Price Discovery of Bitcoin in the ETF and Futures Markets

by

Kiana Kia

MSS in Economics, University of Tehran, Iran, 2017

An Extended Essay Submitted in Partial Fulfillment of the Requirements for the Degree of

MASTER OF ARTS

in the Department of Economics

We accept this extended essay as conforming

to the required standard

Ke Xu

Dr. Ke Xu, Supervisor (Department of Economics)

Dr. Kenneth Stewart, Member (Department of Economics)

Dr. Tao Wang, Member (Department of Economics)

Abstract

In this research, we explore the price discovery dynamics across three bitcoin markets: spot, futures, and exchange-traded funds (ETFs), using minute-level bitcoin price data. Given the recent emergence of the bitcoin ETF market, our study emphasizes its price discovery contribution in comparison to the established spot and futures markets. Utilizing the Fractionally Cointegrated Vector Autoregressive (FCVAR) model, we analyze data from October 19, 2021—the launch date of the first US bitcoin ETF—to December 30, 2022. Our empirical results confirm fractional cointegration across the markets, highlighting persistent long-run relationships between market pairs. More importantly, the findings reveal the ETF market's dominant role in leading the price discovery process in the bitcoin market.

Contents

List of Figures

List of Tables

1 Introduction

The process of price discovery, central to financial markets, empowers market participants to make informed investment decisions and allocate resources effectively. At its core lies an asset's fundamental value, which is derived from underlying economic indicators such as earnings, cash flow, dividends, financial ratios, and broader macroeconomic factors (O'Hara, 2003). These fundamentals guide market participants in discerning an asset's true value. However, the market price can occasionally deviate from this intrinsic value due to various factors, including market sentiment, supply-demand imbalances, liquidity constraints, or unexpected market shocks. Historically, the lens of price discovery has been focused on the timing interplay between spot and futures prices. Yet, contemporary research methodologies, notably the Gonzalo-Granger decomposition and the fractionally cointegrated vector autoregression (FCVAR), have paved the way for a deeper understanding of the interrelationships and adjustments among asset prices over time, encompassing both commodities (Dolatabadi, Nielsen and Xu, 2015) and crypto markets (Wu et al., 2021).

Bitcoin's significant growth as a leading digital cryptocurrency highlights the complexities of the price discovery process across markets. As Bitcoin carved its niche and began influencing traditional financial markets, a continuous influx of new information—from its technological advantage to its hedging potential—has been instrumental in shaping its price trajectory. Cryptocurrency ETFs, particularly Bitcoin ETFs, have emerged as significant players in this landscape. These exchange-traded funds invest in cryptocurrencies or companies involved in the cryptocurrency industry, offering a more accessible and regulated avenue for investors to gain exposure to cryptocurrencies (Singh, 2022). Specifically, Bitcoin ETFs hold Bitcoin derivatives and are structured as a series of trusts and funds managed by a sponsor, with futures contracts serving as the benchmark for the fund (Brown, 2019). While many investors opt for holding spot Bitcoin directly, others are drawn to Bitcoin ETFs due to factors such as tax considerations, the security of a regulated exchange, and the flexibility of trading (Arslanian, 2022). The introduction of Bitcoin ETFs has underscored the need to understand how this financial innovation could alter price discovery mechanisms in the Bitcoin markets.

Building upon this context, our study seeks to enhance the existing literature by examining the influence of the newly introduced Bitcoin ETFs on the price discovery process within Bitcoin markets. Our objective is to shed light on the evolving dynamics of this emerging digital asset class. To address our research question, we apply the Fractionally Cointegrated VAR (FCVAR) model by Johansen (2008) to estimate price discovery in three distinct Bitcoin market pairs: spot-futures, spot-ETF, and futures-ETF. We utilize high-frequency 1-minute intraday Bitcoin price data spanning from October 19, 2021, to December 30, 2022, to pinpoint the role of the Bitcoin ETF market in the price discovery mechanism. By extending the CVAR model to allow for fractional cointegration, our findings align with those of Dolatabadi, Nielsen and Xu (2015). We confirm that the FCVAR model fits Bitcoin data better than the CVAR model, giving a more precise estimate of price discovery among different Bitcoin markets.

Our primary focus is the ProShares Bitcoin ETF which tracks the performance of Bitcoin futures price. Our empirical findings highlight the significant role of the ProShares Bitcoin ETF in price discovery, especially in its interactions with Bitcoin futures and spot market. As depicted in Figure 1, the dollar value of the Bitcoin ETF frequently surpasses that of futures, reinforcing its substantial influence in price discovery. ETFs, particularly those akin to the ProShares Bitcoin ETF, are celebrated for their liquidity, granting investors exposure to the underlying asset without holding it directly. This liquidity, when contrasted with the potential inaccessibility of futures markets, positions ETFs as key players in the price discovery process because it facilitates investment by mitigating the complexities and risks associated with direct futures trading.

While trading volume is a salient metric, it doesn't fully encapsulate the nuances of price discovery. Other determinants, such as market efficiency, accessibility, and the composition of market participants, play crucial roles. The ETF market, potentially drawing a larger portion of informed traders, may exert a greater influence on price discovery compared to the spot market. The structured nature of the ETF, coupled with its alignment with the futures market, could make it more responsive to changes in sentiment or new information. For example, Duffy, Rabanal and Rud (2021) find that ETFs do not harm, and may in fact improve, price discovery and liquidity in the underlying asset market. Such ETFs not only prevent detrimental effects on price discovery but can also enhance it, thereby improving market efficiency in asset markets.

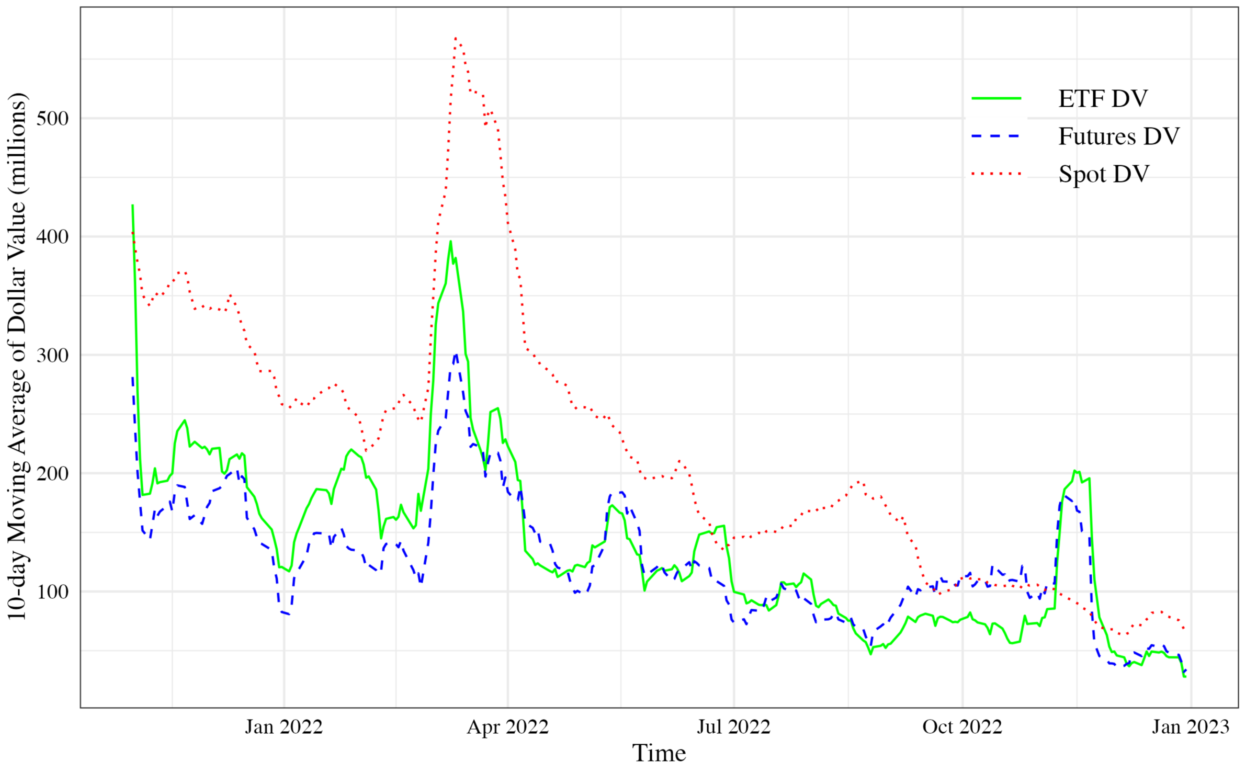

Figure 1: The 10-day moving average of the trading volume of Bitcoin spot, futures, and ETF from https://barchart.com, for the period October 19, 2021, to December 30, 2022.

Furthermore, for securities that are less liquid than the ETF, the intra-day price discovery of securities can be improved by ETFs (Ivanov, Jones and Zaima, 2013). This is because ETFs respond faster to new information and are traded at lower costs, which can enhance liquidity and price discovery of the underlying securities, and increase market efficiency (Duffy, Rabanal and Rud, 2021). Madhavan and Sobczyk (2016) developed a model to analyze ETF price dynamics and ETFs' unique creation and redemption mechanisms. Through empirical testing, they found that ETFs can accelerate the price discovery of securities if there are no frictions in the arbitrage process. (Glosten, Nallareddy and Zou, 2021) investigates the effect of ETF activity on the short-run informational efficiency of underlying stocks and finds that greater ETF activity leads to an improvement in short-run informational efficiency, particularly for firms with weak information environments. The increase in correlation linked to ETF activity can be partially explained by systematic fundamental information.

Shrestha, Philip and Peranginangin (2020) investigates the contributions of three crude oil-based ETFs in the price discovery process. Using daily data on the crude oil spot price, near-month[1] crude oil futures, and three crude-oil-based ETFs, the researchers analyze the price discovery contributions of the five price series. They found that the futures market dominates the price discovery process. However, they also found that the crude-oil-based ETFs significantly contribute to the price discovery

[1]"Near month" in the context of futures contracts refers to the contract that is closest to its expiration date.

process. Thus, they concluded that additional ETFs play a significant role in price discovery.

Hegde and McDermott (2004) and Richie and Madura (2007) agree that ETFs improve the liquidity of their component stocks during non-turbulent market times. However, in times of financial distress, (Pan and Zeng, 2017) believes the liquidity provision of ETFs can deteriorate. Ben-David, Franzoni and Moussawi (2018) provides two hypotheses regarding how ETFs affect market efficiency. The price discovery hypothesis argues that ETFs increase market efficiency, while the liquidity trading hypothesis argues that they decrease it. Both mechanisms may coexist as ETFs increase liquidity and information efficiency but also reflect non-fundamental information in prices due to noise traders. Our research findings enhance the exiting literature, demonstrating that in the Bitcoin market, ETFs play a more significant role in price discovery compared to spot and futures.

The structure of this paper unfolds as follows: Section 2 delves into a detailed exploration of ETFs, with a particular emphasis on Bitcoin ETFs. In Section 3, we outline the research methodology, paving the way for an in-depth empirical analysis presented in Section 4. Section 5 illustrates our empirical results, while Section 6 include a discussion on the structural break test. We wrap up the paper in the concluding section, summarizing our findings and reflecting on their broader implications.

2 Bitcoin ETF

An ETF is a type of investment fund that trades on stock exchanges like a stock. Its purpose is to track the performance of a particular index, commodity, or other asset. ETFs can be bought and sold throughout the day, just like individual stocks, and their prices fluctuate based on supply and demand. While ETFs share similarities with mutual funds, such as representing a collection of investments, they differ in key ways. ETFs are typically passively managed, aiming to match the performance of an index or benchmark, whereas mutual funds are often actively managed by a professional fund manager who selects and manages individual investments, resulting in higher fees than ETFs.

The growth of the overall ETF market, which encompasses various asset classes like stocks, bonds, commodities, and more recently, cryptocurrencies like Bitcoin, has been substantial. The number of ETFs has increased from 276 in 2003 to 8,754 in 2022, and assets under management have grown to nearly 10 trillion U.S. dollars, highlighting the popularity of these investment vehicles. ETFs have gained significant traction among investors due to their high liquidity, as they can be traded throughout the trading day, transparency from daily disclosures of holdings, and potential tax advantages. However, it is important to recognize that ETFs can also have downsides, including the possibility of overvaluation or undervaluation and potential market imbalances and instability, as discussed by Duffy, Rabanal and Rud (2021). They explained that if investors focus more on buying and selling ETFs than the underlying assets they represent, or if the price of the ETF becomes disconnected from its actual value, represented by its net asset value (NAV), it could signal that the ETF is overvalued or undervalued relative to its underlying assets, leading to market imbalances and negative consequences for asset market. For a more detailed understanding of ETFs, Deville (2008) conduct a comprehensive study on the history of ETFs.

The market for Bitcoin ETFs is rapidly developing and evolving with the proposal and regulation of diverse types of Bitcoin ETFs by various authorities. The regulation of Bitcoin ETFs varies depending on the country and the type of ETF. Generally, Bitcoin ETFs are subject to the same rules and regulations as other ETFs in their respective jurisdictions. Currently, Bitcoin ETFs are only available in a few countries, including several European nations, the US, Canada, and Brazil.

Bitcoin ETFs come in different types, depending on how they track the price of Bitcoin. In the United States, the Securities and Exchange Commission (SEC) regulates Bitcoin ETFs and is responsible for protecting investors and ensuring fair and orderly markets. To date, the SEC has only approved a derivatives-based Bitcoin ETF, which tracks the price of Bitcoin indirectly through the futures market. This first type of ETF invests in Bitcoin futures contracts instead of owning spot Bitcoin, and it enables investors to buy or sell Bitcoin at a predetermined price and date in the futures market.

Bitcoin futures, regulated by the Commodity Futures Trading Commission (CFTC)[2], are a type of futures contract that allows investors to bet on the price of Bitcoin at a future date. The CFTC considers Bitcoin a commodity, so Bitcoin futures are commodity futures. These contracts trade on several platforms, such as the Chicago Mercantile Exchange (CME), Bakkt, and Bitnimial. Meanwhile, Bitcoin ETFs are regulated by the Securities and Exchange Commission (SEC) in the United States and the Canadian Securities Administrators (CSA) in Canada.

The second type of Bitcoin ETF is spot Bitcoin ETF, which invests directly in Bitcoins held by a custodian. These ETFs aim to expose investors to Bitcoin's actual performance without having to deal with its technical challenges, allowing them to track the current spot market price of Bitcoin more precisely and accurately. While spot Bitcoin ETFs usually have lower fees, they may face higher regulatory hurdles than Bitcoin futures ETFs and have a greater likelihood of tracking errors. An example of this type of ETF is the Grayscale Bitcoin Trust, which is not technically an ETF but a trust that issues shares backed by Bitcoin. However, the SEC has rejected or delayed several proposals for a spot-based Bitcoin ETF that would directly track the spot price of Bitcoin. The SEC has expressed concerns about the potential for fraud, manipulation, and a lack of transparency in the spot market for Bitcoin, making it challenging for spot-based Bitcoin ETFs to be approved.

[2] CFTC is an independent federal agency that oversees the U.S. derivatives markets.

A third type of Bitcoin ETF uses derivatives other than futures, such as options, swaps, or contracts for difference (CFDs). These financial instruments derive their value from the price of another asset, such as Bitcoin. These ETFs also do not own Bitcoin directly but use derivatives to gain exposure to its price fluctuations. An example of this type of ETF is the WisdomTree Enhanced Commodity Strategy Fund, which allocates a portion of its portfolio to CFDs on Bitcoin. In this paper, we use the futures-based Bitcoin ETF to analyze the price discovery in Bitcoin market.

3 Methodology

3.1 The FCVAR model and interpretation of the parameters

There are two popular measures used for price discovery, namely component share (Gonzalo and Granger, 1995) and information share (Hasbrouck, 1995). The first is constructed on estimates from a cointegrated vector autoregressive (CVAR) model. (Johansen, 2008) improved upon the CVAR model by proposing the fractionally cointegrated VAR (FCVAR) model. (Johansen and Nielsen, 2012) prove the consistency of the maximum likelihood estimators of the FCVAR model. Our empirical analysis follows the FCVAR model proposed by (Dolatabadi, Nielsen and Xu, 2016), accommodating a deterministic trend. For a p-dimensional $I(1)$ time series of $Y_t, t = 1, \ldots, T$, the CVAR model (Dolatabadi, Nielsen and Xu, 2016) is

$$\Delta Y_t = \alpha\beta' Y_{t-1} + \sum_{i=1}^{k} \Gamma_i \Delta Y_{t-i} + \varepsilon_t = \alpha\beta' L Y_t + \sum_{i=1}^{k} \Gamma_i \Delta L^i Y_t + \varepsilon_t \tag{3.1}$$

This model is commonly used to analyze long-run economic relationships represented by stationary combinations of Y_t (Johansen and Nielsen, 2012). In our analysis, we consider three time series, comprising two pairs of integrated time series (of order one), including $Y_t = [f_t, e_t]$ and $Y_t = [s_t, f_t]$. Here, f_t represents the futures price, e_t represents the ETF price, and s_t represents the spot price. The parameter β represents a $p \times r$ matrix of cointegrating vectors showing the long-run equilibrium relationship between variables where $0 \leq r \leq p$ represents the cointegration rank. In the case of just two variables, such as futures and ETF, $Y_t = [f_t, e_t]$, the cointegration relationship (Xu, Stewart and Cao, 2022) is expressed as follows:

$$f_t = \beta_2 e_t + \rho \tag{3.2}$$

and then

$$f_t - \beta_2 e_t - \rho = \beta' Y_t - \rho \tag{3.3}$$

While $\beta' = [1, -\beta_2]$ is the cointegrating vector in the FCVAR model, which implies that the two variables in the model are cointegrated with a long-run equilibrium relationship defined by the linear combination of the two variables with coefficients 1 and $-\beta_2$, respectively. And any deviations from this long-run equilibrium relationship will eventually be corrected in the long run, as the cointegrating vector represents the long-run equilibrium relationship between the variables in the model. Between the $I(1)$ variables e_t and f_t, there exists a long-run equilibrium relationship that makes this $I(0)$ linear combination (Xu, Stewart and Cao, 2022).

The simplest way to derive the FCVAR model from the CVAR model is to replace the difference and lag operators, Δ and L, in equation (1) by their fractional counterparts, Δ^b and $L_b = 1 - \Delta^b$, respectively. This yields the following FCVAR model:

$$\Delta^b Y_t = \alpha\beta' L_b Y_t + \sum_{i=1}^{k} \Gamma_i \Delta^b L_b^i Y_t + \varepsilon_t \tag{3.4}$$

where b represents fractional parameter and we apply $Y_t = \Delta^{d-b} X_t$ to obtain the FCVAR model,

$$\Delta^d X_t = \alpha\beta' L_b \Delta^{d-b} X_t + \sum_{i=1}^{k} \Gamma_i \Delta^d L_b^i X_t + \varepsilon_t \tag{3.5}$$

This means that if the time series vector X_t has components $x_{1t} \sim I(d)$ and $x_{2t} \sim I(d)$, and there exists a vector $\beta' = [1, -\beta_2]$ (Saha, Madhavan and Chandrashekhar, 2022) such that the linear combination $\beta' X_t = [x_{1t} - \beta_2 x_{2t}]$ is integrated of order $d - b$, then the time series x_{1t} and x_{2t} are said to be fractionally cointegrated of order (d, b), and β_2 is the coefficient of x_{2t} in this relationship. The presence of fractional cointegration (or long memory) in the equilibrium relation between ETF and futures prices implies that although the prices themselves are I(1), $\beta' X_t$, the linear combination of the variables in the system is stationary $I(1 - b)$. Based on (Johansen and Nielsen, 2012), when $0 \leq r \leq p$, X_t is fractional of order d and cointegrated of order $d - b$; that is, $\beta' X_t$ is $I(d - b)$.

The order $d - b$ of the linear combination $\beta' X_t$ represents the degree of persistence in the long-run relationship between the two time series, with higher values indicating stronger persistence. Following (Dolatabadi, Nielsen and Xu, 2015) , we assume asset prices are $I(1)$ and we estimate the fractional parameter b, which determines the degree of fractional cointegration. Therefore, as b converges to one, i.e., $b \to 1$, $1 - b \to 0$, the FCVAR model reduces to a special case known as the CVAR model.

The parameter α in FCVAR equation (3.4) is the error correction term and represents the speed of adjustment towards equilibrium for each variable. The orthogonal of the speed of adjustment parameters, $\alpha_\perp$, would be the price discovery contribution of each market. The short-run dynamics of the variables are captured by the parameters Γ_i, with k representing the number of lags used to capture short-term dynamics. The error term, ε_t, is p-dimensional, independent, and identically distributed, with a mean of zero and covariance matrix Ω.

We also include a restricted constant ρ as shown in equation (3.5). The restricted constant term ρ represents the mean of the long-run equilibrium in the FCVAR model. The equation $E[\beta' X_t + \rho'] = 0$ implies that the linear combination of the variables in the system, $\beta' X_t + \rho'$, has a mean of zero. The parameter ξ is the unrestricted constant term and can produce a deterministic trend in the levels of the variables. This deterministic trend captures systematic patterns or movements in the data that are not due to the stochastic processes represented by the other terms in the model.

$$\Delta^d X_t = \alpha \Delta^{d-b} L_b \left(\beta' X_t + \rho'\right) + \sum_{i=1}^{k} \Gamma_i \Delta^d L_b^i X_t + \xi + \varepsilon_t \tag{3.6}$$

Thus, the FCVAR model allows simultaneous modeling of the long-run equilibrium, the adjustment responses to deviations from the equilibrium, and the short-run dynamics of the system (Cárcel and Gil-Alana, 2017).

In our empirical analysis, we follow (Johansen and Nielsen, 2016) by splitting the observed sample into initial values to be conditioned upon and then applying maximum likelihood inference. We assume a sample of length $T + N$ is available on X_t, where N denotes the number of observations used for conditioning. Finally, after estimating the FCVAR parameters and α, we can find the value of $\alpha_\perp$ by normalizing and applying two conditions: $\alpha' \alpha_\perp = 0$ and $\alpha_{\perp,1} + \alpha_{\perp,2} = 1$, where $\alpha_{\perp,1}$ and $\alpha_{\perp,2}$ represent each market's contribution to the price discovery process.

3.2 Hypothesis tests

We conduct the relevant hypothesis tests for the FCVAR model, and the results are shown in Tables 4, 5, and 6 for each market pair. The first hypothesis test determines whether the FCVAR model is a better fit for the data than the CVAR model. The null hypothesis is $H_0 : b = d = 1$. The second hypothesis test is on the cointegration vectors β, which represent the long-run equilibrium relationship between two or more time series. The null hypothesis is that there is no long-run contango or backwardation relationship between the series, with β set to $(1, -1)$, while the alternative hypothesis is that there is a long-run contango or backwardation relationship between the series. This implies that there is a stationary linear combination of non-stationary variables (i.e., it has a constant mean and variance over time). The last two hypothesis tests focus on $\alpha_\perp$, which quantifies the price discovery process between the two markets. These tests are predicated on the assumption that price discovery is exclusive to one of the paired markets. Given a vector of futures and ETF as $[f_t, e_t]$, and assuming that price discovery occurs solely in the futures market, the hypothesis is formulated as $H_0 : \hat{\alpha}_\perp = [a_{\perp,1}, 0]'$. Conversely, if price discovery is believed to be exclusive to the ETF market, the hypothesis is expressed as $H_0 : \hat{\alpha}_{\perp} = [0, a_{\perp,2}]'$.

4 Empirical analysis

We analyze intraday price data denominated in US dollars at one-minute intervals, focusing on three markets: the Bitcoin spot prices (ticker: BTCUSD), the Bitcoin futures (ticker: BTC) from the Chicago Mercantile Exchange (CME), and the ProShares Bitcoin Strategy ETF, which tracks Bitcoin futures contracts on the New York Stock Exchange (NYSE Arca) under the ticker BITO. BITO, launched by ProShares, a leading ETF provider, on October 19, 2021, became the first ETF in the US to offer exposure to Bitcoin returns. On its debut day, over 24 million shares traded hands, and the assets under management surged to $\$1$ billion in just two days.

While the Bitcoin spot (BTCUSD) operates 24/7 without breaks on holidays or weekends, the Bitcoin ETF trades from Monday through Friday, 9:30 a.m. to 3:59 p.m. The CME futures market, on the other hand, is inactive on Saturdays and certain holidays, with trading hours spanning 5:00 p.m. Sunday evening to 4:00 p.m. Friday afternoon. Additionally, the CME futures market observes a daily maintenance period from 4:00 p.m. to 5:00 p.m. CT, Monday through Thursday. By aligning the trading hours across all three markets, we compiled a total of 10,9854 observations from October 19, 2021, to December 30, 2022. As illustrated in Figure 2, Bitcoin's spot and futures prices are often closely matched, at times mirroring each other to such an extent that they appear almost indistinguishable.

In Figure 2, the ProShares Bitcoin Strategy ETF (BITO), which tracks Bitcoin futures prices, mirrors the price trajectories of the underlying futures, a testament to the effective price discovery mechanism shared by both instruments (Lin, Chou and Wang, 2018). This overlap underscores the importance of market anticipations in setting the futures and, by proxy, the ETF's prices. An expected upswing in Bitcoin's value can result in surging demand for the ETF, thereby elevating its price. This intensified demand, reflective of the market's shared vision for Bitcoin's imminent value, plays a pivotal role in the price discovery process. In contrast, if the market's foresight proves errant and Bitcoin's price doesn't soar as predicted, the ETF's valuation would suffer a setback, entailing losses for its investors. This underlines the significance of accurately gauging market projections when dissecting the performance metrics of futures-centric ETFs like the ProShares Bitcoin Strategy ETF.

The log prices of Bitcoin spot, futures, and ETF, all three, share a high degree of correlation. Factors such as market sentiment, novel information, and shifts in supply and demand dynamics influence them together, signaling their cointegration. Yet, discrepancies in pricing between the ETF

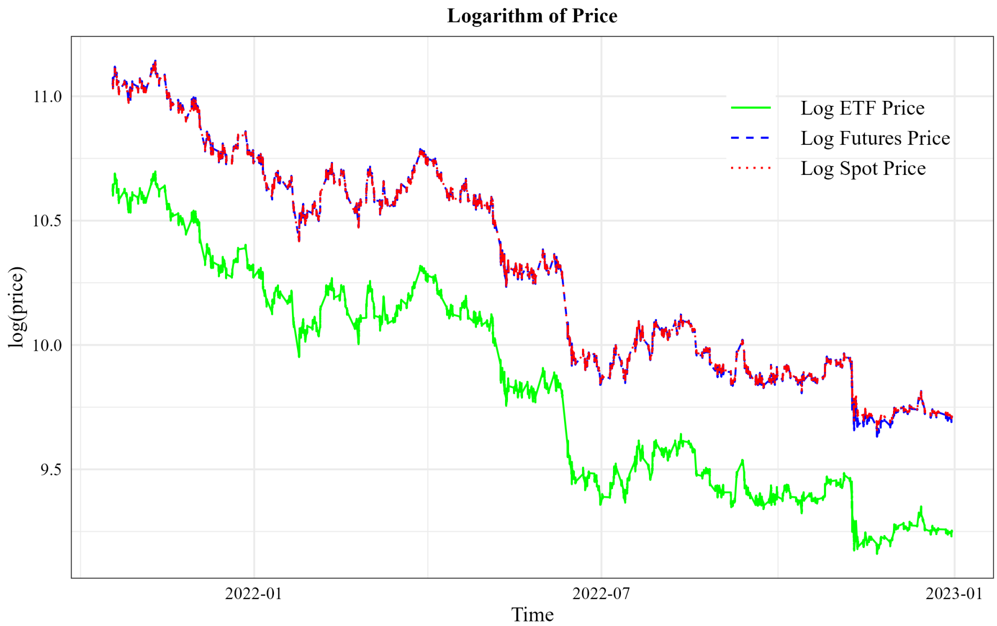

Figure 2: Log prices of Bitcoin spot, futures, and ETF from October 19, 2021, to December 30, 2022.

and its associated futures might surface, attributable to elements beyond the asset's core value, such as transaction costs, liquidity fluxes, or market inefficiencies.

4.1 Summary statistics

Table 1 offers detailed summary statistics for the log prices of Bitcoin in the spot, futures, and ETF markets. The analysis employs one-minute intraday data from October 19, 2021, to December 30, 2022. The objective of this study is to unravel the intricate relationships between these markets and pinpoint the contribution of price discovery in each market. The futures prices align closely with the spot prices on average. In an ideally efficient market, futures prices should mirror expected spot prices. The ETF's marginally lower mean log price, relative to the spot and futures markets, hints that the ETF might not mirror futures price movements on a 1:1 scale. Such disparities can emerge from management fees or tracking errors. The slim average spread between spot and futures signifies a well-functioning market. Nonetheless, the pronounced kurtosis indicates occasional sizeable shifts, which could be influenced by market anomalies, liquidity issues, or impactful news events.

The negative mean returns suggest that, within the sample period, the Bitcoin prices in all three markets leaned more towards downward shifts. For investors, this points towards a predominantly bearish trend. The high standard deviation in Bitcoin returns reaffirms its volatile nature, highlighting the associated investment risks. Though the ETF is designed to track the futures market, its slightly elevated standard deviation indicates a potentially heightened reaction to market shifts.

The statistics illustrate that while Bitcoin's spot, futures, and ETF prices are tightly intercon-

Table 1: Summary statistics of Bitcoin spot, futures, and ETF log prices and log returns

	p_t^S	p_t^F	p_t^E	$p_t^S - p_t^F$	$p_t^S - p_t^E$	$p_t^F - p_t^E$	r_t^S	r_t^F	r_t^E
mean	10.329	10.328	9.858	0.001	0.472	0.471	-0.129	-0.130	-0.133
median	10.346	10.345	9.870	0.000	0.472	0.471	-0.110	-0.104	-0.097
Maximum	11.137	11.143	10.698	0.106	0.553	0.498	2.877	3.270	3.175
Minimum	9.657	9.629	9.160	-0.043	0.407	0.428	-5.801	-5.842	-5.868
Standard Deviation	0.421	0.424	0.434	0.006	0.015	0.012	1.077	1.142	1.139
Skewness	0.063	0.058	0.084	3.277	-0.362	-0.831	-0.616	-0.572	-0.582
Excess Kurtosis	-1.385	-1.372	-1.361	18.406	1.179	0.485	3.083	2.715	2.770

The series of prices are based on one-minute intraday data from https://www.Barchart.com from October 19, 2021, to December 30, 2022. The first three columns are log prices of Bitcoin spot, futures, and ETF. The next three columns are the spread between spot-futures, spot-ETF, and futures-ETF prices. The last three columns are the log returns of Bitcoin prices in three spot, futures, and ETF markets.

nected, there exists a notable variation in the spread between spot and futures prices. The extreme values and non-normal distribution for certain variables mirror the Bitcoin market's volatile and unpredictable nature. Such dynamics are characteristic of cryptocurrency markets, where factors like shifting market sentiment, regulatory changes, and significant news events can instigate sudden and robust price movements.

To assess the stationarity of our time series, we employ the augmented Dickey-Fuller (ADF) test. The results in the first section of Table 2 show the results of the ADF test for the log prices of ETFs, spots, and futures. The p-value exceeds the 5% significance level, suggesting that the unit root hypothesis cannot be rejected for the log prices of Bitcoin ETF, futures, and spot. Given this non-stationarity, we turn to cointegration analysis. This approach allows us to explore potential long-term relationships between these time series and estimate price discovery contribution of each market.

The subsequent sections of Table 2 highlight the stationarity of both spreads and returns. The stationary nature of the spreads indicates that the relationships between the spot, futures, and ETF prices remain relatively stable over time. Given these consistent relationships, it becomes appropriate to employ the FCVAR model to estimate the price discovery among the spot, futures, and ETF prices.

4.2 The FCVAR model specification

Before delving into the FCVAR model estimation and hypothesis testing, we must address several model selection decisions. The correct specification of the vector autoregressive (VAR) model is paramount, as all empirical inferences drawn from it hinge on this specification (Gutierrez, Souza and de Carvalho Guillén, 2009). One of the pivotal steps in this specification for error correction models is the selection of the lag length (Agunloye and Shangodoyin, 2014). We employed various criteria, including the Bayesian Information Criterion (BIC), the Akaike Information Criterion (AIC),

Table 2: Augmented Dickey-Fuller Test

	t Statistics	P-values
Spot log prices	-2.675	0.290
Futures log prices	-2.738	0.263
ETF log prices	-2.671	0.292
Spread $p_t^S - p_t^F$	-7.776	0.010
Spread $p_t^S - p_t^E$	-4.187	0.010
Spread $p_t^F - p_t^E$	-4.143	0.000
Spot log returns	-47.052	0.010
Futures log returns	-47.317	0.010
ETF log returns	-47.245	0.010

The test statistic and the corresponding P-value of the augmented Dickey-Fuller test for spot, futures, and ETF log prices, log spread, and log returns are reported.

likelihood ratio test, and univariate Ljung-Box Q tests to ensure white noise residuals (Dolatabadi, Nielsen and Xu, 2015).

With the lag length set, we proceeded to test the cointegration rank, which would indicate the stationary equilibrium of linear combinations of prices, signifying a long-term relationship between pairs of time series. The choice of deterministic components, either a restricted or unrestricted constant, hinges on the presence of a trend component. Table 3 outlines the three key elements in our FCVAR model specification: the lag length (k), the deterministic component, and the cointegration rank (r). Our analysis revealed that fractional cointegration might require up to 6 lags to aptly model the data while preserving white noise residuals. This choice aligns with Dolatabadi, Nielsen and Xu (2015), who posited that fewer lags are needed in the autoregressive formulation when fractional integration is incorporated. We determined the rank for our models using a series of LR tests, sequentially testing the null hypothesis of rank $= r$ against the alternative hypothesis of rank $> r$, commencing with $r = 0$. The results suggest a cointegration rank of 1 for all three models, indicating a single stationary cointegrated long-term equilibrium relationship between the estimations of spot-futures, spot-ETF, and futures-ETF prices.

The decision to include appropriate deterministic terms in the FCVAR model is pivotal, especially when the cointegration parameter $b \neq 1$. To determine the necessity of incorporating both restricted and unrestricted constants in our model, we tested the null hypothesis $H_0 : \hat{\rho}, \hat{\xi} = 0$. The results presented in Table 3 suggest that for the spot-futures model, there isn't enough evidence to reject the hypothesis, implying that ρ should be excluded from our FCVAR estimation. However, the data does

Table 3: Model selection

	Spot and Futures	Spot and ETF	Futures and ETF
Selected lag	6	6	6
Cointegrated rank	1	1	1
b	0.560	0.580	0.480
restricted constant	no	yes	yes
unrestricted constant	yes	yes	yes

Model selection for two FCVAR models, spot-futures, spot-ETF, and futures-ETF models.

not support the exclusion of ξ from the model. On the other hand, for the spot-ETF and futures-ETF models, we include both the restricted and unrestricted constant terms in the FCVAR model.

5 Estimation of the FCVAR model

This section focuses on the interpretation of the empirical results of price discovery for spot-futures, spot-ETF, and futures-ETF FCVAR model parameters presented in Tables 4, 5, and 6. One table is provided for each market parity, and the panel organization is the same for both tables.

5.1 Bitcoin spot and futures markets

The first panel of Table 4 presents several hypothesis tests discussed in Section 3 regarding the model estimation for spot and futures markets. The first column pertains to the hypothesis test on whether the CVAR model would better fit the data than the FCVAR model. Based on the high LR statistics of 2693.807 and a P-value of 0.000, the null hypothesis of $H_0 : d = b = 1$ is strongly rejected, providing evidence to suggest that the FCVAR model may be a superior choice for modeling our data in comparison to the CVAR model. The second column of panel A examines the hypothesis test $H_0 : \beta = (1, -1)'$. With an LR value of 41.033 and a P-value of zero, we decisively reject the null hypothesis of $H_0 : \beta = (1, -1)'$.

The third column presents a hypothesis test asserting that price discovery is predominantly driven by the futures market. Using the standard likelihood ratio test with an $LR = 5.121$ and a P-value of 0.024, we reject this hypothesis. Conversely, the fourth column tests the hypothesis that price discovery is solely a function of the spot markets. With an $LR = 11.147$ and a P-value of zero, we reject this hypothesis across all conventional significance levels, indicating that price discovery isn't exclusive to the spot market. Consequently, in the subsequent panel of Table 4, we will estimate the

Table 4: Spot and Futures FCVAR Estimation

Panel A: The Null Hypothesis			
CVAR or FCVAR	**no contango/backwardation**	**Exclusively Futures**	**Exclusively Spot**
$H_0: b = d = 1$	$H_0: \beta = (1, -1)'$	$\hat{\alpha}_{\perp} = [s, f]'$	$\hat{\alpha}_{\perp} = [s, f]'$
$\hat{\alpha}_{\perp} = [0.399, 0.600]$	$\hat{\alpha}_{\perp} = [0.814, 0.186]$	$H_0: \hat{\alpha}_{\perp} = [0, a_{\perp 2}]'$	$H_0: \hat{\alpha}_{\perp} = [a_{\perp 1}, 0]'$
LR = 2693.807	*LR = 41.033*	*LR = 5.121*	*LR = 41.034*
P-value = 0.000	*P-value = 0.000*	*P-value = 0.024*	*P-value = 0.000*
Panel B: Estimated FCVAR Model with Restricted Constant			
$\Delta [s_t, f_t] = \hat{\alpha}\Delta^{1-\hat{b}} L_{\hat{b}} \left(\hat{\beta}' [s_t, f_t] + \hat{\rho}'\right) + \sum_{i=1}^{2} \Gamma_i L_{\hat{b}}^i \Delta [s_t, f_t] + \hat{\varepsilon}_t$			
Parameters Estimation			
$\hat{\alpha} = [-0.007, 0.061]$ $(0.023)(0.027)$	$\hat{\alpha}_{\perp} = [0.891, 0.108]$	$\hat{\beta} = [1, -0.997]$	$\hat{b} = 0.560$ (0.013)
Equilibrium relationship			
$s_t = 0.997 f_t + \nu_t$			

Panel A represents different hypothesis tests related to the spot-futures market relationship, and Panel B provides the parameters estimation of the FCVAR model for spot and futures markets.

FCVAR model's parameters to delve deeper into the contributions of both futures and spot markets to the price discovery mechanism. This will furnish a holistic view of how these markets collaboratively influence Bitcoin prices.

In Panel B of Table 4, we present the estimated adjustment coefficients for the spot-futures market as $\alpha = [-0.007, 0.061]$, with standard errors provided in parentheses. The normalized price discovery estimates are denoted as $\hat{\alpha}\perp = [0.891, 0.108]$. This suggests that the spot market, contributing 89%, plays a predominant role in the price discovery of Bitcoin. Such dominance indicates that the spot market participants are more adept at assimilating new information and forecasting shifts in Bitcoin's value, leading to swifter adjustments in the spot market in response to new information compared to the futures market. Various factors, such as enhanced market participation, superior liquidity, or more efficient access to information in the spot market, might account for this. For context, Panel A depicts the futures market's contribution to price discovery at 60% in the CVAR estimation, whereas the FCVAR model suggests a contribution of 10%. This disparity underscores the potential overestimation of the Bitcoin futures market's price discovery in the non-fractional CVAR model.

The next column of panel B presents the estimate of the cointegration coefficient, $-\hat{\beta}_2 = -0.997$. According to the Efficient Market Hypothesis, as stated by Westerlund and Narayan (2013), spot and futures prices should be cointegrated. Upon rejecting the null hypothesis $H_0 : \beta = (1, -1)'$ and observing the estimated cointegration coefficient $-\hat{\beta}_2 = -0.997$, it's evident that while the coefficient is close to -1, it's not exactly -1. This deviation from unity suggests the presence of contango in the Bitcoin market. Contango is a situation where the futures price of an asset is higher than the expected spot price. In essence, the deviation of $-\hat{\beta}_2 = -0.997$ from -1, even if slight, provides valuable insights into the market dynamics and participants' expectations in the Bitcoin spot and futures markets.

Moreover, the estimate of the fractional parameter $b = 0.56$ in the last column suggests that the integration order $1 - b$ is less than 0.5. In the FCVAR framework, the order of integration for the error correction term is $I(1 - b)$. A larger value of b implies a smaller value of $1 - b$, indicating less memory or lower persistence in the error correction term. This suggests that the markets are more efficient as the order of integration decreases. Given this, the FCVAR model specification is more appropriate for capturing the dynamics between Bitcoin spot and futures markets compared to the non-fractional CVAR model. The estimated equilibrium relationship between spot and futures prices is provided in the last part of Panel B.

5.2 Bitcoin Spot and ETF markets

The initial panel of Table 5 show the results of a series of hypothesis tests, as elaborated in Section 3, tailored for the spot and ETF markets. The LR statistic, standing at a notable 2672.488 coupled with a P-value of 0.000, leads us to firmly reject the null hypothesis $H_0 : d = b = 1$, which posits that the CVAR model might be a better fit of our data than the FCVAR model. The subsequent column in panel A rejects the null hypothesis $H_0 : \beta = (1, -1)'$. The third column in the panel broaches the hypothesis that the ETF market is the primary driver of price discovery. With an LR value of 5.626 and a P-value of 0.131, we fail to reject this hypothesis, hinting at the possibility that the ETF market might be at the forefront of the price discovery process. In contrast, the fourth column tests the proposition that the spot market contributes 100% to price discovery. With an LR value of 11.147 and a P-value of zero, we reject this notion across standard significance thresholds, suggesting that the spot market isn't the sole determinant of price discovery. In the following panel of Table 5, we will delve deeper, estimating the FCVAR model parameters to discern the relative contributions of both the spot and ETF markets to the price discovery process, offering a comprehensive perspective on their combined influence on Bitcoin prices.

In Panel B of Table 4, the adjustment coefficients for the spot-futures market are detailed, with α estimated as $[-0.059, 0.008]$. Standard errors are also included in parentheses. The normalized price discovery metrics are represented by $\hat{\alpha}_{\perp} = [0.078, 0.921]$. From these figures, we conclude that the spot prices contribute 7.8%, while the futures prices have a more substantial contribution of 92.1%. Notably, the ETF market, with a commanding 92% contribution, emerges as the primary force in the Bitcoin price discovery process. This significant contribution underscores the efficiency

Table 5: Spot and ETF FCVAR Estimation

Panel A: The Null Hypothesis			
CVAR or FCVAR	**no contango/backwardation**	**Exclusively Futures**	**Exclusively Spot**
$H_0 : b = d = 1$	$H_0 : \beta = (1, -1)'$	$\hat{\alpha}_\perp = [s, e]'$	$\hat{\alpha}_\perp = [s, e]'$
$\hat{\alpha}_\perp = [-0.020, 1.020]$	$\hat{\alpha}_\perp = [-0.623, 1.623]$	$H_0 : \hat{\alpha}_\perp = [0, a_{\perp 2}]'$	$H_0 : \hat{\alpha}_\perp = [a_{\perp 1}, 0]'$
LR = 2672.488	*LR = 47.438*	*LR = 5.626*	*LR = 11.147*
P-value = 0.000	*P-value = 0.000*	*P-value = 0.131*	*P-value = 0.001*

Panel B: Estimated FCVAR Model with Restricted Constant

$\Delta [s_t, e_t] = \hat{\alpha}\Delta^{1-\hat{b}} L_{\hat{b}} \left(\hat{\beta}' [s_t, e_t] + \hat{\rho}' \right) + \sum_{i=1}^{2} \Gamma_i L_{\hat{b}}^{i} \Delta [s_t, e_t] + \hat{\varepsilon}_t$

Parameters Estimation

$\hat{\alpha} = [-0.059, 0.008]$ $(0.017)(0.019)$	$\hat{\alpha}_\perp = [0.078, 0.921]$	$\hat{\beta} = [1, -0.957]$	$\hat{b} = 0.580$ (0.013)

Equilibrium relationship

$s_t = -0.910 + 0.957 e_t + \nu_t$

Panel A represents different hypothesis tests related to the spot-futures market relationship, and Panel B provides the parameters estimation of the FCVAR model for spot and futures markets.

of the ETF market in integrating new information and anticipating Bitcoin price movements, thereby allowing the ETF market to react more promptly to new information than its spot counterpart. The dominance of the ETF market in the Bitcoin price discovery process can be attributed to several intertwined factors. Firstly, the structure of ETFs inherently offers a more diversified exposure to assets, which might attract a broader range of institutional and retail investors compared to the spot market. Secondly, the ETF market often provides higher liquidity, ensuring smoother and more efficient price adjustments in response to new information. Additionally, the regulatory framework surrounding ETFs might instill greater confidence among investors, fostering a more active trading environment. Moreover, the ease of trading Bitcoin ETFs, akin to trading stocks on traditional exchanges, might appeal to both seasoned traders and newcomers, further enhancing market participation and, consequently, its role in price discovery.

In Panel A, we observe an intriguing discrepancy between the CVAR and FCVAR estimations regarding the ETF market's contribution to price discovery. The difference underscores the potential overestimation inherent in the non-fractional CVAR model when assessing the Bitcoin ETF market's role in price discovery. In Panel B, the subsequent column showcases the estimated cointegration coefficient, $-\hat{\beta}_2 = -0.957$, for the FCVAR model concerning the spot and futures-based ETF. In the final column, the fractional parameter's estimate is given as $b = 0.58$. When analyzing the interplay between Bitcoin's spot and ETF markets, the FCVAR model, with its ability to capture these dynamics, proves to be a more fitting choice than the CVAR model. The concluding section of Panel B offers insights into the equilibrium relationship between the spot and ETF prices.

5.3 Bitcoin futures and ETF markets

Table 6 presents the results of the FCVAR model estimation for futures and ETF markets with a restricted and unrestricted constant term. The hypothesis test results are reported in Panel A of Table 6. We begin by examining the first test, which measures the adequacy of the CVAR model in capturing the price discovery contribution in futures and ETF markets. The very high LR statistic of 2380.834 and zero P-value lead us to reject the null hypothesis thereby concluding that the FCVAR model provides a more precise estimation for price discovery. Furthermore, we reject the null hypothesis of $H_0 : \beta = (1, -1)'$ at all conventional significance levels. In the next panel, we estimate $\hat{\beta}_2$ to examine the nature of the relationship between futures and ETF prices.

After conducting the hypothesis test in the third column of Panel A in Table 6, we fail to reject the null hypothesis that price discovery exclusively occurs in the ETF market, as indicated by the low LR statistic of 0.031 and a P-value greater than all significance levels. Although we fail to reject the notion that price discovery exclusively happens in the ETF market, we will proceed to estimate the FCVAR model to further analyze the price discovery contributions of both ETF and futures markets. Conversely, in the fourth column, we have a high LR statistic of 25.448 and a zero P-value less than all significance levels, indicating that we reject the notion that price discovery exclusively happens in the futures market.

Panel B of Table 6 reports the results of the FCVAR model estimation for futures and ETF markets. The $\hat{b} = 0.480$ parameter suggests that using the FCVAR model can improve the analysis of price discovery. As we observe in the last part of Panel B of Table 6, the fractionally cointegrated equilibrium between futures and ETF shows that the cointegration coefficient, $-\hat{\beta}_2 = -0.938$, is less than one. The cointegration coefficient of 0.938 suggests a long-run relationship between the futures and futures-based ETF prices, where the futures prices are, on average, slightly higher than the ETF prices. This discrepancy could be due to factors such as ETF management fees, trading costs, liquidity differences, and other market inefficiencies.

The estimated speed of adjustment coefficient for futures is -0.133, while for ETFs, it stands at 0.027. After normalizing the price discovery parameters, the contribution proportion of futures is 16.7%, and for ETFs, it is 83.2%. These results highlight the dominant influence of the ETF market on Bitcoin's price, accounting for 83% of the price discovery process. In contrast, only about 17% of the price discovery occurs in the futures market. The observed dominance of the futures-based ETF market over the direct futures Bitcoin market in the price discovery process is noteworthy. The ETF market's contribution of 83% to the price discovery underscores its central role in integrating and

Table 6: Futures and ETF FCVAR Estimation

Panel A: The Null Hypothesis			
CVAR or FCVAR	**no contango/backwardation**	**Exclusively ETF**	**Exclusively Futures**
$H_0 : b = d = 1$	$H_0 : \beta = (1, -1)'$	$\hat{\alpha}_\perp = [f, e]'$	$\hat{\alpha}_\perp = [f, e]'$
$\hat{\alpha}_\perp = [-0.903, 1.093]$	$\hat{\alpha}_\perp = [-0.007, 1.0075]$	$H_0 : \hat{\alpha}_\perp = [0, a_{\perp 2}]'$	$H_0 : \hat{\alpha}_\perp = [a_{\perp 1}, 0]'$
LR = 2380.834	*LR = 58.978*	*LR = 0.031*	*LR = 25.448*
P-value = 0.000	*P-value = 0.000*	*P-value = 0.860*	*P-value = 0.000*

Panel B: Estimated FCVAR Model with Restricted Constant

$\Delta [f_t, e_t] = \hat{\alpha}\Delta^{1-\hat{b}} L_{\hat{b}} \left(\hat{\beta}' [f_t, e_t] + \hat{\rho}'\right) + \sum_{i=1}^{2} \Gamma_i L_{\hat{b}}^{i} \Delta [f_t, e_t] + \hat{\varepsilon}_t$

Parameters Estimation

$\hat{\alpha} = [-0.133, 0.027]$ $(0.034)(0.032)$	$\hat{\alpha}_\perp = [0.167, 0.832]$	$\hat{\beta} = [1, -0.938]$	$\hat{b} = 0.480$ (0.018)

Equilibrium relationship

$f_t = -1.135 + 0.938 e_t + \nu_t$

Panel A represents different hypothesis tests related to the futures-ETF market relationship, and Panel B provides the parameters estimation of the FCVAR model for futures and ETF markets.

reflecting new information in Bitcoin market. This dominance can be attributed to several factors. The ProShares Bitcoin Strategy ETF (BITO) tracks the price of Bitcoin primarily through futures markets. While futures contracts are derivatives that derive their value from the underlying asset, in this case, Bitcoin spot price, the ETF's reliance on these derivative instruments could influence its price discovery process. The ETF's price might be shaped by factors beyond the direct market demand and supply for Bitcoin.

The observed dominance of the ETF market in the price discovery process suggests that the ETF market, despite its derivative nature, might be more adept at assimilating and reflecting new valuation information about Bitcoin than the direct futures market. This could be attributed to the ETF's broader investor base, which might lead to enhanced trading volumes and liquidity. Additionally, the inherent structure of ETFs, which aggregate various assets including futures contracts, might offer a more comprehensive view of market sentiments. This comprehensive view, combined with the diverse investor base, might make the ETF market more responsive to new information, leading to quicker price adjustments compared to the futures market. Furthermore, the futures market might be influenced by transient trading strategies or other market-specific factors, which could render it less influential in the long-term price discovery mechanism of Bitcoin. In contrast, the ETF market, with its diverse composition and broader investor base, might offer a more stable and accurate reflection of the long-term market sentiment and price of Bitcoin. This distinction underscores the pivotal role of the ETF market in the overarching price discovery for Bitcoin, even when it derives its value from futures price.

5.4 The structural break

We employed the Bai-Perron structural breakpoints method, as described by Bai and Perron (2003), to evaluate the robustness of our empirical analysis. This method helps determine if the contribution to price discovery varies across different time periods within our data timeline. In the Bai-Perron breakpoint test, a breakpoint signifies a moment when a structural shift in the analyzed relationship takes place. The robustness checks were conducted using the hourly frequency of Bitcoin prices from October 19, 2021, to December 30, 2022. Five breakpoints were detected in this hourly frequency data, indicating moments of structural change in Bitcoin prices. The specific dates of these breakpoints were ascertained by the Bai-Perron test, which minimizes the Residual Sum of Squares (RSS) and uses the Bayesian Information Criterion (BIC) for selection.

Upon identifying five breakpoints in my model and segmenting the data into five distinct time periods, we explored how the contribution to price discovery of Bitcoin prices evolved over time within each period. This methodology enabled us to trace the progression of the price discovery process and understand the shifting dynamics of the Bitcoin market throughout the study duration. As depicted in Figure 3, from October 19, 2021, Bitcoin's price exhibited a general downward trajectory, interspersed with occasional fluctuations. The initial breakpoint occurred on December 17, 2021. A confluence of factors appears to have influenced the decline in Bitcoin's price during December 2021. Potential U.S. government interventions in regulating digital assets might have instigated the sell-off. The U.S. Federal Reserve's incremental tightening could have further impacted Bitcoin's price decline. Additionally, the surge in Covid Omicron variant cases in the U.S. likely played a significant role in the cryptocurrency price downturn that month.

From the first breakpoint to May 06, 2022, which marks the second breakpoint, Bitcoin underwent a notably volatile phase. On May 6th, the price witnessed a sharp decline, marking the second breakpoint. This decline can be linked to anxieties about escalating inflation and potential hikes in interest rates. This sentiment was fueled by the publication of a report indicating that the US inflation rate had hit a 40-year peak in April 2022. Subsequently, the Federal Reserve hinted at the possibility of elevating interest rates to curb inflation. These apprehensions triggered a widespread sell-off across financial markets, encompassing cryptocurrencies like Bitcoin. In the third period, the prices remained turbulent and saw another abrupt decline. This downturn was ascribed to a mix of elements, including worries about China's intensified regulatory clampdown on cryptocurrencies and a more extensive sell-off in the global markets. Following a period of relative stability interspersed with minor fluctuations, the fourth and fifth breakpoints emerged on August 18th and October 26, 2022, respectively.

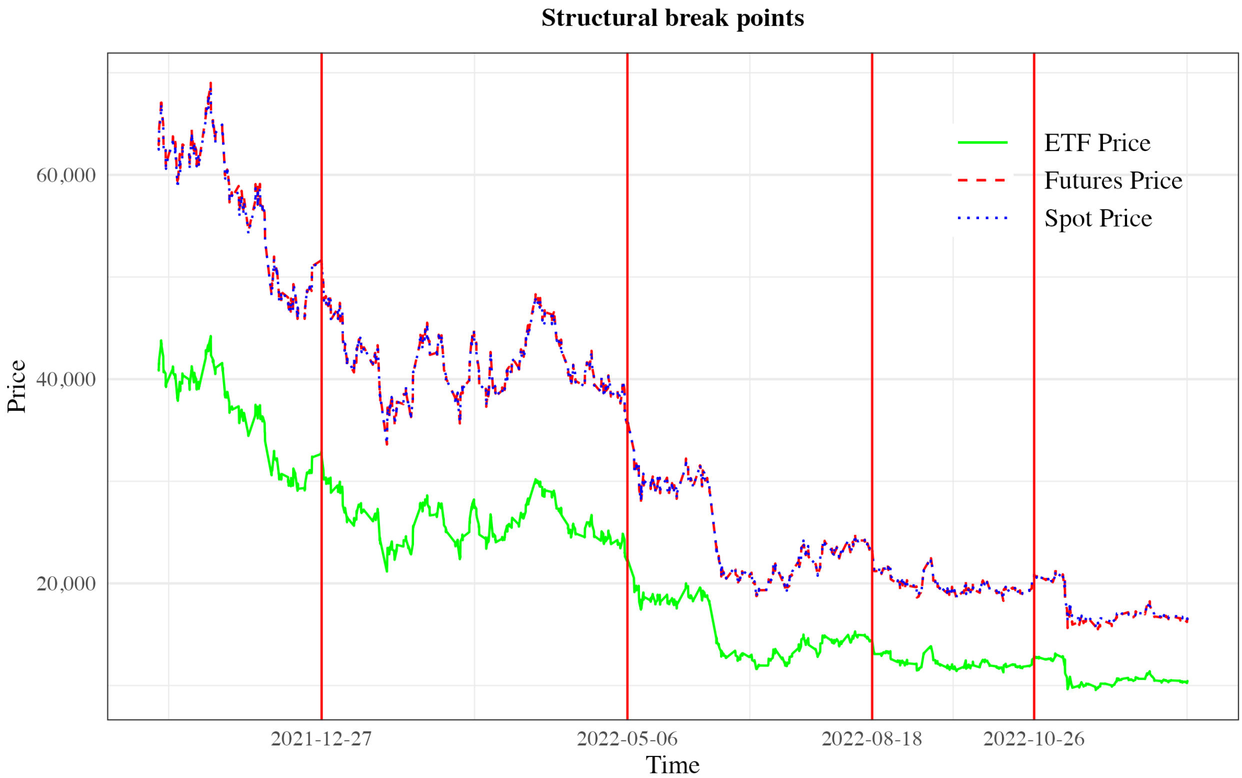

Figure 3: Structural breakpoints for Bitcoin spot, futures, and ETF price series from October 19, 2021, to December 30, 2022.

Starting with the spot-futures price discovery as presented in Table 7, during the initial subperiod, the spot market contributed 73% to the price discovery. However, in the subsequent three subperiods, the futures market played a more dominant role in price discovery than the spot market. Yet, by the final period, the preeminence in price discovery reverted to the spot market. As depicted in Graph 3, the Bitcoin prices during the second and third subperiods exhibited greater volatility compared to other periods, with futures having a more pronounced influence on the price discovery process. Such observations imply that, especially during volatile phases, the futures market might be adept at mirroring anticipations about forthcoming market trends or assimilating external influences, be it macroeconomic updates, regulatory shifts, or technological advancements. The futures markets, often frequented by traders and investors for hedging or speculating on price trajectories, tend to integrate a broader spectrum of information and market expectations into their pricing mechanisms. In times of heightened volatility, the futures market's proficiency in swiftly assimilating and reflecting new data can become especially pivotal, thereby enhancing its contribution to the price discovery mechanism.

Conversely, the results indicate that the Bitcoin spot market exhibits greater efficiency during stable periods. During such times, the prevailing uncertainty might be reduced, allowing market

Table 7: FCVAR Estimation in Different Subperiods

Break point	Sample size	Spot and Futures			Spot and ETF			Futures and ETF		
		$\hat{b}$	$\hat{\alpha}_{\perp}$	$H_0: b=1$	$\hat{b}$	$\hat{\alpha}_{\perp}$	$H_0: b=1$	$\hat{b}$	$\hat{\alpha}_{\perp}$	$H_0: b=1$
1	$15,657$	0.630	$[0.731, 0.268]$	0.000	0.540	$[0.182, 0.817]$	0.000	0.617	$[-1.394, 2.394]$	0.000
2	$37,768$	0.510	$[0.146, 0.853]$	0.000	0.560	$[0.014, 0.987]$	0.000	0.450	$[0.178, 0.821]$	0.000
3	$25,347$	0.445	$[0.220, 0.779]$	0.000	0.550	$[0.192, 0.807]$	0.000	0.400	$[0.325, 0.674]$	0.000
4	$16,853$	0.420	$[0.335, 0.664]$	0.000	0.530	$[-0.300, 1.300]$	0.000	0.470	$[0.296, 0.703]$	0.000
5	$14,225$	0.700	$[0.7575, 0.242]$	0.000	0.701	$[-0.291, 1.291]$	0.000	0.526	$[-1.133, 2.332]$	0.000

FCVAR estimation of spot-futures and futures-ETF markets in six different subperiods based on the Bai-Perron structural test result.

participants to have heightened confidence in the current prices genuinely mirroring the asset's intrinsic value. Under these circumstances, the spot market, which facilitates the immediate exchange of Bitcoin for fiat currency or other cryptocurrencies, adeptly assimilates new information, steering the price discovery process. This enhanced efficiency in tranquil periods can be attributed to the spot market's direct engagement with the trading of the underlying asset. Such direct involvement often results in more precise pricing, as it remains relatively insulated from the speculative undertakings and risk management considerations that characterize futures markets.

Upon testing the hypothesis of $b = 1$, which implies that the CVAR adequately captures the data for price discovery, we reject the null hypothesis across all five subperiods for the spot-futures, spot-ETF, and futures-ETF market pairs. This finding aligns with the study by Wu et al. (2021), which posits that the FCVAR model offers superior insights into the price discovery contribution. The FCVAR model accommodates long-memory dynamics, a feature that seems especially pertinent for cryptocurrency markets, given their inherent high volatility and pronounced persistence. As delineated in Table 7, the fractional parameter b for both spot-ETF and futures-ETF markets remains relatively consistent. However, for the spot-futures market, b is notably higher during the first and last subperiods—intervals where the spot market predominates—compared to the other subperiods marked by futures market dominance. This indicates that the long-memory dynamics between spot and futures were less accentuated during these periods, suggesting a swifter reversion to equilibrium post disturbances or shocks. The diminished long-memory dynamics could be influenced by a myriad of factors, encompassing shifts in market dynamics, investor behavior, regulatory modifications, or external events.

In the subsequent market pairs of Table 7, specifically for the spot-ETF and futures-ETF markets, the ETF market consistently dominates across all subperiods. This dominance suggests that traders and investors might be more inclined to utilize the Bitcoin ETF market to assimilate new information and adjust Bitcoin prices accordingly. The congruence of results across all five subperiods with the overarching period analysis underscores the relative stability of market dynamics across these distinct subperiods. Such consistency further bolsters the argument that the futures-based ETF plays a more pivotal role in the Bitcoin price discovery process than its benchmark, irrespective of the specific temporal segment under scrutiny.

6 Conclusion

In this research, we empirically explore the price discovery contributions across Bitcoin ETF, spot, and futures markets. Using 1-minute intraday data from October 19, 2021, to December 30, 2022, our study augments the work of Wu et al. (2021) by exploring the role of the Bitcoin ETF market in price discovery. By employing the fractionally cointegrated VAR (FCVAR) model, we examine the long-run equilibrium relationships across three distinct Bitcoin market pairs: spot-futures, spot-ETF, and futures-ETF. The FCVAR model is an appropriate tool for studying price discovery mechanisms due to its better fit for Bitcoin data.

Our FCVAR model estimation reveals that in the spot-futures pair a substantial 89% of price discovery is anchored in the spot market, leaving the futures market with a contribution of approximately 11%. In the spot-ETF pair, the ETF market is dominant, accounting for 92% of the price discovery, relegating the spot market to a modest 8% contribution. Similarly, in the futures-ETF pair, the ETF market accounts for 83% of the price discovery, while the futures market contributes the residual 17%. We argue that the dominance of the ETF in the price discovery process can be attributed to several factors. First, ETFs, such as the ProShares Bitcoin ETF, are highly liquid, allowing investors exposure to the underlying index, which contrasts with the potential inaccessibility of futures markets. Second, the well regulated nature of ETFs may attract more institutional and retail investor participation therefore assimilating shifts in sentiment or new information. The ability to respond faster to new information can enhance liquidity and price discovery of the underlying securities.

We conduct a series of robustness checks using the Bai-Perron breakpoint test and verify the ETF markets' preeminence in Bitcoin price discovery across five subperiods. Our robustness tests show that in the spot-ETF and futures-ETF pairs, the ETF is dominant in all subperiods. Furthermore, our analysis discerns a dichotomy in dominance patterns in spot-futures pair: the spot market is dominant during the first and last subperiods which were characterized by relative stability, whereas the futures market is dominant during the intervening three volatile periods. To summarize, our research underscores the pivotal role of the futures-based Bitcoin ETF market in the price discovery process.

7 References

References

Agunloye, Oluokun Kasali and Dahud Kehinde Shangodoyin. 2014. "Lag length specification in Engle-Granger cointegration test: a modified Koyck mean lag approach based on partial correlation." *Statistics in Transition new series* 15(4).

Arslanian, Henri. 2022. Crypto Funds. In *The Book of Crypto*. Springer pp. 351–362.

Bai, Jushan and Pierre Perron. 2003. "Computation and analysis of multiple structural change models." *Journal of Applied Econometrics* 18(1):1–22.

Ben-David, Itzhak, Francesco Franzoni and Rabih Moussawi. 2018. "Do ETFs increase volatility?" *The Journal of Finance* 73(6):2471–2535.

Brown, Marion A. 2019. "Cryptocurrency and Financial Regulation: The SEC's Rejection of Bitcoin-Based ETPs." *NC Banking Inst.* 23:139.

Cárcel, Hector and Luis Gil-Alana. 2017. "A FCVAR model for the Central American economy." *Foro de Investigadores de Bancos Centrales del Consejo Monetario Centroamericano* .

Deville, Laurent. 2008. "Exchange traded funds: History, trading, and research." *Handbook of financial engineering* pp. 67–98.

Dolatabadi, Sepideh, Morten Ørregaard Nielsen and Ke Xu. 2015. "A fractionally cointegrated VAR analysis of price discovery in commodity futures markets." *Journal of Futures Markets* 35(4):339–356.

Dolatabadi, Sepideh, Morten Ørregaard Nielsen and Ke Xu. 2016. "A fractionally cointegrated VAR model with deterministic trends and application to commodity futures markets." *Journal of Empirical Finance* 38:623–639.

Duffy, John, Jean Paul Rabanal and Olga A Rud. 2021. "The impact of ETFs in secondary asset markets: Experimental evidence." *Journal of Economic Behavior & Organization* 188:674–696.

Glosten, Lawrence, Suresh Nallareddy and Yuan Zou. 2021. "ETF activity and informational efficiency of underlying securities." *Management Science* 67(1):22–47.

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

Gutierrez, Carlos Enrique Carrasco, Reinaldo Castro Souza and Osmani Teixeira de Carvalho Guillén. 2009. "Selection of optimal lag length in cointegrated VAR models with weak form of common cyclical features." *Brazilian Review of Econometrics* 29(1):59–78.

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

Hegde, Shantaram P and John B McDermott. 2004. "The market liquidity of DIAMONDS, Q's, and their underlying stocks." *Journal of Banking & Finance* 28(5):1043–1067.

Ivanov, Stoyu I, Frank J Jones and Janis K Zaima. 2013. "Analysis of DJIA, S&P 500, S&P 400, NASDAQ 100 and Russell 2000 ETFs and their influence on price discovery." *Global Finance Journal* 24(3):171–187.

Johansen, Søren. 2008. "A representation theory for a class of vector autoregressive models for fractional processes." *Econometric Theory* 24(3):651–676.

Johansen, Søren and Morten Ørregaard Nielsen. 2012. "Likelihood inference for a fractionally cointegrated vector autoregressive model." *Econometrica* 80(6):2667–2732.

Johansen, Søren and Morten Ørregaard Nielsen. 2016. "The role of initial values in conditional sum-of-squares estimation of nonstationary fractional time series models." *Econometric Theory* 32(5):1095–1139.

Lin, Chu-Bin, Robin K Chou and George HK Wang. 2018. "Investor sentiment and price discovery: Evidence from the pricing dynamics between the futures and spot markets." *Journal of Banking & Finance* 90:17–31.

Madhavan, Ananth and Aleksander Sobczyk. 2016. "Price dynamics and liquidity of exchange-traded funds." *Journal of Investment Management* 14(2):1–17.

O'Hara, Maureen. 2003. "Presidential address: Liquidity and price discovery." *The Journal of Finance* 58(4):1335–1354.

Pan, Kevin and Yao Zeng. 2017. "ETF arbitrage under liquidity mismatch." *Available at SSRN 3723406* .

Richie, Nivine and Jeff Madura. 2007. "Impact of the QQQ on liquidity and risk of the underlying stocks." *The Quarterly Review of Economics and Finance* 47(3):411–421.

Saha, Kunal, Vinodh Madhavan and GR Chandrashekhar. 2022. "Relationship between ETFs and underlying indices: a fractional cointegration approach." *Applied Economics* pp. 1–10.

Shrestha, Keshab, Sheena Philip and Yessy Peranginangin. 2020. "Contributions of Crude Oil Exchange Traded Funds in Price Discovery Process." *American Business Review* 23(2):10.

Singh, Pooja. 2022. "Is the Financial Market ready for Cryptocurrency ETFs?-A critical evaluation." *The Journal of Risk Finance* 23(4):456–460.

Westerlund, Joakim and Paresh Narayan. 2013. "Testing the efficient market hypothesis in conditionally heteroskedastic futures markets." *Journal of Futures Markets* 33(11):1024–1045.

Wu, Jinghong, Ke Xu, Xinwei Zheng and Jian Chen. 2021. "Fractional cointegration in bitcoin spot and futures markets." *Journal of Futures Markets* .

Xu, Ke, Kenneth G Stewart and Zeyang Cao. 2022. "Fractional cointegration and price discovery in Canadian commodities." *The North American Journal of Economics and Finance* 63:101799.