics*, *17*(4), 293–299.

Hilpisch, Y. (2015). *Derivatives analytics with python: data analysis, models, simulation, calibration and hedging*. John Wiley & Sons.

Hsu, C.-W., Chang, C.-C., Lin, C.-J., et al. (2003). *A practical guide to support vector classification.* Taipei, Taiwan.

Hull, J. C. (2003). *Options futures and other derivatives*. Pearson Education India.

Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International journal of forecasting*, *22*(4), 679–688.

Ince, H., & Trafalis, T. B. (2006). A hybrid model for exchange rate prediction. *Decision Support Systems*, *42*(2), 1054–1062.

Jakub, B. (2015). Does bitcoin follow the hypothesis of efficient market. *International Journal of Economic Sciences*, *4*(2), 10–23.

Kim, K.-j. (2003). Financial time series forecasting using support vector machines. *Neurocomputing*, *55*(1-2), 307–319.

Kramer, O. (2013). *Dimensionality reduction with unsupervised nearest neighbors* (Vol. 51). Springer.

Lien, D., & Shrestha, K. (2009). A new information share measure. *Journal of Futures Markets: Futures, Options, and Other Derivative Products*, *29*(4), 377–395.

Lora, A. T., Santos, J. M. R., Expósito, A. G., Ramos, J. L. M., & Santos, J. C. R. (2007). Electricity market price forecasting based on weighted nearest neighbors techniques. *IEEE Transactions on Power Systems*, *22*(3), 1294–1301.

Lynch, T. E. (2011). Derivatives: A twenty-first century understanding. *Loy. U. Chi. LJ*, *43*, 1.

Miller, K. D. (1972). The relation between volatility and maturity in futures contracts.

Milonas, N. T. (1986). Price variability and the maturity effect in futures markets. *Journal of futures markets*, *6*(3), 443–460.

Moosa, I. A., & Bollen, B. (2001). Is there a maturity effect in the price of the s&p 500 futures contract? *Applied Economics Letters*, *8*(11), 693–695.

Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system. *Cryptography Mailing list at metzdowd.com*.

Noble, W. S. (2006). What is a support vector machine? *Nature biotechnology*, *24*(12), 1565–1567.

Okasha, M. K. (2014). Using support vector machines in financial time series forecasting. *International Journal of Statistics and Applications*, *4*(1), 28–39.

Pernice, I. G. A., & Scott, B. (2021). Cryptocurrency. *Internet Policy Review, Glossary of decentralised technosocial systems*, *10*(2).

Pradhan, A. (2012). Support vector machine-a survey. *International Journal of Emerging Technology and Advanced Engineering*, *2*(8), 82–85.

Samuelson, P. A. (1965). Proof that properly anticipated prices fluctuate randomly. *Management Review*, *6*(2).

Serletis, A. (1992). Maturity effects in energy futures. *Energy Economics*, *14*(2), 150–157.

Sezer, O. B., Gudelek, M. U., & Ozbayoglu, A. M. (2020). Financial time series forecasting with deep learning: A systematic literature review: 2005–2019. *Applied Soft Computing*, *90*, 106181.

Smola, A. J., & Schölkopf, B. (2004). A tutorial on support vector regression. *Statistics and computing*, *14*(3), 199–222.

Srinath, K. (2017). Python–the fastest growing programming language. *International Research Journal of Engineering and Technology (IRJET)*, *4*(12), 354–357.

Stange, S., & Kaserer, C. (2009). Market liquidity risk-an overview. *CEFS working paper series*, *4*.

Stoll, H. R., & Whaley, R. E. (1987). Program trading and expiration-day effects. *Financial Analysts Journal*, *43*(2), 16–28.

Su, E., & Tokmakcioglu, K. (2021). A comparison of bid-ask spread proxies and determinants of bond bid-ask spread. *Borsa Istanbul Review*, *21*(3), 227–238.

Suits, D. B. (1957). Use of dummy variables in regression equations. *Journal of the American*

Statistical Association, *52*(280), 548–551.

Sun, X., Liu, M., & Sima, Z. (2020). A novel cryptocurrency price trend forecasting model based on lightgbm. *Finance Research Letters*, *32*, 101084.

Suthaharan, S. (2016). Support vector machine. *Machine learning models and algorithms for big data classification*, 207–235.

Tang, L., Pan, H., & Yao, Y. (2018). K-nearest neighbor regression with principal component analysis for financial time series prediction. *Proceedings of the 2018 International Conference on Computing and Artificial Intelligence*, 127–131.

Taunk, K., De, S., Verma, S., & Swetapadma, A. (2019). A brief review of nearest neighbor algorithm for learning and classification. *2019 International Conference on Intelligent Computing and Control Systems (ICCS)*, 1255–1260.

Tay, F. E., & Cao, L. (2001). Application of support vector machines in financial time series forecasting. *omega*, *29*(4), 309–317.

Tsantekidis, A., Passalis, N., Tefas, A., Kanniainen, J., Gabbouj, M., & Iosifidis, A. (2017). Forecasting stock prices from the limit order book using convolutional neural networks. *2017 IEEE 19th Conference on Business Informatics (CBI)*, *1*, 7–12.

Whaley, R. E. (2009). Understanding the vix. *The Journal of Portfolio Management*, *35*(3), 98–105.

Wu, Y. (2021). A quantitative analysis on bitmex perpetual inverse futures xbtusd contract. *Undergraduate Economic Review*, *17*(1), 12.

Yan, H., & Ouyang, H. (2018). Financial time series prediction based on deep learning. *Wireless Personal Communications*, *102*(2), 683–700.

Appendix A

Appendix

A.1 Metrics

A.1.1 Liquidity metrics

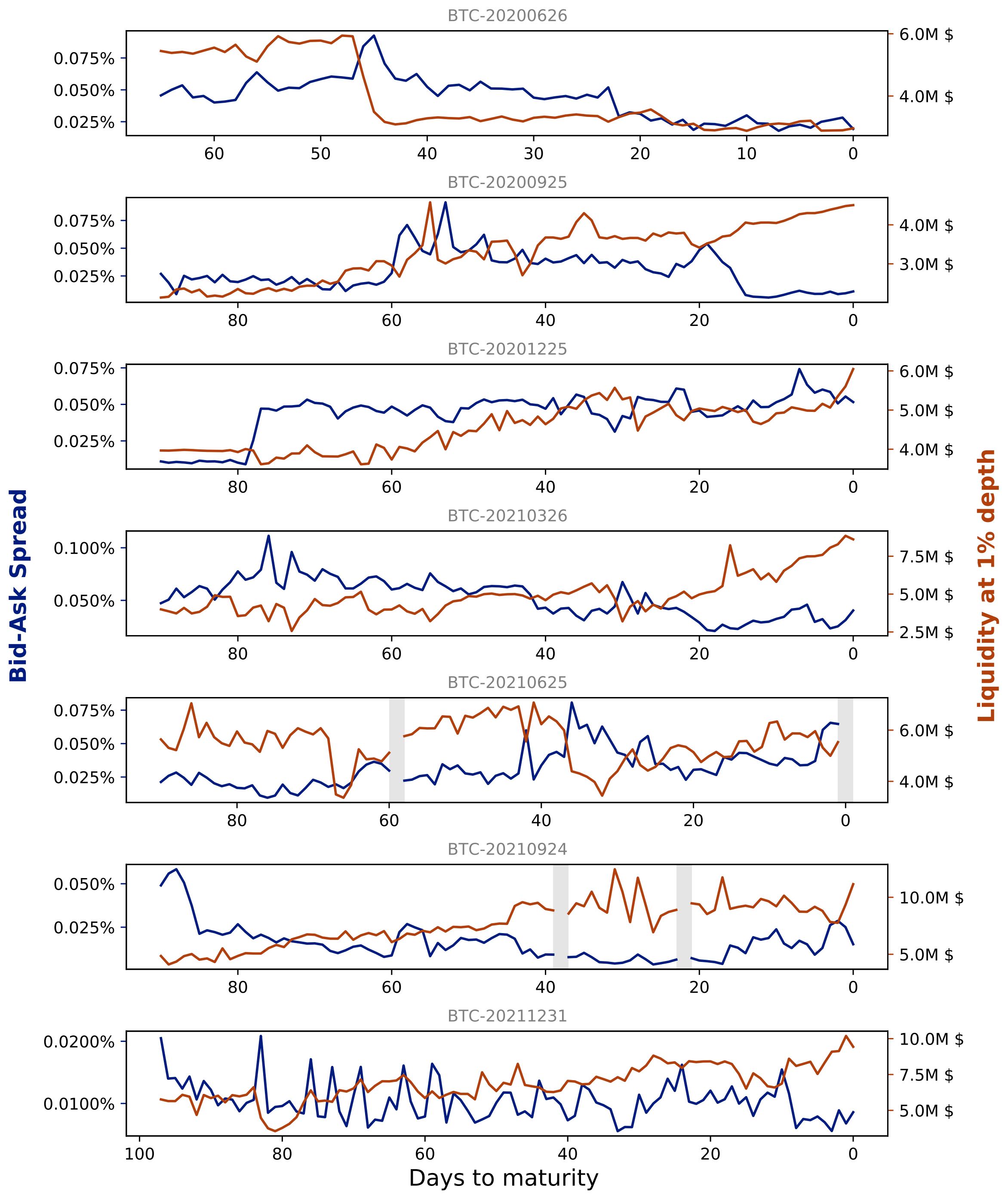

Figure A.1: Liquidity metrics of FTX's Bitcoin Futures contracts

A.1.2 Volatility metrics

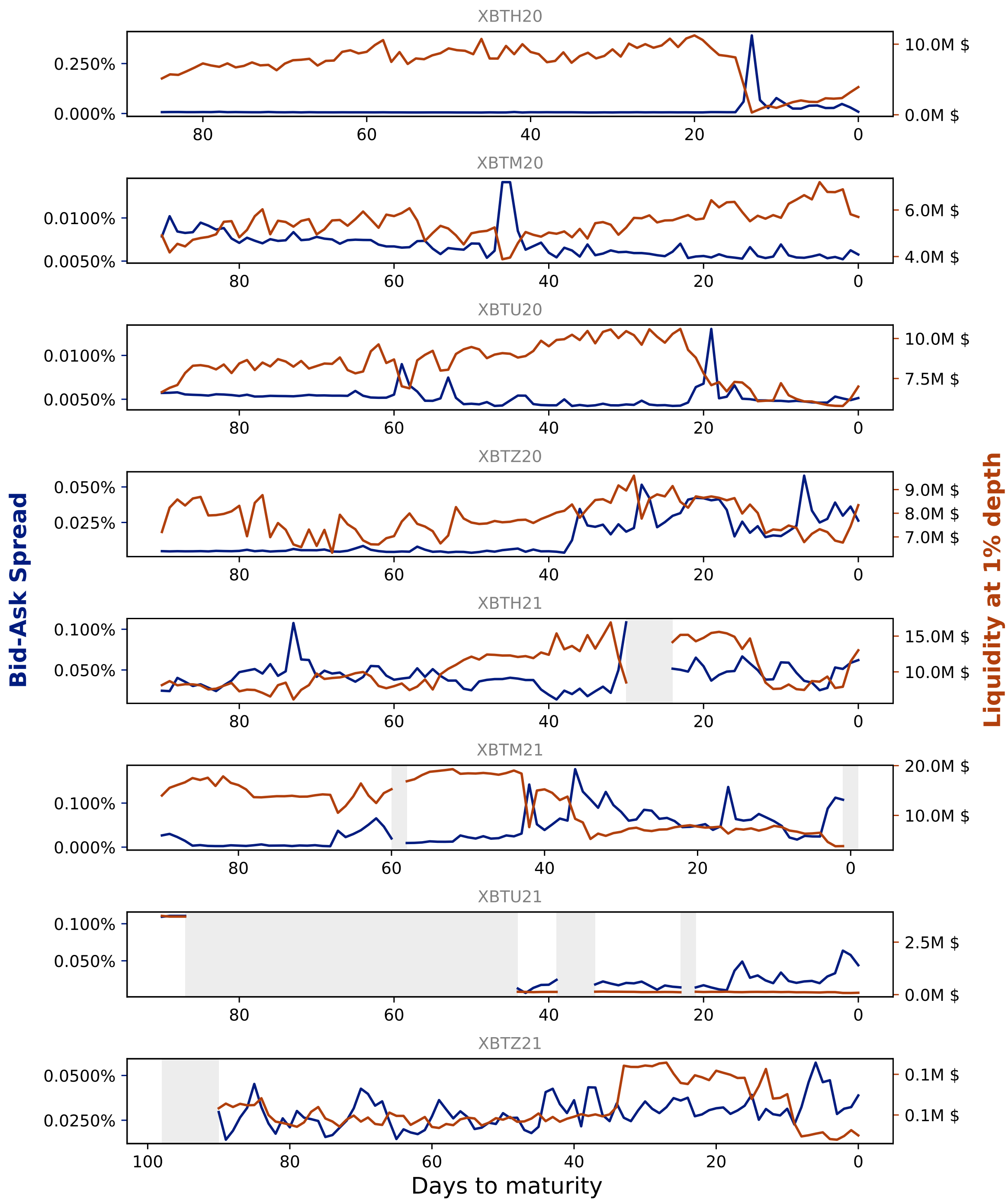

Figure A.2: Liquidity metrics of Bitmex's Bitcoin Futures contracts

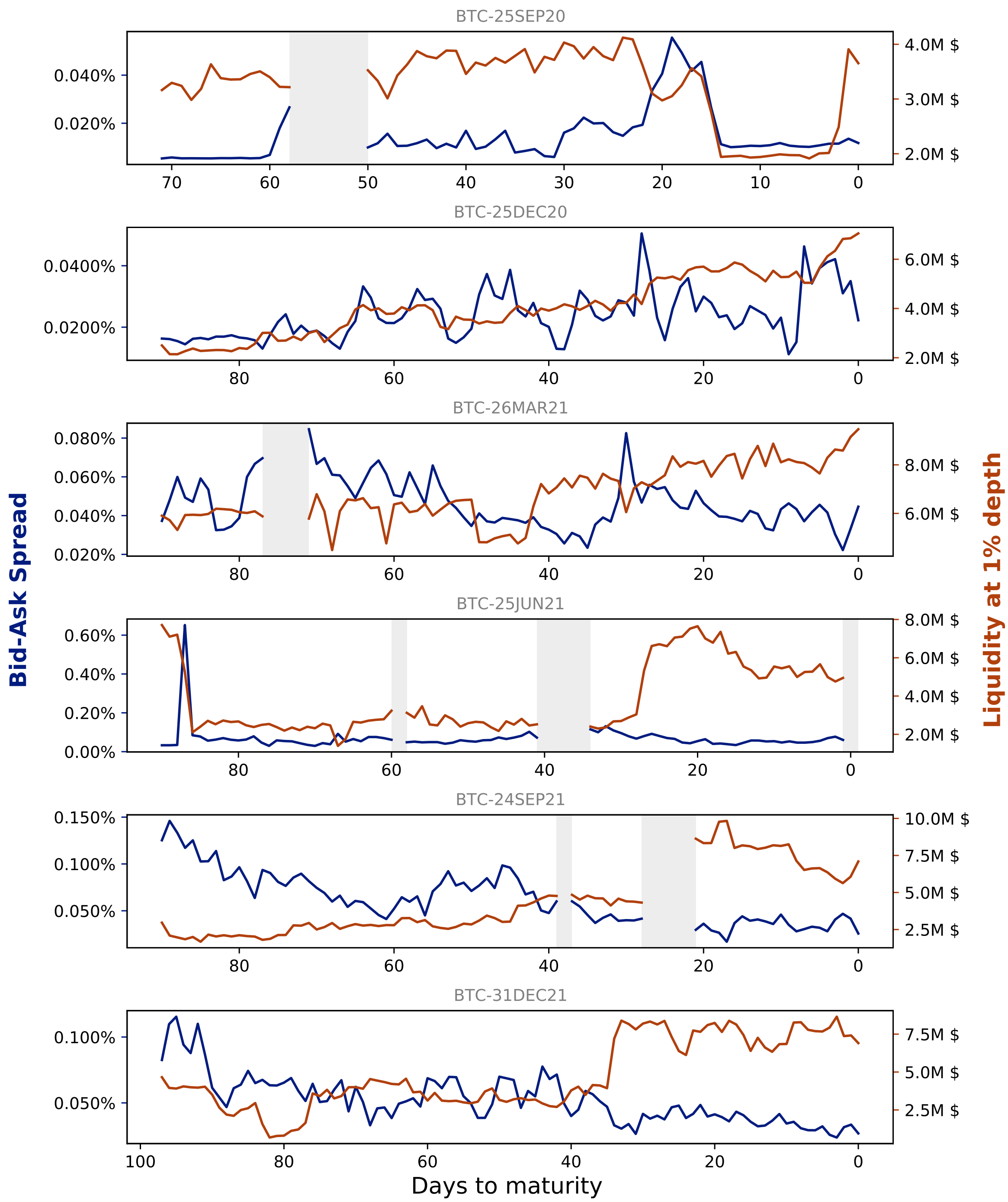

Figure A.3: Liquidity metrics of Deribit's Bitcoin Futures contracts

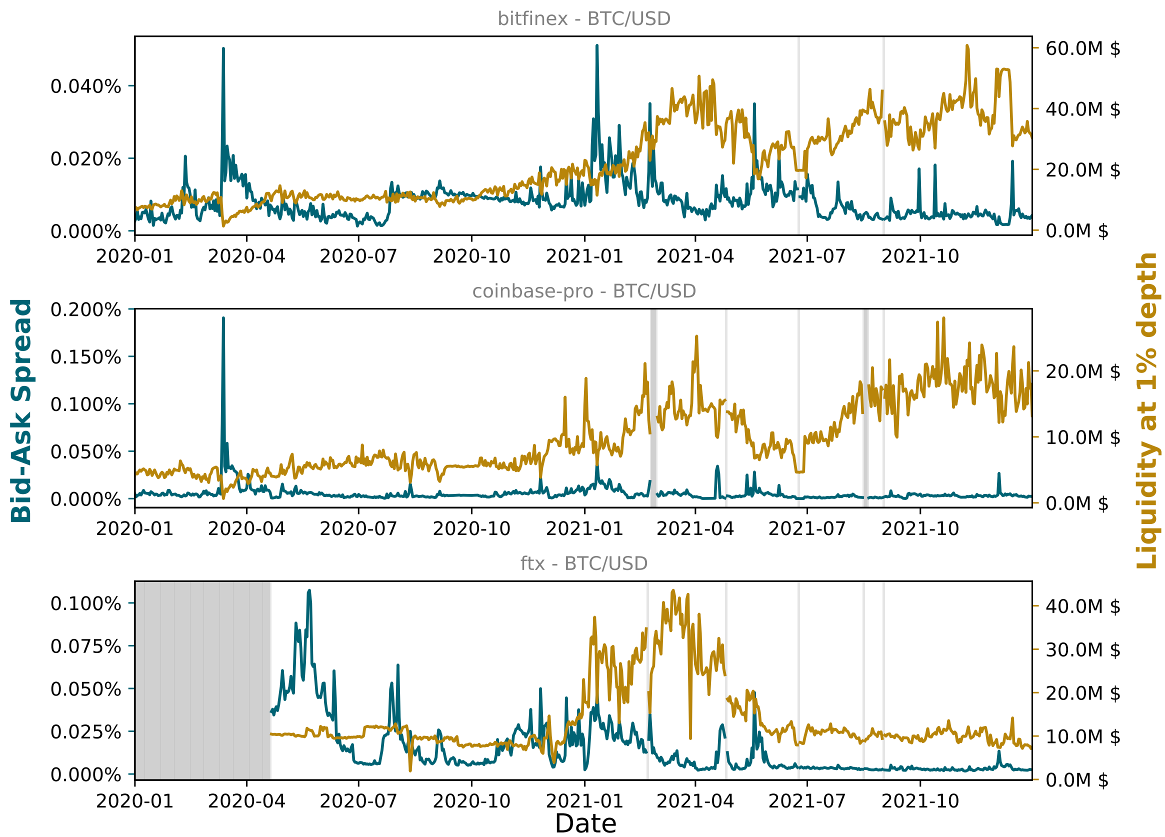

Figure A.4: Liquidity metrics of Bitcoin Spot markets

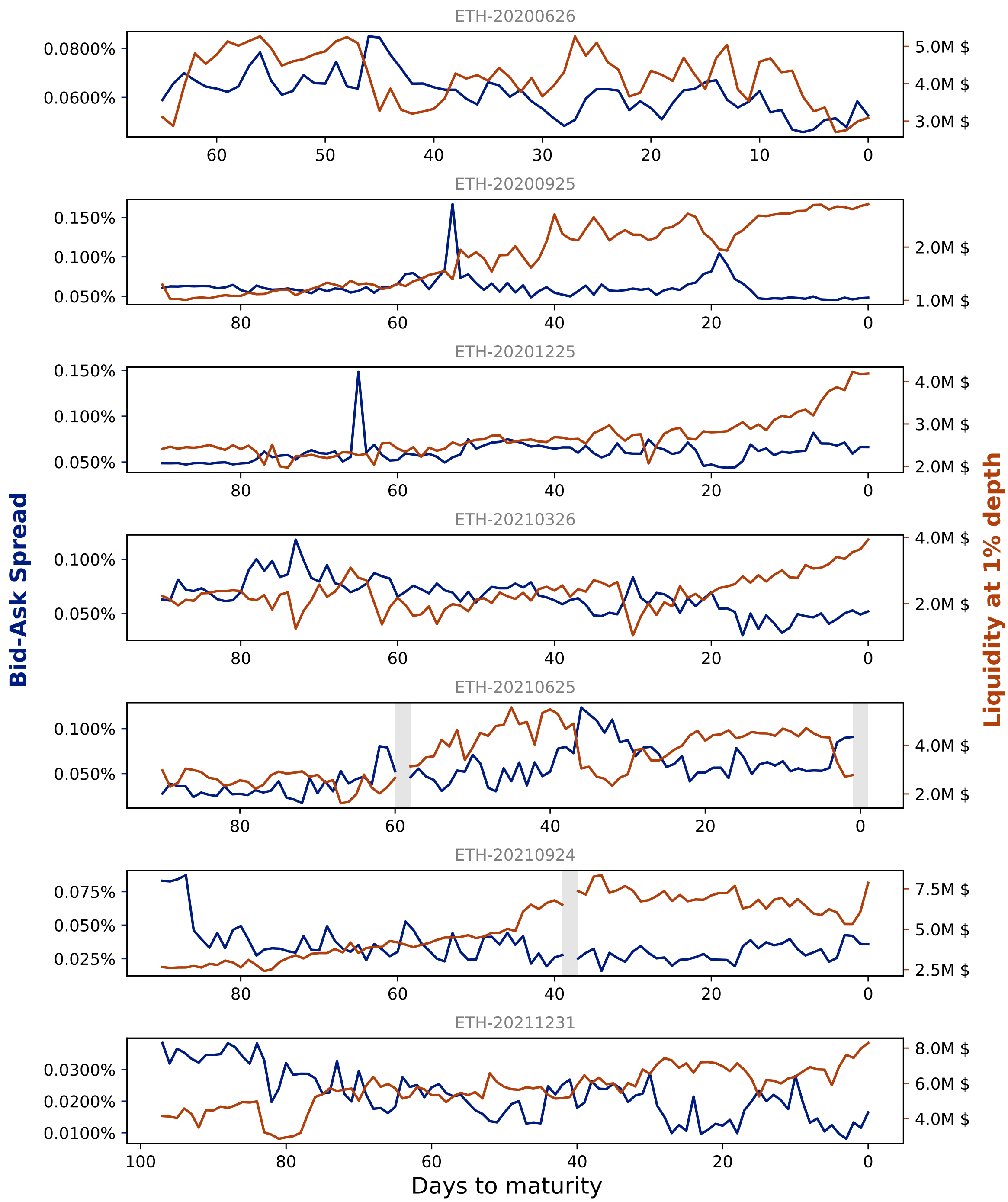

Figure A.5: Liquidity metrics of FTX's Ethereum Futures contracts

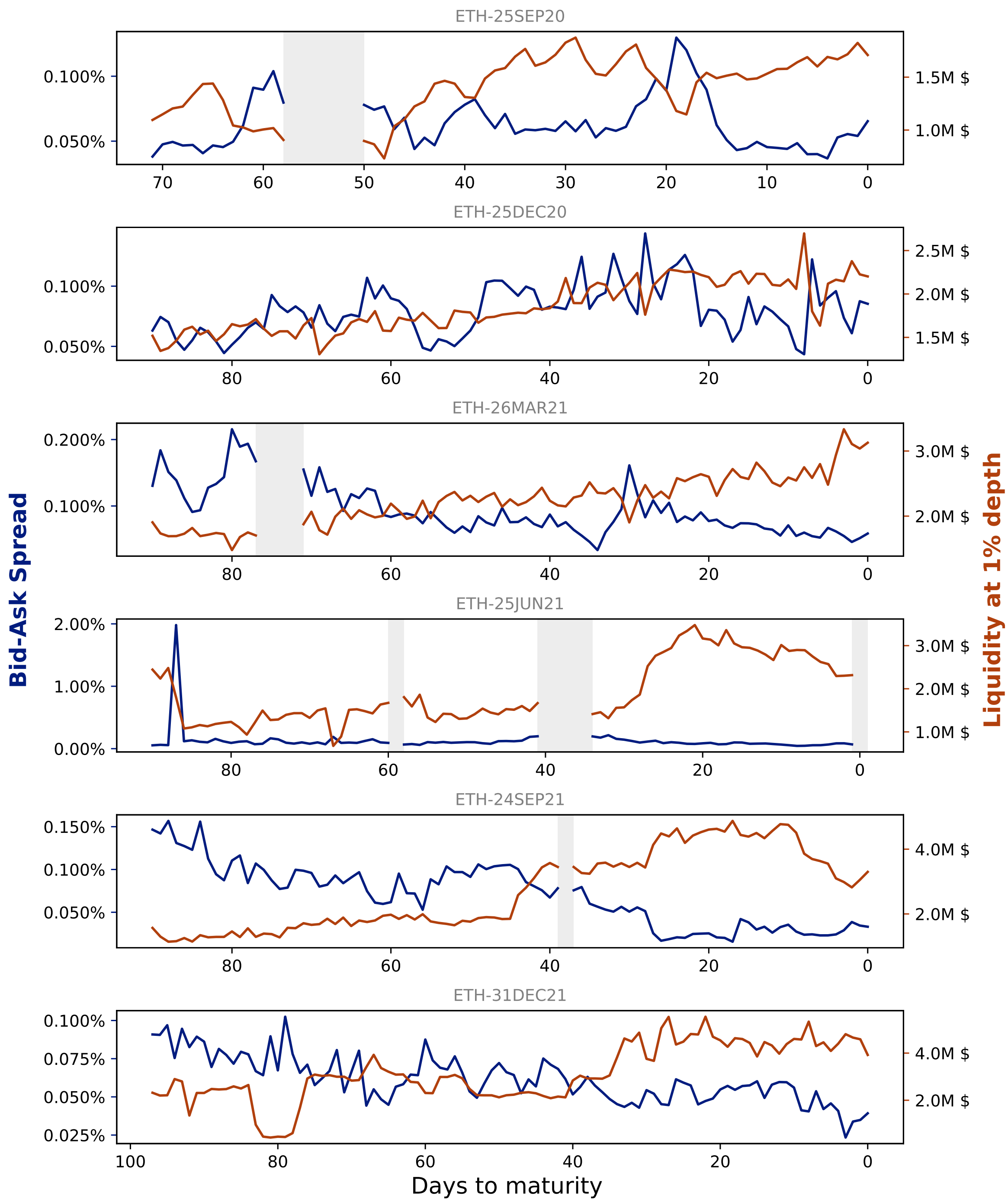

Figure A.6: Liquidity metrics of Deribit's Ethereum Futures contracts

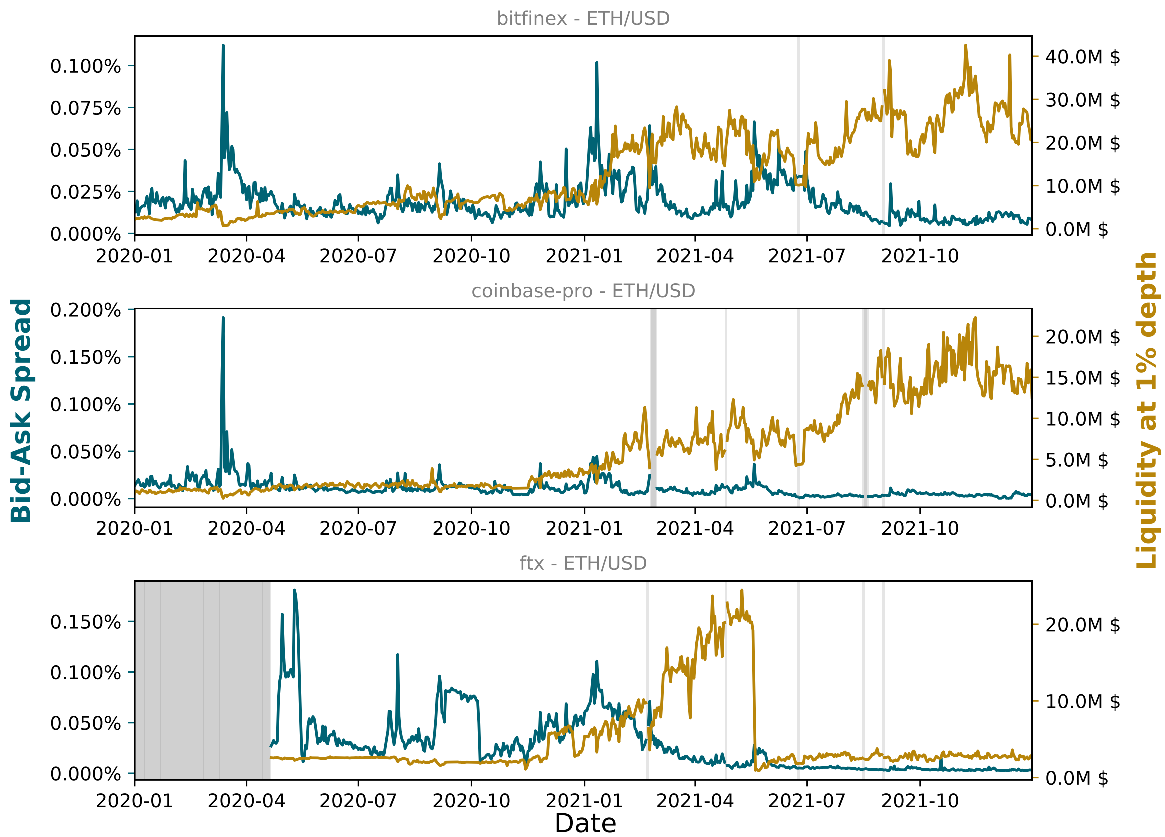

Figure A.7: Liquidity metrics of Ethereum Spot markets

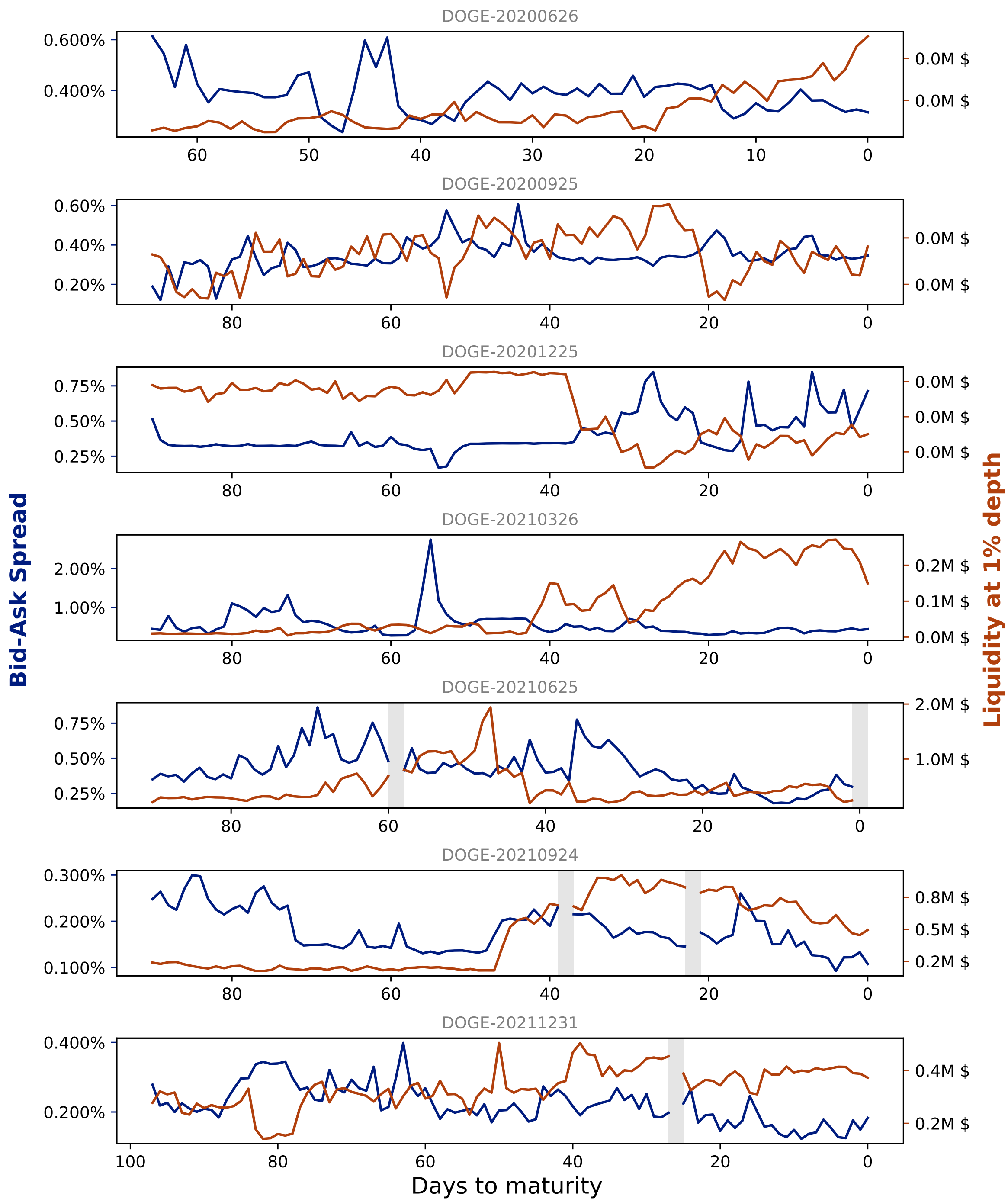

Figure A.8: Liquidity metrics of FTX's Dogecoin Futures contracts

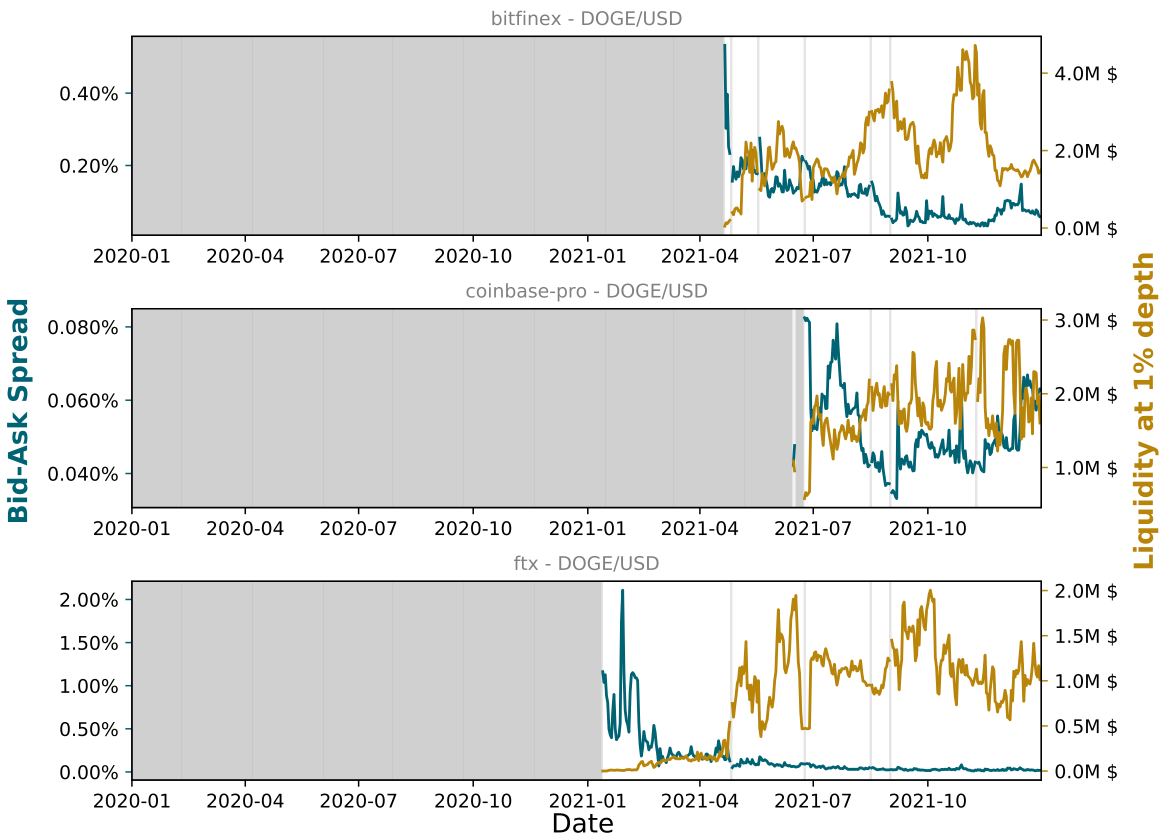

Figure A.9: Liquidity metrics of Dogecoin Spot markets

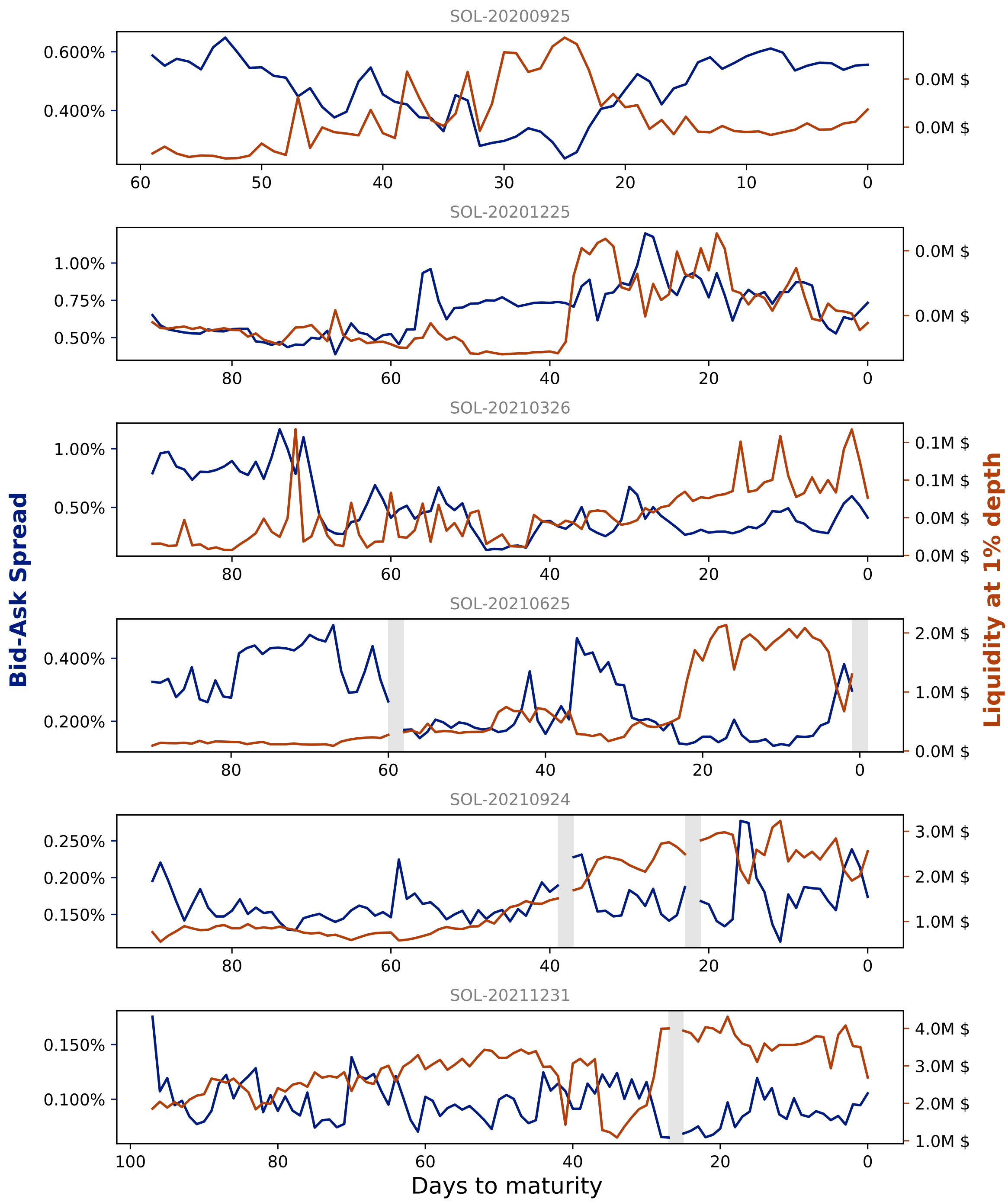

Figure A.10: Liquidity metrics of FTX's Solana Futures contracts

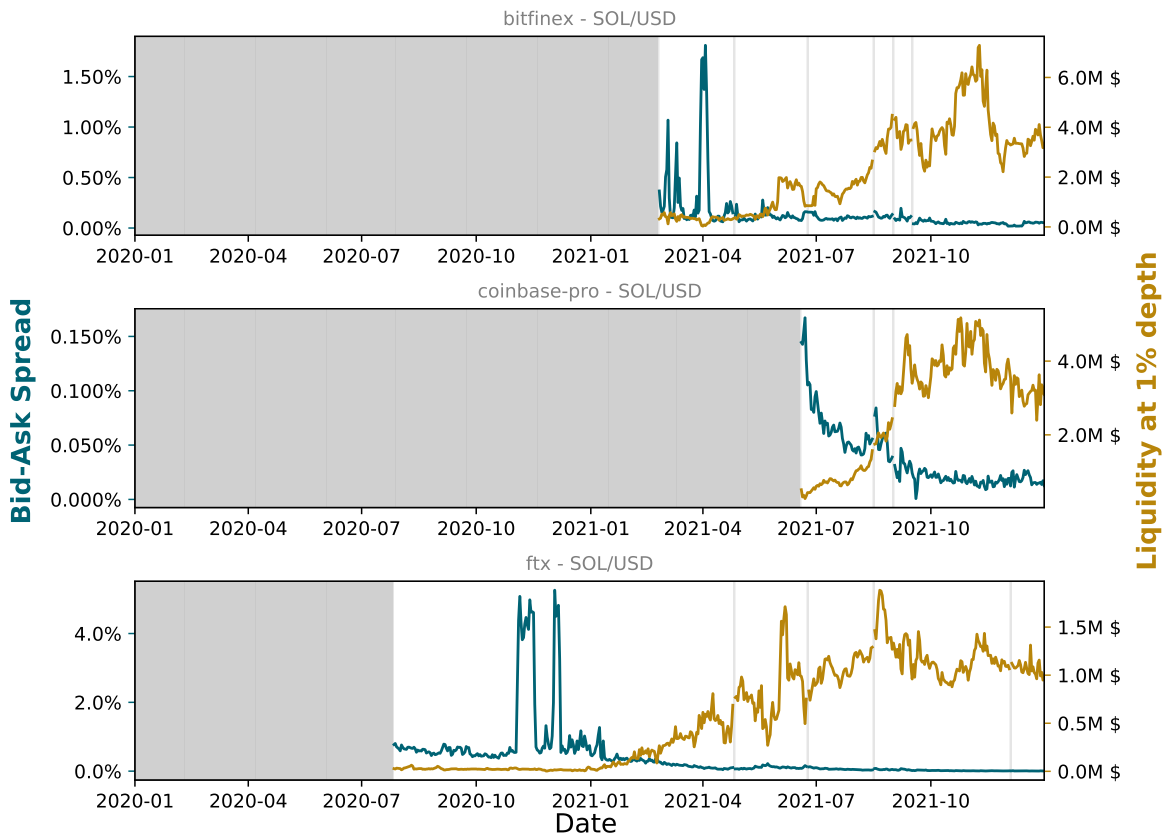

Figure A.11: Liquidity metrics of Solana Spot markets

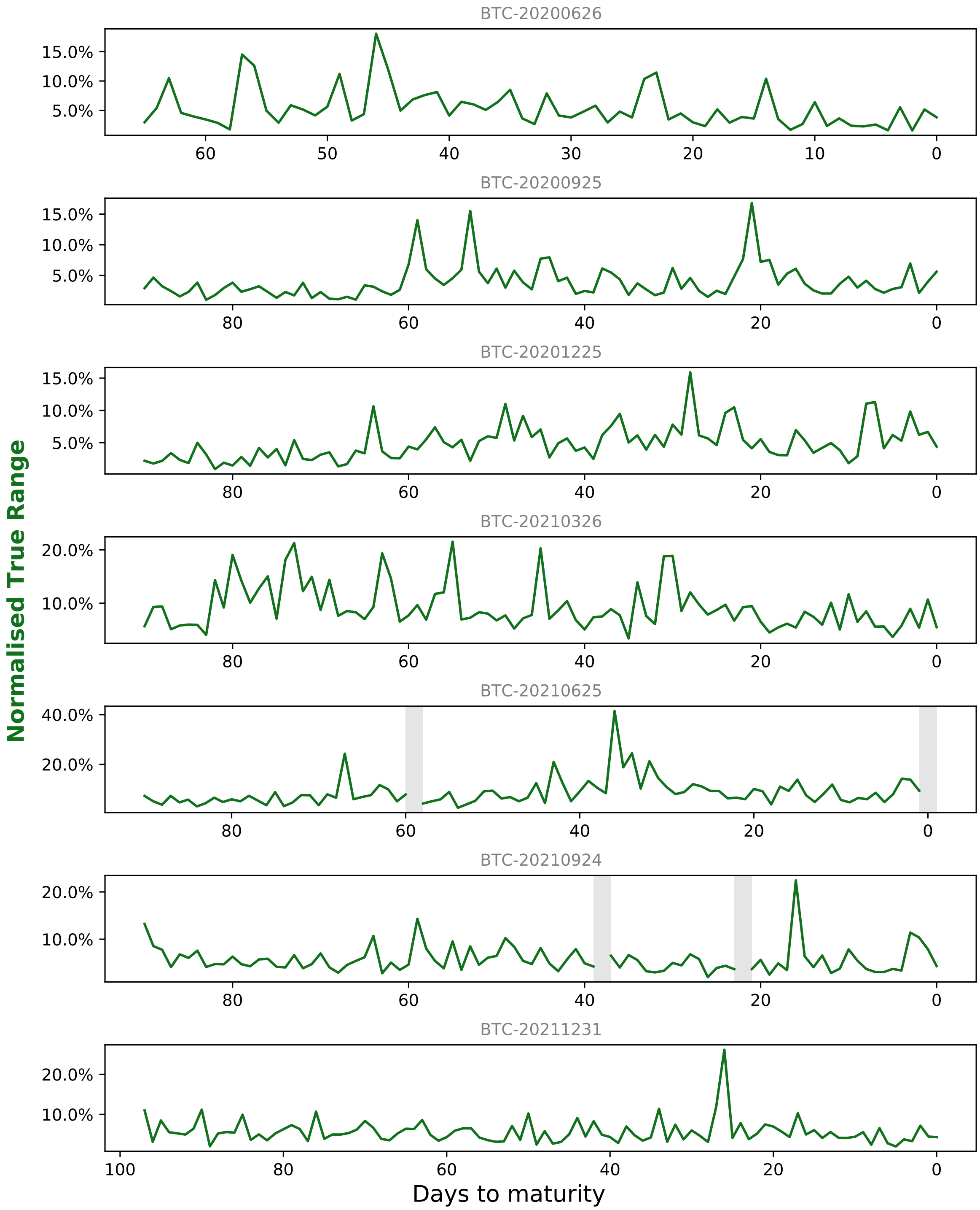

Figure A.12: Normalised True Range of FTX's Bitcoin Futures contracts

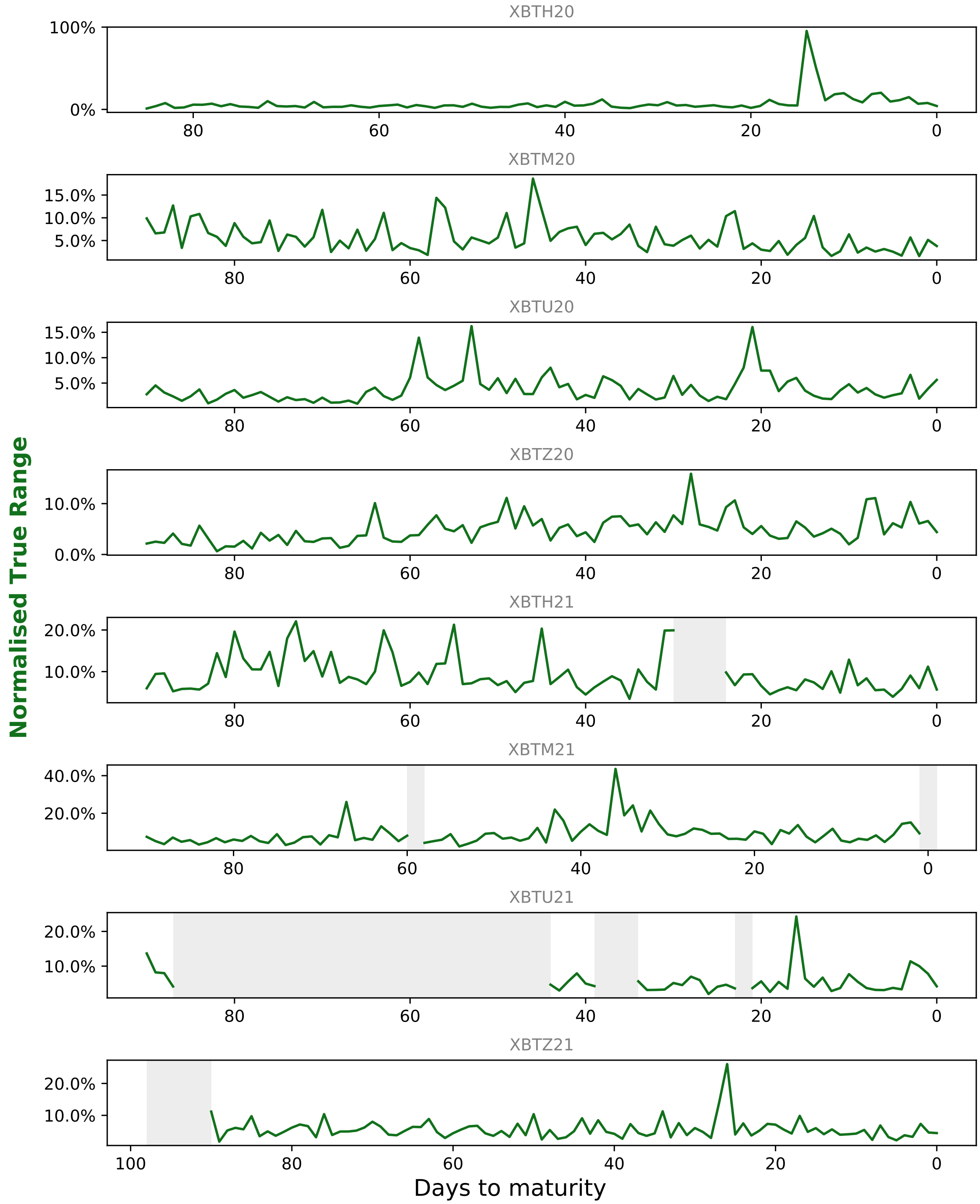

Figure A.13: Normalised True Range of Bitmex's Bitcoin Futures contracts

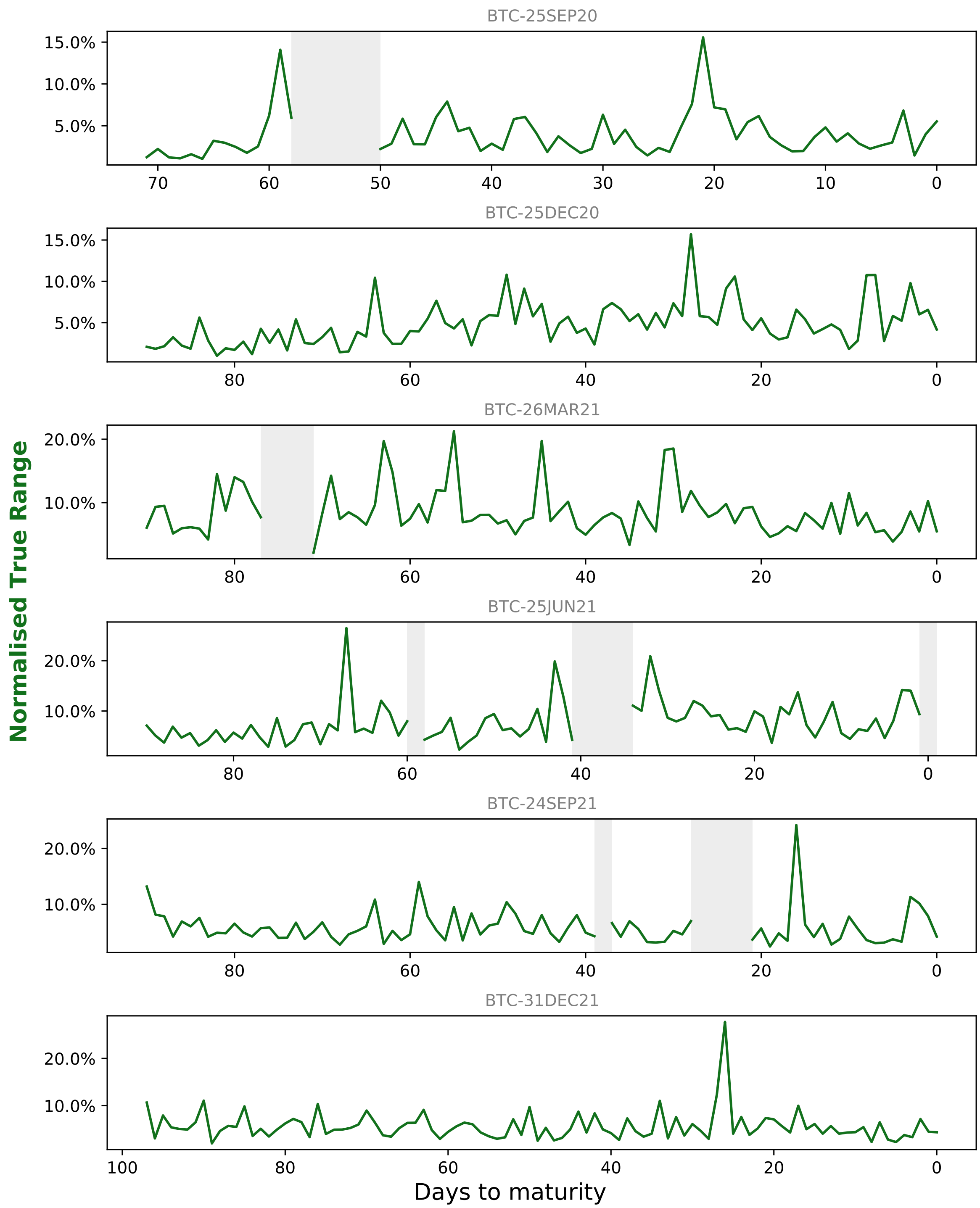

Figure A.14: Normalised True Range of Deribit's Bitcoin Futures contracts

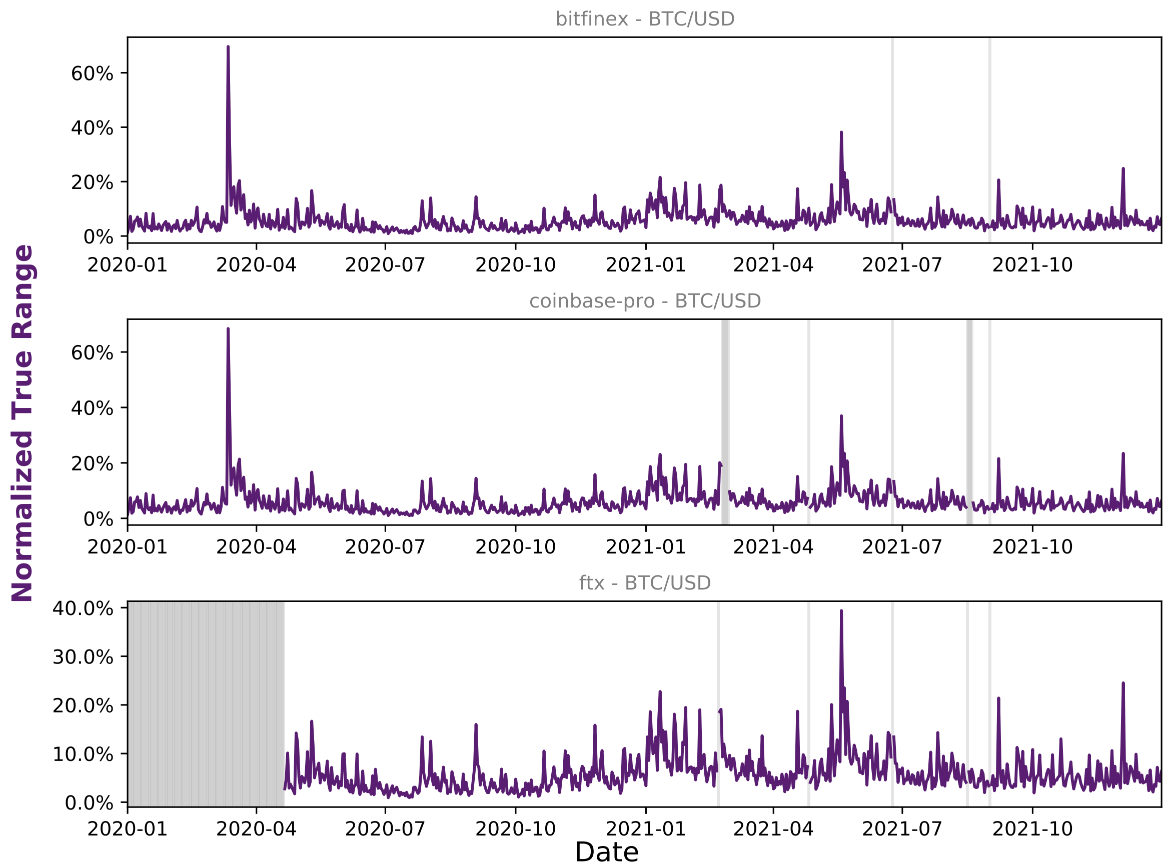

Figure A.15: Normalised True Range of Bitcoin Spot markets

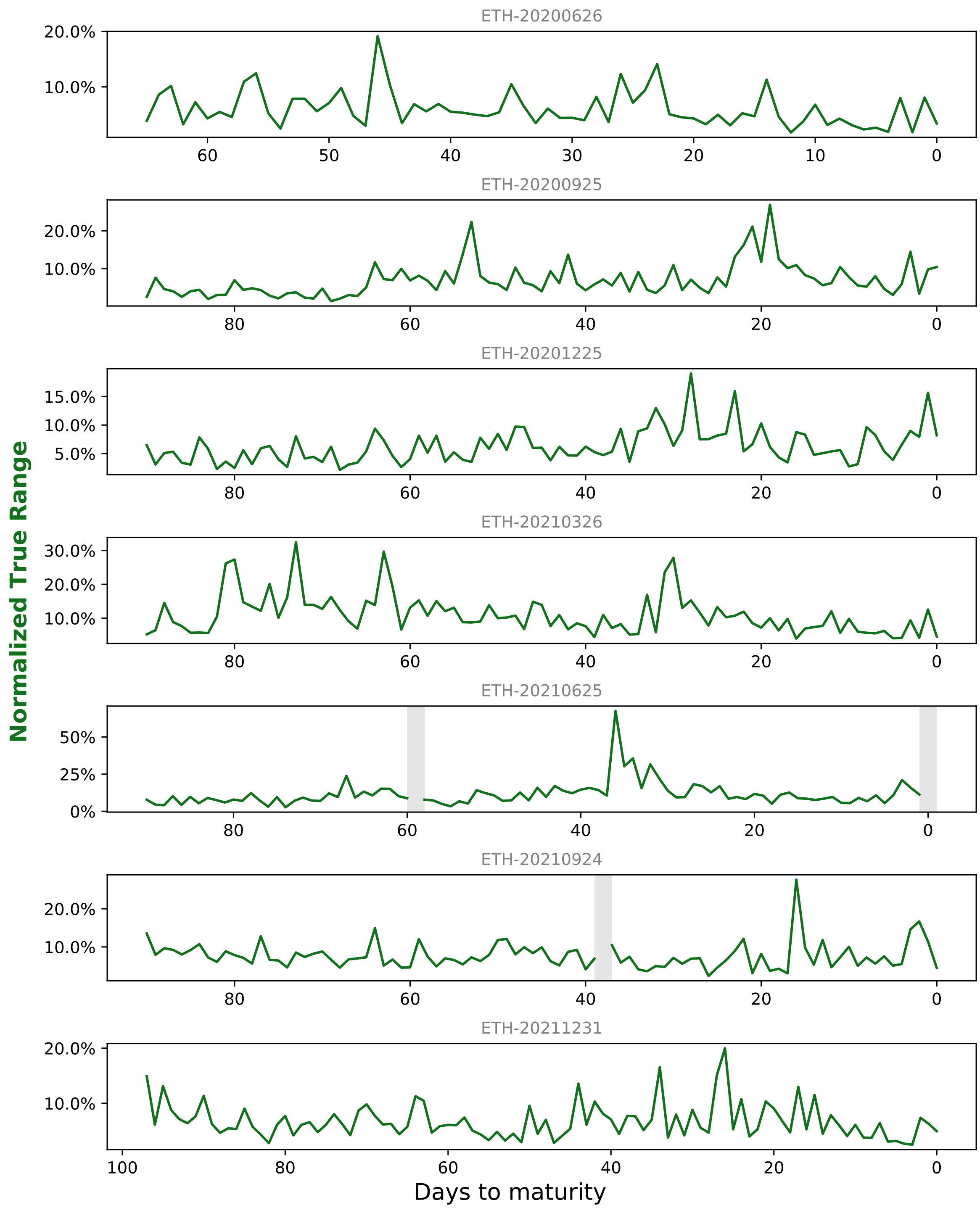

Figure A.16: Normalised True Range of FTX's Ethereum Futures contracts

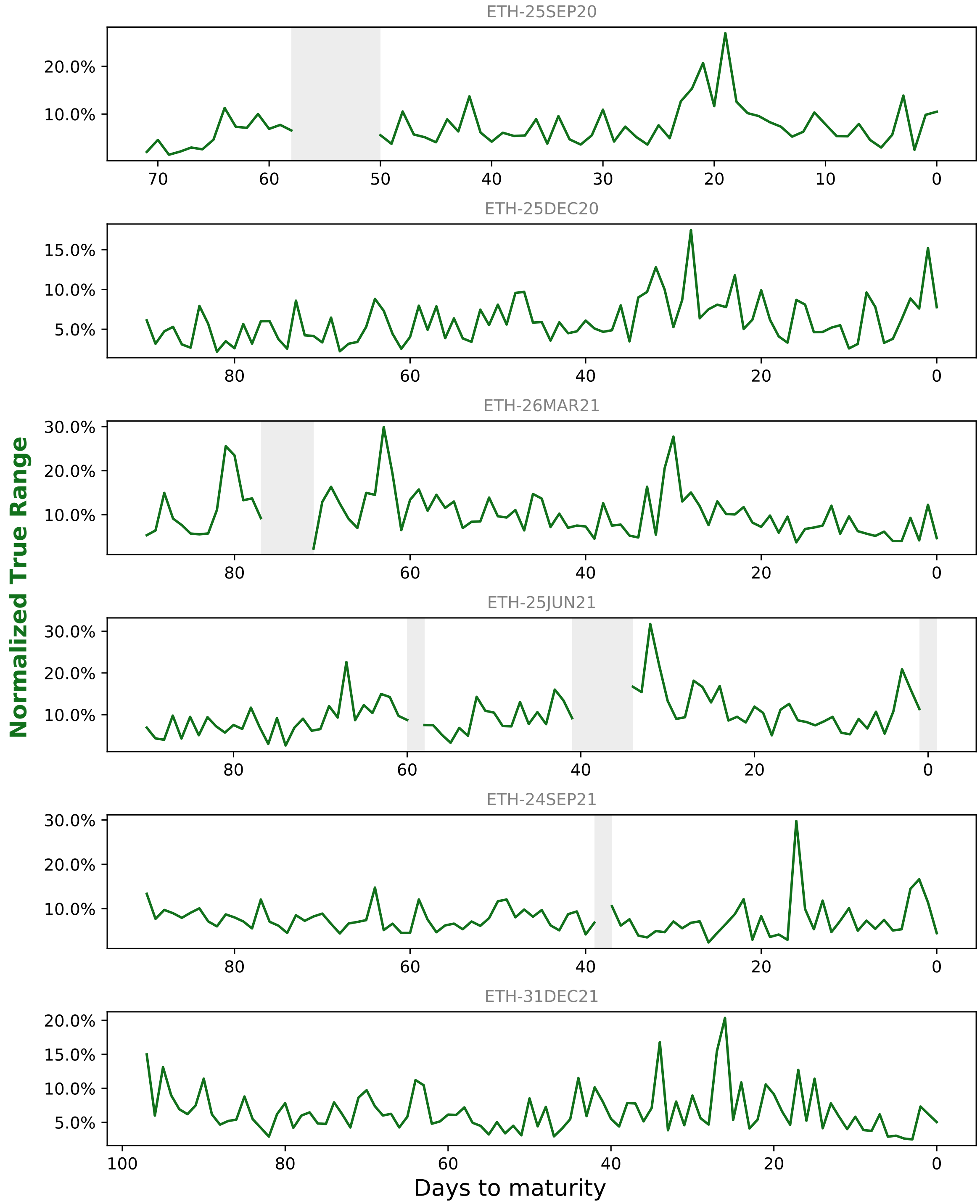

Figure A.17: Normalised True Range of Deribit's Ethereum Futures contracts

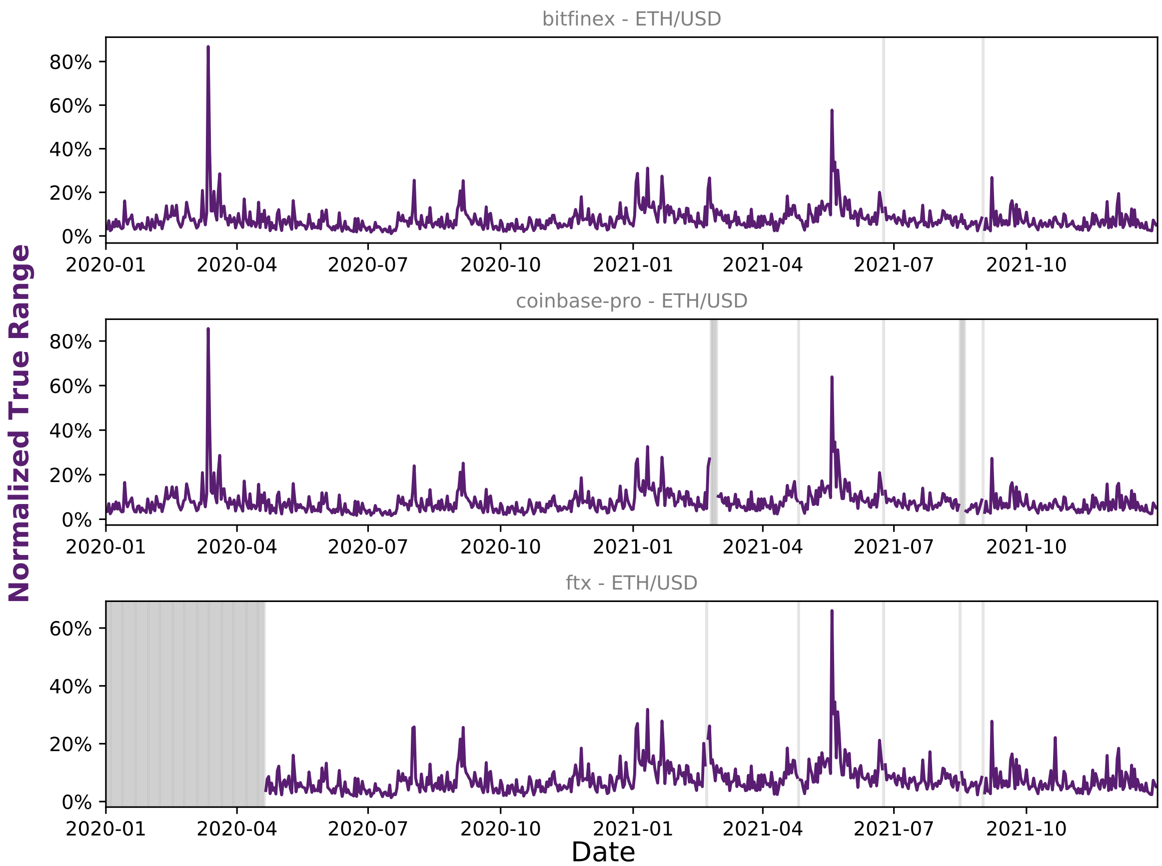

Figure A.18: Normalised True Range of Ethereum Spot markets

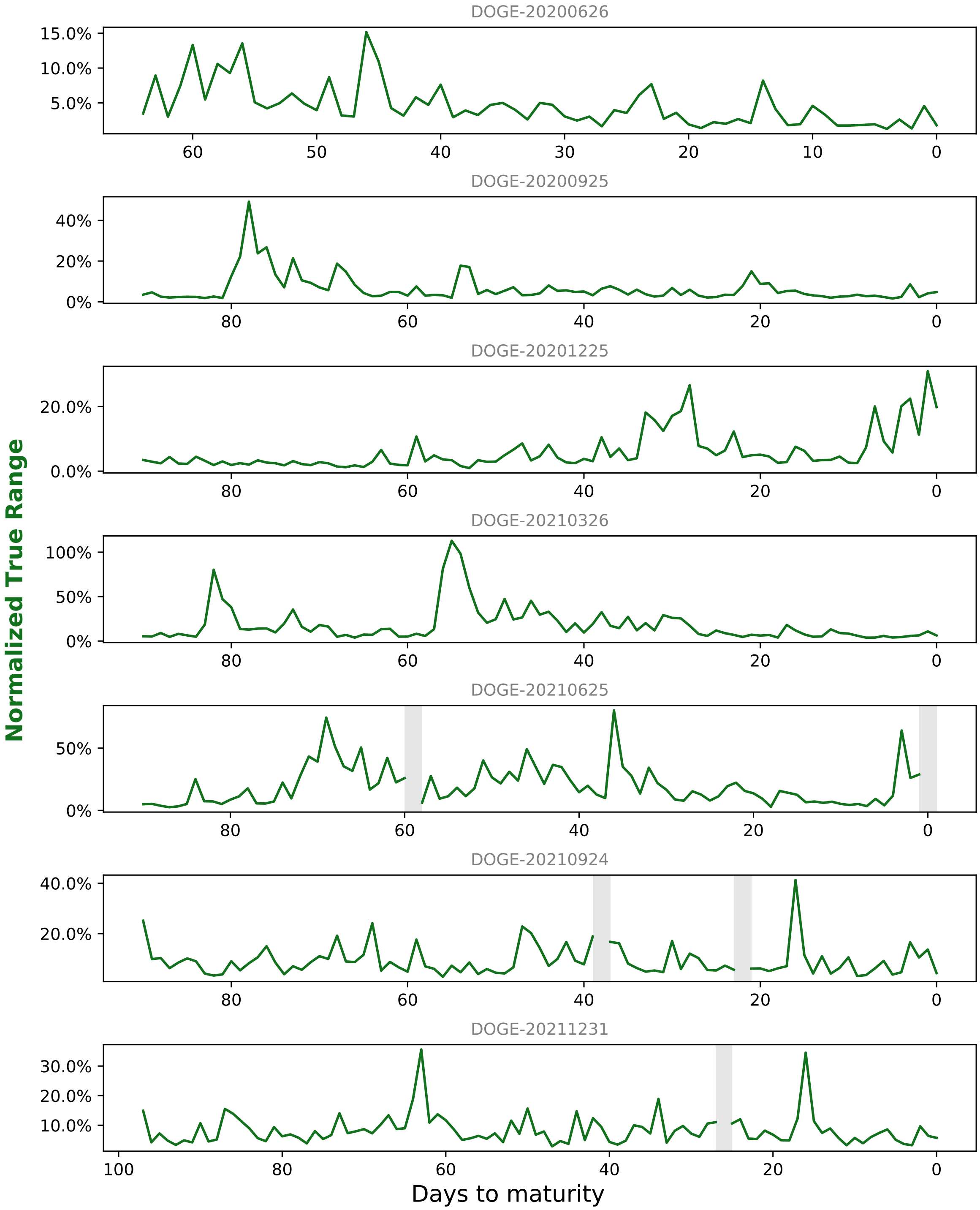

Figure A.19: Normalised True Range of Ftx's Dogecoin Futures contracts

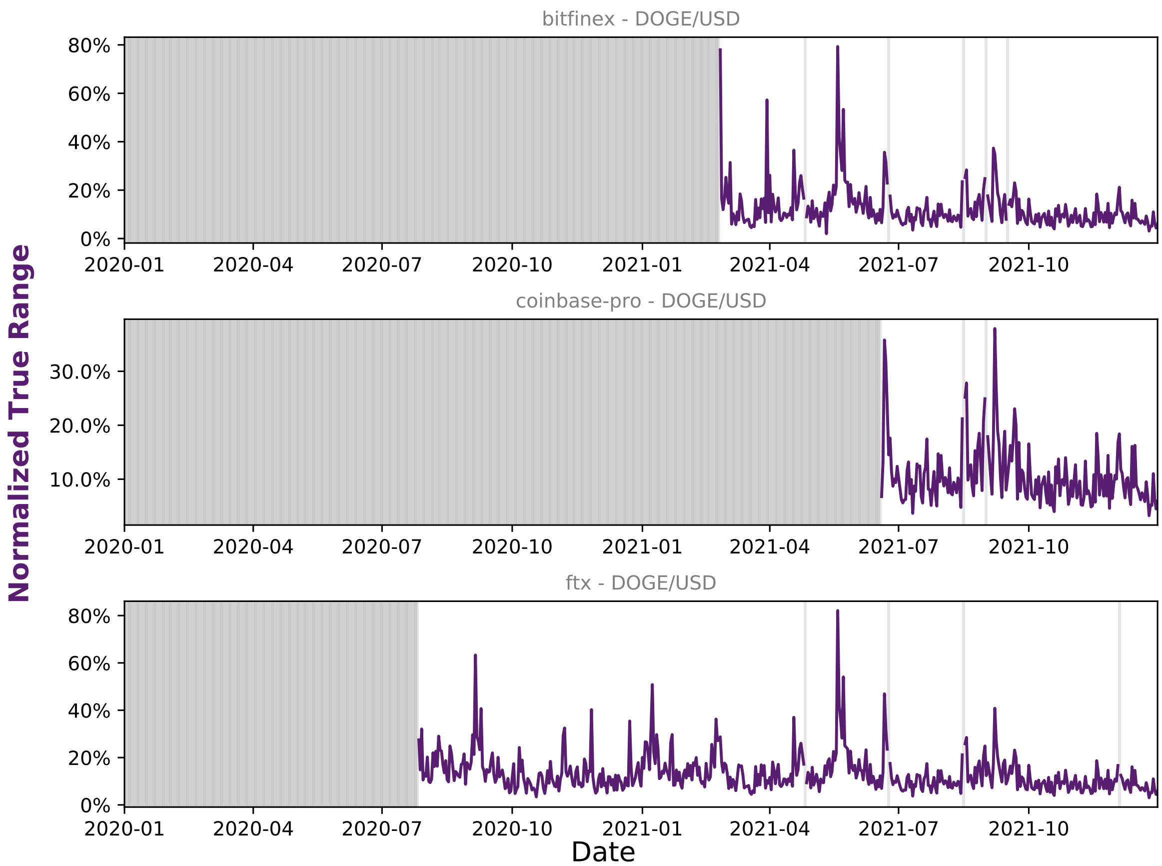

Figure A.20: Normalised True Range of Dogecoin Spot markets

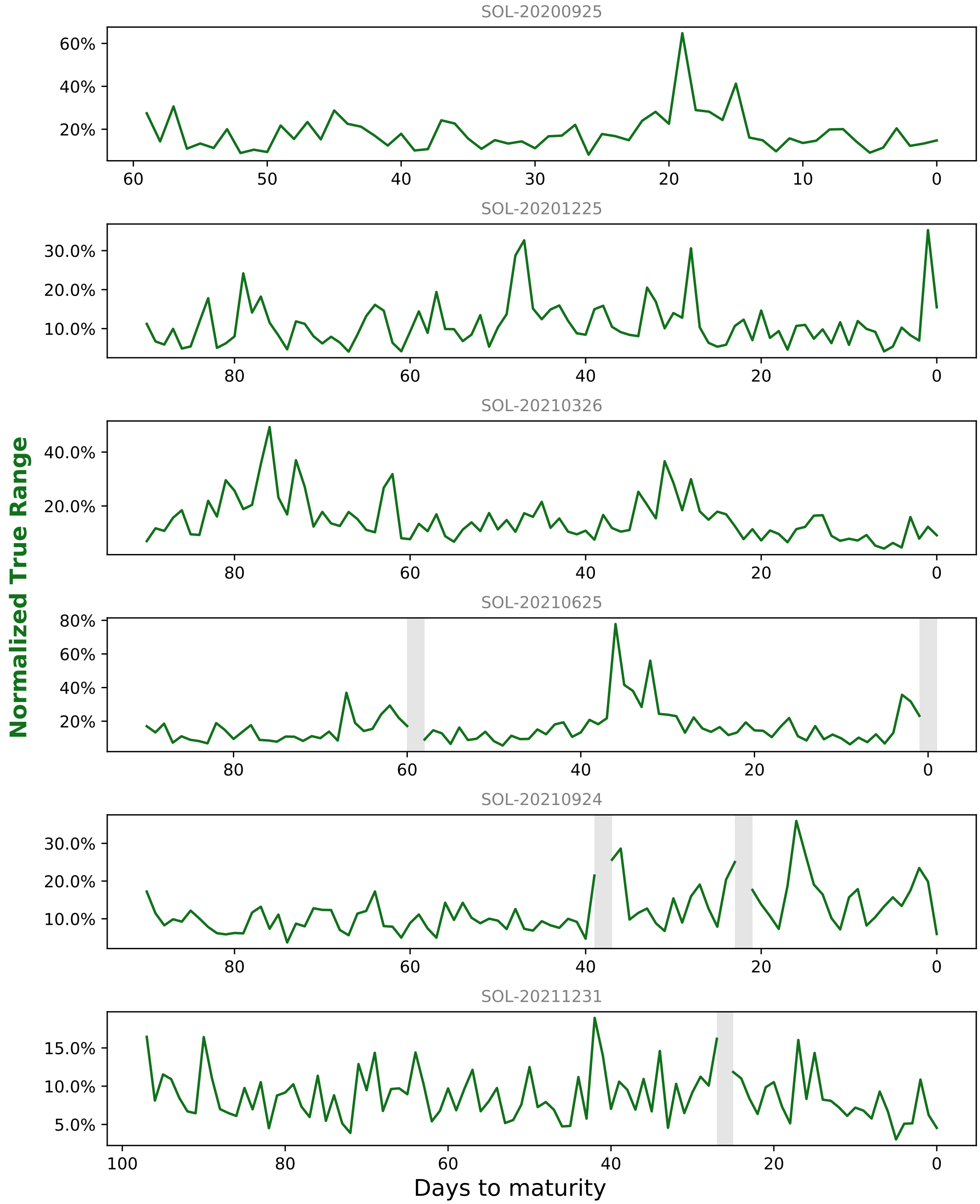

Figure A.21: Normalised True Range of Ftx's Solana Futures contracts

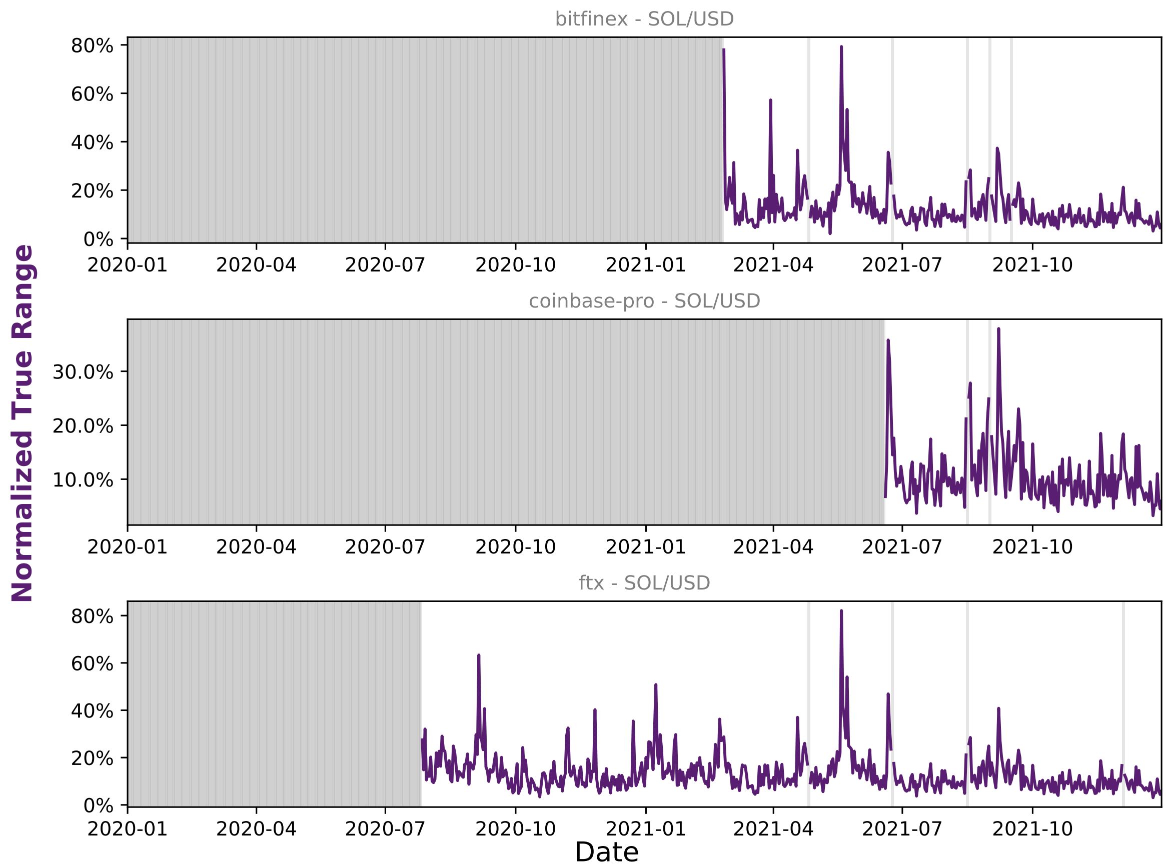

Figure A.22: Normalised True Range of Solana Spot markets

A.2 OLS values

A.2.1 Spread

Asset	Exchange	Contract	Beta	Beta P-Value	C	C P-Value
BTC	bitmex	XBTH20	-5.12E-06	0.006	0.000377298	0
BTC	bitmex	XBTM20	3.43E-07	0	5.29E-05	0
BTC	bitmex	XBTU20	6.86E-08	0.125	4.90E-05	0
BTC	bitmex	XBTZ20	-3.83E-06	0	0.000320837	0
BTC	bitmex	XBTH21	-6.99E-07	0.274	0.000458575	0
BTC	bitmex	XBTM21	-8.54E-06	0	0.00080636	0
BTC	bitmex	XBTU21	7.57E-06	0	0.000108557	0.038
BTC	bitmex	XBTZ21	-1.32E-06	0	0.000353796	0
BTC	deribit	BTC-25SEP20	-1.85E-06	0.004	0.000208929	0
BTC	deribit	BTC-25DEC20	-1.57E-06	0	0.00031049	0
BTC	deribit	BTC-26MAR21	1.95E-06	0	0.00037723	0
BTC	deribit	BTC-25JUN21	3.18E-06	0.253	0.000534919	0
BTC	deribit	BTC-24SEP21	8.50E-06	0	0.000239114	0
BTC	deribit	BTC-31DEC21	4.92E-06	0	0.00028398	0
BTC	ftx	BTC-20200626	6.19E-06	0	0.000230419	0
BTC	ftx	BTC-20200925	4.97E-07	0.462	0.000272088	0
BTC	ftx	BTC-20201225	-3.67E-06	0	0.000603875	0
BTC	ftx	BTC-20210326	5.26E-06	0	0.000288546	0
BTC	ftx	BTC-20210625	-3.41E-06	0	0.000477498	0
BTC	ftx	BTC-20210924	1.89E-06	0	7.65E-05	0
BTC	ftx	BTC-20211231	2.93E-07	0.007	8.97E-05	0
ETH	deribit	ETH-25SEP20	2.51E-08	0.983	0.000631765	0
ETH	deribit	ETH-25DEC20	-2.66E-06	0.001	0.000919801	0
ETH	deribit	ETH-26MAR21	1.00E-05	0	0.000488352	0
ETH	deribit	ETH-25JUN21	1.55E-05	0.07	0.000537105	0.234

ETH	deribit	ETH-24SEP21	1.19E-05	0	0.000163684	0
ETH	deribit	ETH-31DEC21	4.10E-06	0	0.000416965	0
ETH	ftx	ETH-20200626	2.72E-06	0	0.00052791	0
ETH	ftx	ETH-20200925	1.02E-06	0.092	0.000566849	0
ETH	ftx	ETH-20201225	-1.13E-06	0.023	0.000657569	0
ETH	ftx	ETH-20210326	4.37E-06	0	0.00046263	0
ETH	ftx	ETH-20210625	-5.22E-06	0	0.000778326	0
ETH	ftx	ETH-20210924	2.24E-06	0	0.00024477	0
ETH	ftx	ETH-20211231	2.05E-06	0	0.000120364	0
DOGE	ftx	DOGE-20200626	1.44E-05	0.005	0.003421548	0
DOGE	ftx	DOGE-20200925	-9.24E-06	0.001	0.003879536	0
DOGE	ftx	DOGE-20201225	-3.13E-05	0	0.005485205	0
DOGE	ftx	DOGE-20210326	3.98E-05	0.002	0.003806507	0
DOGE	ftx	DOGE-20210625	2.44E-05	0	0.003138594	0
DOGE	ftx	DOGE-20210924	8.74E-06	0	0.001409372	0
DOGE	ftx	DOGE-20211231	1.25E-05	0	0.001620796	0
SOL	ftx	SOL-20200925	-3.00E-06	0.709	0.004823965	0
SOL	ftx	SOL-20201225	-3.92E-05	0	0.008636062	0
SOL	ftx	SOL-20210326	5.82E-05	0	0.002309969	0
SOL	ftx	SOL-20210625	2.43E-05	0	0.001540975	0
SOL	ftx	SOL-20210924	-3.04E-06	0.01	0.001808724	0
SOL	ftx	SOL-20211231	1.86E-06	0.005	0.000870296	0
BTC	bitfinex	spotA	-1.68E-06	0	0.000152894	0
BTC	bitfinex	spotB	8.33E-07	0	2.53E-05	0
BTC	bitfinex	spotC	-1.10E-06	0	0.000127114	0
BTC	bitfinex	spotD	1.16E-07	0.098	8.30E-05	0
BTC	bitfinex	spotE	1.03E-06	0	9.24E-05	0
BTC	bitfinex	spotG	5.92E-07	0	3.09E-05	0
BTC	bitfinex	spotH	1.58E-07	0.101	3.83E-05	0

BTC	bitfinex	spotF	-8.33E-07	0	0.000134731	0
BTC	coinbase-pro	spotA	-3.69E-06	0	0.000273757	0
BTC	coinbase-pro	spotB	9.59E-07	0	1.77E-05	0.017
BTC	coinbase-pro	spotC	-1.29E-07	0.113	3.51E-05	0
BTC	coinbase-pro	spotD	-3.87E-07	0.001	6.71E-05	0
BTC	coinbase-pro	spotE	1.13E-06	0	3.26E-05	0.005
BTC	coinbase-pro	spotF	-7.71E-08	0.772	4.78E-05	0.001
BTC	coinbase-pro	spotG	-1.55E-07	0.004	2.54E-05	0
BTC	coinbase-pro	spotH	1.01E-08	0.914	3.32E-05	0
BTC	ftx	spotB	5.55E-06	0	0.00029517	0
BTC	ftx	spotC	3.65E-07	0.414	0.00012344	0
BTC	ftx	spotD	-1.60E-06	0	0.000228625	0
BTC	ftx	spotE	2.39E-06	0	7.77E-05	0
BTC	ftx	spotG	1.02E-07	0	2.77E-05	0
BTC	ftx	spotH	-1.07E-07	0.04	3.63E-05	0
BTC	ftx	spotF	2.44E-10	0.999	7.76E-05	0
DOGE	bitfinex	spotG	1.50E-05	0	0.000479778	0
DOGE	bitfinex	spotH	-3.28E-06	0	0.000795911	0
DOGE	bitfinex	spotF	1.97E-05	0	0.001196002	0
DOGE	coinbase-pro	spotG	3.64E-06	0	0.000362177	0
DOGE	coinbase-pro	spotH	-1.30E-06	0	0.000560291	0
DOGE	ftx	spotE	0.000127877	0	0.000423875	0.562
DOGE	ftx	spotG	5.09E-06	0	0.000190414	0
DOGE	ftx	spotH	1.72E-07	0.658	0.000251058	0
DOGE	ftx	spotF	1.62E-05	0	0.000543804	0
ETH	bitfinex	spotB	3.90E-07	0.109	0.000135534	0
ETH	bitfinex	spotC	-8.51E-07	0	0.000209142	0
ETH	bitfinex	spotD	-1.10E-06	0	0.000207314	0
ETH	bitfinex	spotE	2.78E-06	0	0.000160865	0

ETH	bitfinex	spotF	-3.35E-06	0	0.000394247	0
ETH	bitfinex	spotG	2.20E-06	0	5.28E-05	0
ETH	bitfinex	spotH	-3.76E-08	0.689	8.92E-05	0
ETH	coinbase-pro	spotB	9.15E-07	0	8.30E-05	0
ETH	coinbase-pro	spotC	-7.96E-07	0	0.000148261	0
ETH	coinbase-pro	spotD	-1.84E-07	0.315	0.00010531	0
ETH	coinbase-pro	spotE	1.65E-06	0	5.72E-05	0.001
ETH	coinbase-pro	spotG	-3.92E-07	0	5.26E-05	0
ETH	coinbase-pro	spotH	3.21E-07	0	2.52E-05	0
ETH	coinbase-pro	spotF	-4.05E-08	0.851	9.28E-05	0
ETH	ftx	spotB	1.04E-05	0	0.00022813	0.011
ETH	ftx	spotC	-6.43E-06	0	0.00071805	0
ETH	ftx	spotD	3.23E-07	0.677	0.000360937	0
ETH	ftx	spotE	6.81E-06	0	0.000168066	0
ETH	ftx	spotG	2.36E-07	0	3.69E-05	0
ETH	ftx	spotH	1.63E-07	0.001	2.84E-05	0
ETH	ftx	spotF	5.26E-07	0.011	8.54E-05	0
SOL	bitfinex	spotE	1.00E-04	0.058	0.001282484	0.13
SOL	bitfinex	spotG	2.25E-06	0.038	0.000921591	0
SOL	bitfinex	spotH	2.36E-06	0	0.000374075	0
SOL	bitfinex	spotF	5.73E-05	0	-0.000413817	0.551
SOL	coinbase-pro	spotG	6.74E-06	0	0.000214115	0
SOL	coinbase-pro	spotH	4.17E-07	0.006	0.000161299	0
SOL	ftx	spotC	1.38E-05	0.033	0.005799217	0
SOL	ftx	spotD	-0.000133353	0.032	0.020076788	0
SOL	ftx	spotE	6.27E-05	0	0.00076841	0.01
SOL	ftx	spotF	-3.44E-06	0.01	0.001161889	0
SOL	ftx	spotG	6.18E-06	0	0.000246661	0
SOL	ftx	spotH	1.67E-06	0	2.80E-05	0

Table A.1: OLS Spread values

A.2.2 Depth

Asset	Exchange	Contract	Beta	Beta P-Value	C	C P-Value
BTC	bitmex	XBTH20	36923.45963	0.001	5670640.966	0
BTC	bitmex	XBTM20	-13460.89775	0	6044001.937	0
BTC	bitmex	XBTU20	11171.97723	0.036	7845792.014	0
BTC	bitmex	XBTZ20	-3816.662207	0.177	7975185.733	0
BTC	bitmex	XBTH21	-50070.40869	0	12726943.12	0
BTC	bitmex	XBTM21	137941.4971	0	5883526.267	0
BTC	bitmex	XBTU21	38144.55565	0	-569095.8033	0
BTC	bitmex	XBTZ21	-253.8421771	0.014	119221.0518	0
BTC	deribit	BTC-25SEP20	14461.95637	0	2747069.216	0
BTC	deribit	BTC-25DEC20	-44008.5039	0	6058811.137	0
BTC	deribit	BTC-26MAR21	-34725.47123	0	8348725.12	0
BTC	deribit	BTC-25JUN21	-35674.37169	0	5512068.922	0
BTC	deribit	BTC-24SEP21	-73692.26633	0	7694000.969	0
BTC	deribit	BTC-31DEC21	-63227.34914	0	7868946.525	0
BTC	ftx	BTC-20200626	47322.70953	0	2363791.94	0
BTC	ftx	BTC-20200925	-25135.62223	0	4419363	0
BTC	ftx	BTC-20201225	-18778.54849	0	5378653.803	0
BTC	ftx	BTC-20210326	-37161.34894	0	6627981.472	0
BTC	ftx	BTC-20210625	4584.579177	0.175	5307497.286	0
BTC	ftx	BTC-20210924	-60450.09524	0	10424586.79	0
BTC	ftx	BTC-20211231	-35251.86275	0	8552487.588	0
ETH	deribit	ETH-25SEP20	-8722.13956	0	1713397.246	0
ETH	deribit	ETH-25DEC20	-9429.060997	0	2277083.899	0
ETH	deribit	ETH-26MAR21	-13242.62672	0	2830397.029	0

ETH	deribit	ETH-25JUN21	-20061.99182	0	2840007.215	0
ETH	deribit	ETH-24SEP21	-40304.35336	0	4522690.795	0
ETH	deribit	ETH-31DEC21	-31743.74284	0	4715215.865	0
ETH	ftx	ETH-20200626	12187.33452	0.006	3777949.204	0
ETH	ftx	ETH-20200925	-21826.58003	0	2834575.112	0
ETH	ftx	ETH-20201225	-13893.3999	0	3312412.886	0
ETH	ftx	ETH-20210326	-10187.0223	0	2809669.849	0
ETH	ftx	ETH-20210625	-22522.00488	0	4550864.85	0
ETH	ftx	ETH-20210924	-57102.94541	0	7745462.217	0
ETH	ftx	ETH-20211231	-32911.09073	0	7286008.819	0
DOGE	ftx	DOGE-20200626	-416.8459353	0	28720.94685	0
DOGE	ftx	DOGE-20200925	-45.9960981	0.022	28690.48376	0
DOGE	ftx	DOGE-20201225	115.3331741	0	5970.347826	0
DOGE	ftx	DOGE-20210326	-2968.313681	0	221560.5442	0
DOGE	ftx	DOGE-20210625	111.4147499	0.932	492778.5626	0
DOGE	ftx	DOGE-20210924	-7847.49202	0	807329.0933	0
DOGE	ftx	DOGE-20211231	-1821.281012	0	425034.4056	0
SOL	ftx	SOL-20200925	-108.7609892	0.026	15008.81585	0
SOL	ftx	SOL-20201225	-99.4450868	0	14385.06187	0
SOL	ftx	SOL-20210326	-1030.324431	0	99662.11586	0
SOL	ftx	SOL-20210625	-20037.02007	0	1558760.124	0
SOL	ftx	SOL-20210924	-27575.54785	0	2781227.242	0
SOL	ftx	SOL-20211231	-14721.14404	0	3614435.545	0
BTC	bitfinex	spotA	19131.72546	0.078	7533285.156	0
BTC	bitfinex	spotB	-35912.06275	0	11938582.49	0
BTC	bitfinex	spotC	12100.13235	0.001	9805370.748	0
BTC	bitfinex	spotD	-103145.5891	0	19281220.96	0
BTC	bitfinex	spotE	-311943.964	0	40885842.12	0
BTC	bitfinex	spotG	-146701.8496	0	39543177.04	0

BTC	bitfinex	spotH	-48079.32725	0.084	41330396.35	0
BTC	bitfinex	spotF	220759.1456	0	21894018.88	0
BTC	coinbase-pro	spotA	22038.09153	0	3257829.389	0
BTC	coinbase-pro	spotB	-28548.73626	0	6347603.242	0
BTC	coinbase-pro	spotC	17915.11961	0	5149713.024	0
BTC	coinbase-pro	spotD	-47811.87211	0	8977546.003	0
BTC	coinbase-pro	spotE	-73362.3767	0	15052338.28	0
BTC	coinbase-pro	spotF	124600.6817	0	5873563.129	0
BTC	coinbase-pro	spotG	-109173.6328	0	18118861.47	0
BTC	coinbase-pro	spotH	21854.61367	0.034	17230923.69	0
BTC	ftx	spotB	9155.254566	0.043	10018876.8	0
BTC	ftx	spotC	35196.14477	0	8609830.079	0
BTC	ftx	spotD	-30955.04762	0	10000267.58	0
BTC	ftx	spotE	-209898.1448	0	37576002.91	0
BTC	ftx	spotG	15110.90789	0	9654661.423	0
BTC	ftx	spotH	21703.23171	0	8716705.609	0
BTC	ftx	spotF	253635.9381	0	7227448.356	0
DOGE	bitfinex	spotG	-22829.35378	0	3128695.972	0
DOGE	bitfinex	spotH	10791.69596	0.002	1725416.827	0
DOGE	bitfinex	spotF	-24712.83624	0	2256695.297	0
DOGE	coinbase-pro	spotG	-8695.347464	0	2071210.551	0
DOGE	coinbase-pro	spotH	-1802.112113	0.253	2095685.167	0
DOGE	ftx	spotE	-2496.608513	0	159153.3187	0
DOGE	ftx	spotG	-4222.082791	0	1349189.113	0
DOGE	ftx	spotH	6568.826158	0	809298.3191	0
DOGE	ftx	spotF	-15540.51513	0	1442117.596	0
ETH	bitfinex	spotB	989.7421981	0.711	3851042.166	0
ETH	bitfinex	spotC	-13042.66603	0.029	6860984.806	0
ETH	bitfinex	spotD	-18380.28396	0	7552907.394	0

ETH	bitfinex	spotE	-206545.9782	0	26127731.77	0
ETH	bitfinex	spotF	46967.084	0.003	16769495.46	0
ETH	bitfinex	spotG	-156546.3442	0	29160926.11	0
ETH	bitfinex	spotH	-36027.5711	0.065	27923854.55	0
ETH	coinbase-pro	spotB	-5988.330446	0	1984575.739	0
ETH	coinbase-pro	spotC	895.1140309	0.502	1804613.935	0
ETH	coinbase-pro	spotD	-20626.31215	0	3129460.201	0
ETH	coinbase-pro	spotE	-49188.82615	0	7975012.925	0
ETH	coinbase-pro	spotG	-98572.03373	0	16152400.89	0
ETH	coinbase-pro	spotH	-1813.449879	0.836	15755344.16	0
ETH	coinbase-pro	spotF	2635.616445	0.701	7593463.955	0
ETH	ftx	spotB	-1252.768396	0.006	2573084.518	0
ETH	ftx	spotC	10131.7756	0	1818676.801	0
ETH	ftx	spotD	-47164.00908	0	5124875.573	0
ETH	ftx	spotE	-117793.8778	0	13782502.28	0
ETH	ftx	spotG	-953.0967466	0.516	2734700.847	0
ETH	ftx	spotH	393.186823	0.711	2737803.092	0
ETH	ftx	spotF	234061.8059	0	1494199.678	0.25
SOL	bitfinex	spotE	2795.121182	0.213	358052.4069	0
SOL	bitfinex	spotG	-37616.85115	0	4123009.812	0
SOL	bitfinex	spotH	6377.137757	0.147	3768823.717	0
SOL	bitfinex	spotF	-18408.75106	0	1566337.795	0
SOL	coinbase-pro	spotG	-42227.43159	0	3624499.635	0
SOL	coinbase-pro	spotH	9649.294162	0	3326902.294	0
SOL	ftx	spotC	232.9263018	0.002	21562.80743	0
SOL	ftx	spotD	163.3406109	0	11872.8024	0
SOL	ftx	spotE	-4487.782868	0	371509.2895	0
SOL	ftx	spotF	-6402.437719	0	1007953.704	0
SOL	ftx	spotG	-4961.442475	0	1414542.39	0

SOL	ftx	spotH	-364.2505365	0.381	1115975.745	0

Table A.2: OLS Depth values

A.2.3 NTR

Asset	Exchange	Contract	Beta	Beta P-Value	C	C P-Value
BTC	Bitmex	XBTH20	-0.001522397	0.002	0.138628352	0
BTC	Bitmex	XBTM20	0.000374291	0.003	0.04044932	0
BTC	Bitmex	XBTU20	-0.000199709	0.07	0.047593304	0
BTC	Bitmex	XBTZ20	-0.000457693	0	0.069473149	0
BTC	Bitmex	XBTH21	0.000454728	0.009	0.071733285	0
BTC	Bitmex	XBTM21	-0.000533648	0.026	0.112202069	0
BTC	Bitmex	XBTU21	0.000215049	0.388	0.050733546	0
BTC	Bitmex	XBTZ21	3.77E-05	0.765	0.054925551	0
BTC	Deribit	BTC-25SEP20	-0.000152859	0.339	0.044481455	0
BTC	Deribit	BTC-25DEC20	-0.00043391	0	0.067684248	0
BTC	Deribit	BTC-26MAR21	0.000230913	0.135	0.07493208	0
BTC	Deribit	BTC-25JUN21	-0.000451098	0.006	0.09808131	0
BTC	Deribit	BTC-24SEP21	3.84E-05	0.763	0.057078415	0
BTC	Deribit	BTC-31DEC21	5.21E-05	0.647	0.054036403	0
BTC	Ftx	BTC-20200626	0.000542525	0.01	0.036024624	0
BTC	Ftx	BTC-20200925	-0.000193066	0.08	0.048076057	0
BTC	Ftx	BTC-20201225	-0.000463554	0	0.06991	0
BTC	Ftx	BTC-20210326	0.000442749	0.007	0.072701972	0
BTC	Ftx	BTC-20210625	-0.000541007	0.017	0.111425464	0
BTC	Ftx	BTC-20210924	7.54E-05	0.528	0.054042429	0
BTC	Ftx	BTC-20211231	7.03E-05	0.523	0.053549954	0
ETH	Deribit	ETH-25SEP20	-0.000599397	0.019	0.094302452	0
ETH	Deribit	ETH-25DEC20	-0.000385406	0	0.077891931	0

ETH	Deribit	ETH-26MAR21	0.000553193	0.011	0.078852687	0
ETH	Deribit	ETH-25JUN21	-0.000517466	0.008	0.124097158	0
ETH	Deribit	ETH-24SEP21	2.29E-06	0.988	0.076588264	0
ETH	Deribit	ETH-31DEC21	8.72E-05	0.449	0.062663921	0
ETH	Ftx	ETH-20200626	0.000473329	0.024	0.045228727	0
ETH	Ftx	ETH-20200925	-0.00065605	0	0.099001112	0
ETH	Ftx	ETH-20201225	-0.000444767	0	0.082979147	0
ETH	Ftx	ETH-20210326	0.000762159	0.001	0.076544248	0
ETH	Ftx	ETH-20210625	-0.000650271	0.056	0.145383824	0
ETH	Ftx	ETH-20210924	2.56E-05	0.857	0.075742516	0
ETH	Ftx	ETH-20211231	9.35E-05	0.417	0.063163246	0
DOGE	Ftx	DOGE-20200626	0.000932856	0	0.016380986	0.012
DOGE	Ftx	DOGE-20200925	0.000810224	0.003	0.029637431	0.034
DOGE	Ftx	DOGE-20201225	-0.001206008	0	0.114998021	0
DOGE	Ftx	DOGE-20210326	0.001603223	0.043	0.110095511	0.008
DOGE	Ftx	DOGE-20210625	0.000299023	0.645	0.183497395	0
DOGE	Ftx	DOGE-20210924	8.22E-05	0.737	0.089065074	0
DOGE	Ftx	DOGE-20211231	7.35E-05	0.698	0.080951588	0
SOL	Ftx	SOL-20200925	-0.000256498	0.705	0.190273498	0
SOL	Ftx	SOL-20201225	-0.000139844	0.565	0.117980197	0
SOL	Ftx	SOL-20210326	0.001092566	0.001	0.102010581	0
SOL	Ftx	SOL-20210625	-0.000712832	0.107	0.196178634	0
SOL	Ftx	SOL-20210924	-0.00094289	0	0.163539779	0
SOL	Ftx	SOL-20211231	0.00012977	0.26	0.081470096	0
BTC	Bitfinex	SpotA	-0.001300656	0	0.120461115	0
BTC	Bitfinex	SpotB	0.000326387	0.007	0.037051994	0
BTC	Bitfinex	SpotC	-0.000207652	0.035	0.045282283	0
BTC	Bitfinex	SpotD	-0.000427375	0	0.064312156	0
BTC	Bitfinex	SpotE	0.000415589	0.007	0.067437981	0

BTC	Bitfinex	SpotG	6.06E-05	0.596	0.052151518	0
BTC	Bitfinex	SpotH	2.17E-05	0.836	0.051871891	0
BTC	Bitfinex	SpotF	-0.000672014	0.001	0.108959493	0
BTC	Coinbase-Pro	SpotA	-0.00131753	0	0.123549434	0
BTC	Coinbase-Pro	SpotB	0.000343776	0.005	0.03768236	0
BTC	Coinbase-Pro	SpotC	-0.000212659	0.035	0.046402111	0
BTC	Coinbase-Pro	SpotD	-0.000457538	0	0.067167496	0
BTC	Coinbase-Pro	SpotE	0.000443002	0.012	0.066746958	0
BTC	Coinbase-Pro	SpotF	-0.000706577	0.001	0.111105786	0
BTC	Coinbase-Pro	SpotG	6.10E-05	0.61	0.052832797	0
BTC	Coinbase-Pro	SpotH	2.88E-05	0.776	0.051891091	0
BTC	Ftx	SpotB	0.000474412	0.02	0.035210409	0
BTC	Ftx	SpotC	-0.000218215	0.03	0.046368383	0
BTC	Ftx	SpotD	-0.000467934	0	0.067548531	0
BTC	Ftx	SpotE	0.000389168	0.02	0.070630751	0
BTC	Ftx	SpotG	5.77E-05	0.622	0.053194432	0
BTC	Ftx	SpotH	4.01E-05	0.709	0.052569442	0
BTC	Ftx	SpotF	-0.000686686	0.001	0.111227369	0
DOGE	Bitfinex	SpotG	0.000105473	0.659	0.087366374	0
DOGE	Bitfinex	SpotH	-4.13E-05	0.851	0.088654282	0
DOGE	Bitfinex	SpotF	0.001749008	0.071	0.134921787	0
DOGE	Coinbase-Pro	SpotG	0.000117698	0.624	0.08713977	0
DOGE	Coinbase-Pro	SpotH	-5.37E-05	0.81	0.089938967	0
DOGE	Ftx	SpotE	0.003198746	0.007	0.068065703	0.157
DOGE	Ftx	SpotG	0.000124931	0.603	0.087501331	0
DOGE	Ftx	SpotH	-1.80E-06	0.994	0.090197982	0
DOGE	Ftx	SpotF	0.000327644	0.618	0.183790906	0
ETH	Bitfinex	SpotB	0.000379824	0.045	0.043241667	0
ETH	Bitfinex	SpotC	-0.000647172	0	0.095322002	0

ETH	Bitfinex	SpotD	-0.000462053	0	0.07964413	0
ETH	Bitfinex	SpotE	0.000761775	0	0.068390411	0
ETH	Bitfinex	SpotF	-0.000749409	0.014	0.141254828	0
ETH	Bitfinex	SpotG	5.68E-06	0.967	0.071944969	0
ETH	Bitfinex	SpotH	4.87E-05	0.659	0.060618079	0
ETH	Coinbase-Pro	SpotB	0.000381872	0.045	0.04360708	0
ETH	Coinbase-Pro	SpotC	-0.000635514	0	0.095407087	0
ETH	Coinbase-Pro	SpotD	-0.000459972	0	0.080258032	0
ETH	Coinbase-Pro	SpotE	0.00076875	0.001	0.067043173	0
ETH	Coinbase-Pro	SpotG	-1.04E-05	0.942	0.073160694	0
ETH	Coinbase-Pro	SpotH	5.48E-05	0.605	0.06083337	0
ETH	Coinbase-Pro	SpotF	-0.000774971	0.018	0.144864387	0
ETH	Ftx	SpotB	0.000354213	0.068	0.044638115	0
ETH	Ftx	SpotC	-0.000624997	0.001	0.096526091	0
ETH	Ftx	SpotD	-0.000475122	0	0.080690006	0
ETH	Ftx	SpotE	0.000693425	0.002	0.074156254	0
ETH	Ftx	SpotG	1.72E-05	0.906	0.073823219	0
ETH	Ftx	SpotH	9.62E-05	0.437	0.061031133	0
ETH	Ftx	SpotF	-0.000777927	0.02	0.145712765	0
SOL	Bitfinex	SpotE	0.008140699	0.006	0.023931968	0.593
SOL	Bitfinex	SpotG	-0.001064012	0	0.170330262	0
SOL	Bitfinex	SpotH	7.96E-05	0.515	0.084318157	0
SOL	Bitfinex	SpotF	-0.000348035	0.469	0.181840905	0
SOL	Coinbase-Pro	SpotG	-0.000939372	0	0.163439256	0
SOL	Coinbase-Pro	SpotH	8.34E-05	0.492	0.085157096	0
SOL	Ftx	SpotC	-7.64E-05	0.909	0.184191964	0
SOL	Ftx	SpotD	-0.000253223	0.32	0.125455426	0
SOL	Ftx	SpotE	0.000908064	0.003	0.109427623	0
SOL	Ftx	SpotF	-0.000868846	0.058	0.20515085	0

SOL	Ftx	SpotG	-0.001011513	0	0.168386186	0
SOL	Ftx	SpotH	0.000128903	0.268	0.081909454	0

Table A.3: OLS NTR values