HEC Montreal

Does Trading Volume Impact the Volatility and Returns in Cryptocurrency Markets?

By

Rajat Kumar

Administrative Science

(Finance)

A Thesis Submitted

in Partial Fulfillment of Requirements

for a Master of Science in Administration

(M.Sc.)

September 2020

Sommaire

Cet article vise à analyser l'effet du volume des transactions sur la volatilité et les rendements sur le marché de la crypto-monnaie. En particulier, en utilisant les variations des frais de négociation comme instrument de volume de négociation, nous analysons l'effet causal du volume de négociation sur les rendements et la volatilité. L'étude est menée sur 15 bourses de crypto-monnaie réputées, comprenant environ 2300 paires de crypto-monnaies de janvier 2015 à mai 2020. Les résultats illustrent une relation négative statistiquement significative entre le volume de négociation et les rendements avec les frais de négociation et le volume retardé comme instruments. En outre, nous observons une relation positive statistiquement significative entre le volume des transactions et la volatilité avec les instruments mentionnés ci-dessus. La littérature empirique trouve une relation positive entre le volume des transactions et les rendements, mais elle ne tient pas compte de la variation des frais de négociation, alors que ce document l'inclut. Cependant, la relation de causalité positive observée entre le volume des transactions et la volatilité est conforme à la littérature. Le document mesure l'exactitude des résultats grâce aux contrôles de robustesse et constate que presque tous les résultats découverts sont robustes dans les deux sous-échantillons créés en fonction des données et du type d'échange.

Abstract

This paper analyzes the effect of trading volume on volatility and returns in the cryptocurrency market. In particular, we analyze the causal effect of trading volume on returns and volatility using the changes in trading fees as an instrument. The study is conducted on 15 reputable cryptocurrency exchanges, comprising approximately 2300 crypto-pairs from January 2015 to May 2020. We find a statistically significant negative relationship between the trading volume and returns. Furthermore, we observe a statistically significant positive relationship between trading volume and volatility. We confirm using robustness checks that almost all the results discovered are robust throughout the sub-samples created based on dates and exchange type.

Acknowledgement

Foremost, I would like to express my sincere gratitude to my thesis directors, Professor. Tolga Cenesizoglu and Professor. Vincent Grégoire for their continuous support, precious guidance, and incredible patience throughout the thesis duration. I thank my colleagues for positive discussion and productive criticism. I wish to express my love and appreciation to my family and closest friends for their understanding and constant support during this journey.

List of Tables

List of Figures

Table of Contents

1. Introduction

Cryptocurrency, the best performing asset class of the decade that flaunted a rise of 9,000,000% through bitcoin, has fostered the research to examine such an asset class's intricacies. Cryptocurrency, primarily developed as a payment method, is now mostly used as a speculative asset, wherein people try to benefit from its appreciation. With its blockchain technology, cryptocurrency has attracted worldwide recognition but still holds a shadow of a doubt by the vast number of (prospective) traders, as it is arduous to ascertain the real value and know as to what drives its prices. Upon seeing a massive upward rally in the cryptocurrency domain, researchers came forward to analyze the causes of such returns and volatility in the market. Having no underlying asset, cryptocurrencies domain of research initially focused on the web search (google searches) and the trading volume as the potential causes of the surge in the market's returns and volatility. Nasir et al. (2019) study the causal relationship between the number of google searches and returns. This study finds a positive relationship between the frequency of bitcoin searches and returns and bitcoin volume. Researchers have also explored the possibility of volume as a potential cause of the fluctuations in returns and volatility. A paper by Bouri et al. (2019) illustrates that volume Granger causes the returns in the seven cryptocurrencies studied and volatility in three of the seven cryptocurrencies studied. However, the scope of the above study is narrow and only focuses on seven cryptocurrencies. Hence, the paper investigates the same relationship using changes in trading fees as an exogenous shock to volume over 2300 crypto-pairs across 15 exchanges.

The paper aims to unearth a causal relationship between the volume, returns, and volatility in the trading fee's presence. In other words, the paper uses the change in trading fees as an instrument for the trading volume, thereby analyzing the causal effect of trading volume on returns and volatility. To study this causal effect, we use the Instrumental Variable Estimation's Two-Stage Least Square method, which is used to control for regressor's endogeneity, i.e., when the regressor is correlated with the error term. An efficient instrument is included in the model to test a causal relationship between the independent and dependent variables. To check for a robust and efficient instrument for a given regressor, we perform the Durbin Wu-Hausman test. The trading fee acts as an instrument in our scenario to establish the relationship between the volume, returns, and volatility. The volume is an endogenous variable, whereas log-return and log volatility are the dependent variables. Using the trading fee as an instrument, we evaluate the change in the trading fee's impact on the volume and further analyze that volume's influence on the returns and volatility. The paper also highlights the general relationship between the fee, volume, returns, and volatility by computing the Pearson Correlation coefficient and OLS regressions.

An extensive database of 15 cryptocurrency exchanges with approximately 2300 crypto-pairs' daily observations has been acquired from Kaiko. Another significant aspect of the data is the historical trading fees collected using the Wayback machine and the exchanges' website. As cryptocurrency is a relatively new asset class and there is a lack of trusted data, that is why the paper uses the data from 15 reputable exchanges that spans over five years from 1st January 2015 to 31st May 2020. The paper also incorporates the Total Fee specification, which aggregates the Maker's fee and Takers' fee. Makers are traders that provide liquidity using the limit order. They make a profit through relatively small spreads. Since the makers inject the liquidity into the market, most exchanges offer them rebates. On the other hand, takers withdraw liquidity from the market, so they pay the premium in the form of higher trading fees. According to the theory, the total fees should exogenously affect the relationship between the volume, returns, and volatility as it will summarize the final impact of fee on the factors mentioned above. Nevertheless, the paper still analyzes and portrays the result for the takers' fee structure and makers' fee structure, so that their respective reactions to the fee change can be observed.

First, we analyze our empirical results in the light of our principal instrument, i.e., the Total fee structure, which studies the cumulated effect of Makers and Takers trading fees. We observe a statistically negative relationship between the trading volume and the returns with the Total trading fee as an instrument. The relationship mentioned above contrasts with the existing literature, which finds a positive relationship between the volume and the returns. However, the literature does not account for fee structure, and the paper includes the same. The negative affiliation between volume and returns suggest the increase in returns through volume could have been neutralized by the trading fee, thereby marking a negative relationship. There may be other possible reasons for this negative relationship between the trading volume and returns, as the cryptocurrency domain has not been fully explored. The paper also unearths a statistically significant positive relationship between the trading volume and volatility with the total trading fee as an instrument. The positive relationship signifies that an increase in volume drives the volatility up. We obtain consistent outcomes in the relationship after the inclusion of lagged volume as an instrument. However, the coefficients estimates decrease upon the inclusion of lagged volume in the model, as the lagged volume reduces the variance from the volume's past values. The results are coherent with Lee and Rui (2002), who find that volume does not Granger cause the returns but discover a positive relationship between the trading volume and volatility in the stock market. Hence, the results of the study are significant and vital in the domain of the cryptocurrency market.

Second, we analyze the empirical results in the presence of the Maker and Taker fee structure. The paper discovers a positive and less significant (at alpha = 0.10) relationship between trading volume and returns with the Taker fee structure as an instrument. In contrast, the paper finds a negative relationship between the trading volume and returns with the Maker fee structure as an instrument. However, upon the inclusion of lagged volume as an additional instrument in the existing model, we notice a negative relationship for both the fee structures. The paper illustrated a positive and statistically significant relationship between the trading volume and volatility for both the Maker's fee structure and Taker's fee structure as instruments. The coefficients estimates decrease upon the inclusion of lagged volume in the model, as the lagged volume reduces the variance from the volume's past values but remains positive for both the fee structures. The magnitude of the positive relationship between the trading volume and volatility is greater for the takers' fee structure than the makers' fee structure, which showcases that the takers tend to escalate the volatility more than the makers. In other words, it seems that the makers have a comparatively stable pattern of trading through the usage of the limit order book, whereas takers are more influenced by the trading volume and the fee changes.

To ensure the accuracy of the results, we perform a set of robustness tests. We divide the data into two dimensions: (1) based on exchange type and (2) based on date. The two types of exchanges on which the data is divided are fiat exchanges and crypto-only exchanges. The paper computes the 2SLS for log return on log volume and log-volatility on log volume (with trading fees and lagged volume as instruments) for both exchange types. The first sub-sample results demonstrate consistency with our study's primary results, as they yield a negative relationship between the volume and returns and a positive relationship between the volume and volatility for both types of exchanges. For the second robust check, the data was divided based on the date. We choose as our split date the 15th of December 2017, when the bitcoin reached an exceptional level of $20,000, we saw a rough drop in the value, losing 80% of worth by September 2018 from its peak value. The date divides the data into a standard period and abnormal period, which is an extreme measure for a robustness check. The second sub-sample's result of the 2SLS for log return on log volume was consistent with our main result. However, the second sub-sample's result of 2SLS for log volatility on log volume was not consistent with our main results after the data sample. The data post the chosen date demonstrated a negative relationship between the volume and volatility, maybe because the volume traded shrunk significantly, driving the price down and increasing market volatility. Overall, we found that the main results are coherent with the robust checks, except for an exceptional situation.

The rest of the paper is as follows. Chapter 2 explores the existing literature about the relationship between fees, volume, returns, and volatility in various financial markets. Chapter 3 describes the data to be used for the analysis process. Chapter 4 portrays the methodology, which is the Instrumental Variable estimation method. Chapter 5 demonstrates the results of the Pearson correlation, OLS regression, Durbin Wu-Hausman, and 2SLS. Chapter 6 presents a series of robust checks, while chapter 7 concludes the studies.

2. **Literature Review**

Despite the extensive study, the relationship between trading volume, return, and volatility has been a prominent subject matter for the scholars so far. Most importantly, the recent inclusion of the volume in the Capital Asset Pricing Model (CAPM) illustrates the significance of the volume. Acharya and Pedersen (2005) derives CAPM using liquidity, a notion of the trading volume, and finds that the required return relies on the expected liquidity and asset's return and liquidity and the market's return and liquidity. The paper also illustrates that a constant negative shock in liquidity (impact on volume) not only affected the market returns but also affected the high expected returns. Upon knowing the importance of the volume in the financial markets, researchers strive hard to unearth the factors that impact the volume and examine the substantial volume's effects on the market's returns and volatility. There has been ample amount of research in the equity market (Gebka and Wohar, 2013, Karpoff, 1987, Chen et al., 2005, Todorova and Souček, 2014.), bonds (Balduzzi et al., 2001), and commodities (Chen et al., 2005) about the relation above but not a lot in the cryptocurrency market.

Blume et al., 1994, were amongst the first researchers to illustrate that the trading volume as an indicator of absorbed market information disclosed by the price, thereby providing a theoretical explanation for the wide use of volume in forecasting future stock returns. Todorova and Souček (2014) examines 26 German stocks' intraday data to establish a relationship between trading volume, overnight returns, and volatility forecasting. The results illustrate that the liquidity factors help in little enhancements in forecasting operation, consistent with Blume et al. (1994). Lee and Swaminathan (2002) also states that the volume predicts long-term market momentum. A study on the same line by Maheshwari and Dhankar (2017) in Indian stock markets analyzes the portfolio of massive trading volume stocks and another portfolio of low trading volume stocks and discovers that the massive trading volume stocks portfolio earned higher returns and showcases continuous momentum as compared to the low trading volume stock. Overall the study concludes that the volume-based momentum strategies and volume-based constrain strategy outperformed the average momentum and constrained based strategies.

Gebka and Wohar (2013) analyzes the relation between the past trading volume and index returns of the Pacific Basin countries and finds no linear relationship between the volume traded and returns. However, there is non-linear positive volume-return causality for high return quantiles and negative for lower ones. The above results are consistent with Chen et al., 2005, who uses Granger causality tests to examine volume-return relation. The paper indicates a positive correlation between trading volume and the absolute value of stock prices changes, but no significant volume-volatility relation. The finding of Chen et al., 2001 is similar to that of Harris and Raviv, 1993, as they also find absolute price changes and volume are positively correlated. Louhichi, 2011, a noteworthy paper, sheds light on the relationship between the trading volume and the volatility. The paper not only discovers a positive and significant relationship between the two but also unearths a positive relationship after controlling for the autocorrelations in the data (controlling the impact of the intraday patterns).

The study of volume-volatility is not only restricted to the equity market but also the Futures & Options (F&O) market. Chiarella et al. (2015) utilize a continuous-time multi-factor stochastic volatility model to examine the possible relationship between the returns and volatility in the commodity futures market. The paper unearths a positive relationship between the gold futures market and negative for the crude oil futures, mainly driven by market-wide shocks in the high volatility period. It illustrates a consistent relation between the returns and volatility, as the positive relationship is detected between the volatility and returns in the gold's future market. An important paper by Sarwar (2007) discovers that the past call options and put options volumes of S&P 500 futures options have a strong predictive power to estimate the market's potential price volatility.

Cheng. F. Lee (2000) argues the consistency of the volume-return causality effect. The paper scrutinizes the causal relationship between the trading volume, stock returns, and return volatility in Chinese exchanges and also across the other various markets. It illustrates that the trading volume does not Granger cause the market returns and finds a weak predictive power of the US and Hongkong financial market for the Chinese markets. Therefore, the paper concludes no relationship between the volume, returns, and volatility at the exchange and cross-exchange levels. F. Douglas Foster et al. (1993) have a similar conclusion about the relationship between the volume, return, and volatility. The paper finds that the volume and returns are not high when the trading cost is low. Instead, it discovers that the volume and volatility are higher for the actively traded firms in the first half-hour of the day. Other researchers who share the common understanding of the complex volume, return and volatility relationship are Campbell et al., 1993 and Llorente et al., 2002. They discuss the convoluted relation of volume and returns and demonstrate that the relation between them is not a simple linear relationship, as the volume is influenced by numerous price movements and traders' motives. The paper by Lee and Rui, 2002, who study the three largest stock markets in the world, namely New York, Tokyo, and London, also find that the trading volume causes no granger causality on the stock market returns but do find a positive relationship between the trading volume and stock return volatility.

Researchers are diligently examining the factors to broaden the horizons of the factors affecting trading volume, thereby influencing the returns and volatility. The two significant factors studied in the literature so far are (1) economic announcement effect (2) change in transaction cost. Both the factors have shown its effect on the volume that have been reported in the empirical research. Balduzzi (1998), Kouwenberg et al. (2008), and Michael J. Barclay (1998) show evidence that factors described above can affect volume, thereby affecting the returns up to a certain extent. Kouwenberg et al. (2008) examine the effect of earning announcements on trading volume, volatility, and spreads in the options market. They conclude that the options market's trading volume responds faster and more robust in earning announcements than stock volume.

Balduzzi (1998) examines government bond data to explore the effects of economic announcements on the prices, trading volume, and bid-ask spread. Ten out of the seventeen economic announcements observed significantly impacted the observable prices within one minute of the announcement. Also, for a 10-year bill, he finds a significant relationship between the announcement and the volume and saw a surge of 1.7 times of average volume as a consequence of the announcement, but results are not consistent for short term bonds, which illustrated a weak relation between the announcement and the volume. Following the 15 minutes after the announcement, the long-term bonds' volume also comes to normal. The result mentioned above about the relationship between the economic announcement and volume is consistent with Fleming and Remolona (1999), who demonstrate a significant effect on the volume for 30 minutes following the economic announcement.

Out of many factors, one significant factor affecting volume, according to empirical research, is the transaction cost. Hence, it is eminent to note the three types of transaction costs, which are:

1. Bid-Ask Spread, which is the difference between the asking price and bid price
2. Broker fee (commissions)
3. The market fluctuation of the asset price

Our focus in the paper is concerned with the Broker fees (commissions), i.e., the trading fee. The exchanges nowadays adopt complicated fee structure, which involves the maker and taker fee. In other words, it implies there are differential fees charged from the liquidity supplier (Maker) and liquidity demander (Taker). Usually, there is a high fee for the liquidity demanded (Taker), as it is a fee charged from the trader to provide him for immediate liquidation. Some stock market exchanges have had adopted an inverted fee structure, which provides rebates to the takers rather than makers. There have been empirical researches in the domain of the complicated maker/taker fee concept and the spread domain.

Michael J. Barclay (1998) studies the impact of transaction costs on the trading volume and stock prices. The paper keeps the bid-ask spread as the proxy for the transaction/trading cost. According to theory, the paper anticipates that the transaction cost increase will lead to an increase in investors' expected return and the average holding period to amortize the high transaction cost. Hence, the results are consistent with the theory, as it highlights the increase in transaction cost significantly affects the volume traded, but there is no evidence on its effect on the prices. Kramer (1999) develops an economic model of rational trader that incorporates the transaction cost and noise trading as variables, and reports that the level of trading affects the traders' marginal cost of the transaction. The paper explicitly states that the trading volume is a source of risk in the light of the marginal cost. This mechanism causes the equilibrium association between volume and returns.

Amihud & Mendelson (1986) portray that expected returns are highly sensitive to the transaction cost. On the contrary, Vayanos (1998) and Constantinides (1998) demonstrates that the principal factors affected by transaction costs are the holding period of security and the trading volume, whereas the returns have second-order effects. Michaely and Vila (1996) also test the relation between the investors' heterogeneity, risk, transaction costs, and trading volume. They find that an increase in risk or transaction cost cuts the volume, consistent with the results above, but does not comment about the returns. The paper also concludes that unsystematic risk is an essential factor in determining the volume of trade. Kadlec and McConnell (1994) find a likewise result, where the effect of transaction cost is little on the expected returns but significant on the volume. The above conclusion is similar to the Admati and Pfleiderer (1988), who discover that the volume and return volatility are high when there is low trading cost. A paper by Cepoi (2014) studies the effect of transaction cost on the intra-day trading data of the Bucharest Stock Exchange and later determine that transaction cost is a significant factor affecting the market liquidity and asset returns. The paper marks the similarity with Michael J. Barclay (1998) and states that the rational investor demands compensation when there is high transaction cost, so they will tend to hold the asset and expect higher returns.

Different types of fees often create confusion in the traders' minds, which leads to the mismanagement of the cost and the risk associated. Angel et al. (2006) portray that the "Maker or Taker" fee structure causes inaccuracy in the market that should be rectified. The fee/rebate is not revealed in the quotation, which is often a determinant of order execution that causes market-wide inefficiency in the risk management system. The study mentions above is affirmative to the article presented by Harris (2013), which discusses the intricacy of the Maker/Taker fees that inaccurately quotes the false and less informative quotations, thereby not taking account of the fee rebate for the liquidity takers and providers.

On the other hand, Foucault (2013) reports that that discrete tick size is already adjusted to neutralize the impact of fees and rebates. This result is consistent with Chao et al. (2017), highlighting the competitive tool's significance for exchanges. The illustrated model in the paper represents the modernization in the discrete tick size, and they also reason that the usage of the maker/taker fees helps fulfill the need for different kinds of trading participants' needs. Another paper by Comorton-Forde, Grégoire, and Zhong (2019), is consistent with the argument above and finds that sub-tick price rendered by the inverted venues strengthens liquidity provision and adds information to the prices, thereby helping the traders to make a decision.

There have been many well-known research papers written in the equity market, bonds, and futures market, but less emphasis is given to the cryptocurrency market. Cryptocurrency has become a widespread new class of investments. It is a novel technique to exchange value through the internet without the need for any intermediary. It has attracted complete acknowledgment from the financial world due to its massive upsurge in transactions and market capitalization. The cryptocurrency market almost touched $1 trillion in the 2017/2018 bubble/bust cycle that currently hovers around $216 billion presently (as on 26th April 2020), according to the CoinMarketCap. Amongst all the currencies, the most dominant is bitcoin (64% dominance), commencing its commendable traction around the financial crisis (2008) and the European sovereign debt crisis (2010-13), when the governments and central banks failed to gear up with the existing situations.

While the governments control the currencies, cryptocurrencies are decentralized that use cryptography for secure transactions and preclude destructive actions that can damage the system. The transactions are stored in the digital form as a unique algorithm in a digital ledger, commonly known as Blockchain. Although the transactions have unique algorithms to prevent counterfeiting, there is still a threat of theft through the system that happened with the world's largest bitcoin exchange, Mt. Gox exchange, in February 2014. Irrespective of illicit activities, cryptocurrencies have observed a significant increase in returns (such as Ethereum, Litecoin, Ripple, and many more) and volume traded since 2017. Some cryptocurrencies, such as Ethereum and Ripple, saw an exponential increase in trading volume.

In addition to all the facts about the crypto market, researchers debate its financial and economic aspects. Kristoufek (2015) illustrates bitcoin as an exclusive asset, having the properties of a standard financial asset and a speculative asset as well. On the other hand, the book by Popper (2015) portrays bitcoin as a digital form of gold. Also, Bouri et al. (2017) highlight bitcoin's significant features as an investment asset. Yermack's (2013) finding is partially consistent with Kristoufek (2014) and mentions Bitcoin as a speculative investment tool and not a currency as its market capitalization does not match the actual transactions.

However, Glaser et al. (2014) argue that digital currencies are not primarily an alternative currency but more of an investable instrument. Consistent with the Glaser et al. (2014), Hanley (2013) suggests that digital currency is distinctive from actual mainstream currencies. In an attempt to uncover the value of bitcoin, Hayes (2006) finds that bitcoin mining adds to bitcoin's value.

A significant amount of work has been done to determine the factors driving bitcoin prices and volume. An important paper by Kristoufek (2013) reports a strong bidirectional causality effect between the bitcoin and its search on Wikipedia and google trends. Bouoiyour and Selmi (2015) also shows a consistent result and quoted that bitcoin prices can be explained by lagged Google searches, and not by the data of transactions. They sum up bitcoin as speculative foolery that is far from long term assets. Polasik et al. (2015) also find that bitcoin returns are primarily driven by the popularity, i.e., by the opinions in the newspaper and the internet. Kristoufek (2014) derives that the trade-exchange ratio plays a vital role in fluctuating bitcoin prices. The paper Bouoiyour and Selmi (2015) also confirms that the bitcoin price Granger causes the trade-exchange ratio. A paper by Ciaian et al. (2016) demonstrates the economics behind bitcoin's price formation, which suggests an increase in demand due to its popularity and time-consuming supply (lengthy mining process), causing the prices to increase. Nevertheless, the above literature presents a partial picture of the trading volume and return and does not consider other factors other than its popularity in the cryptocurrency market.

Upon observing bitcoin's history, the price-volume relationship is deduced as when the price plunged, a surge in volume is seen. Such events happened on 19th November 2013, when the price of bitcoin plunged by approximately 20%, a surge in the trading volume in bitcoins (71,560) was seen, another on 7th December 2013, when the price plunged by 15%, a new high volume of 79,852 coins were recorded, and on 18th December 2013, when the price plunged by another 23% and a new high volume of 137,070 was observed. The evidence mentioned above advocates a strong relationship between the magnitude of price movements and volume traded. So, to determine whether the volume has its effect on returns and volatility, Balcilar (2017) came up with the same.

Balcilar (2017) discusses the causal relation between trading volume and bitcoin returns and volatility over its conditional distributions. The paper shows that the quantiles test's causality shows that the volume can determine the returns except in the two-extreme scenario, e.g., in a bearish and bullish market. This paper also captures the causal relation at the tails and the help of the causality-in-quantiles test. However, there is no relationship between the trading volume and the volatility at any point in the conditional distribution. The paper by Balcilar (2017), however, suffers from two critical disadvantages (1) neglect copula-based dependence, that allows decomposing a joint probability distribution into marginals and a function that couples them together, thereby giving separate correlation, (2) concentration on bitcoin, and overlooking other major cryptocurrencies such as Ethereum, Litecoin, and Ripple. To overcome the shortcomings of the Balcilar (2017) paper, Bouri et al. (2019) highlights both of the above ideas in its paper.

In the paper, Bouri et al. (2019) uses the copula approach, powerful in modeling tail dependence, and also captured all the major cryptocurrencies such as Bitcoin, Litecoin, Ripple, Ethereum, Nem, Dash, and Stellar. The paper is also crucial because it captures the generic market reaction of 7 top traded cryptocurrencies and not only of bitcoin, whose dominance in the market is diminishing. The bitcoin dominance plunged from 95% to approximately (in April 2013) to 63% approx. (in April 2020), touching a low of 38% approximately in June 2017. The paper's empirical results highlight the volume granger cause the returns of all the seven cryptocurrencies in bullish (upper quantile) and bearish (lower quantile) market phases. However, trading volume had no granger cause on volatility except for Litecoin, Nem, Dash in very low volatility scenario, that too when squared returns are used as a proxy for volatility and not GARCH volatility.

The literature in cryptocurrency has not been still fully explored when it comes to what causes the fluctuations in volume, thereby affecting returns and volatility. Upon observing the vast number of papers in equity, commodity, and bond markets, it breeds the inquisitiveness to discover some useful finding in the cryptocurrency market and contribute to its literature. To broaden the perspective about the crypto-market, the paper initially evaluates the existence of a general relationship between the change in trading fee, volume, returns, and volatility and then will analyze the effect of trading fee on volume, thereby investigating the resultant volume's effect on the expected returns and volatility in the crypto market. The paper plans to keep the transaction cost an instrumental variable in understanding the transaction cost's shock on volume and volume's impact on the crypto market's returns and volatility. The study aims to build on the existing literature about the relationship between the volume, returns, and volatility in the presence of trading fees in the cryptocurrency market.

3. Data

Lack of reliable data has often constrained study in the cryptocurrency domain. There have been significant and trustworthy sources who drafted such allegations against the cryptocurrency markets, such as by Bitwise Inc., who published a report in March 2019 and trusted news channels such as CNBC, Forbes, and Blackrock. The report explicitly stated that the false and unregistered exchanges fake 95% of the bitcoin trading volume. The company also estimated only $273 million of legitimate transactions out of the $6 billion stated in the CoinMarketCap.com. The reports warn researchers about the data inaccuracy, which is one reason that restricts the study scope. The paper examines the trusted and the registered exchanges and analyses two types of data, (1) Fee structure, which is directly collected from the reputed exchanges' websites, and (2) the exchange wise intraday trading data on the cryptocurrency pairs, which is acquired from Kaiko, a digital assets data provider.

3.1 Fee Structure

The first part of the data captures the fee structures, promotional schemes, and fee changes for all exchanges from January 2015 to May 2020. The data is manually collected from the respective websites of the exchanges. The most trusted direct source of information is the exchange websites, but it shows only the current fee structure. To overcome this hurdle, we used the Wayback machine; it is a form of a free digital library for scholars, researchers, and historians, who can find old archived internet pages. Similarly, in order to capture the last fee changes, the Wayback machine was used. However, the Wayback machine has a limitation that it does not have archived web pages for all the dates, so by seeing the archived page, one can only compare the old archived page and observe a change but cannot determine the exact date of fee change or promotional scheme. The 'News' or 'Announcements' section on the exchanges' websites was studied thoroughly to resolve this issue. Among the many exchanges that were researched, comprehensive information for fifteen reputable exchanges was found. This paper judiciously examines fifteen exchanges and evaluates the fee's impact on the volume, returns, and volatility in the respective exchanges.

Every exchange fee structure has different fee schedules based on the trading volume, i.e., the high trading fee for the less volume traded and low fee for high volume. The trading fee is bifurcated into three parts: High trading volume fee, Median trading volume fee, and Low trading volume fee. The data expands from 1st January 2015 or from the date of commencement of the exchange (whichever is early) to 31st May 2020.

Table 1: Summary Statistics of the Maker Fees

Particulars	Maker		
	Low volume Trading Fee	Median volume Trading Fee	High Volume Trading Fee
Mean Fee	0.1404	0.0731	0.0397
Maximum Fee	1 (Gemini)	0.34 (Bitstamp)	0.25 (Bittrex)
Minimum Fee	-0.1 (HitBTC)	-0.1 (HitBTC)	-0.1 (HitBTC)
Standard Deviation Fee	0.036	0.0166	0.0171

Table 2: Summary Statistics of the Taker Fees

Particulars	Taker		
	Low volume Trading Fee	Median volume Trading Fee	High Volume Trading Fee
Mean Fee	0.1981	0.1486	0.0912
Maximum Fee	1 (Gemini)	0.375(Gemini)	0.25 (Bittrex)
Minimum Fee	0 (Bitstamp, Kraken, BTCbox)	0 (Bitstamp, Kraken, BTCbox)	0 (Bitstamp, Kraken, BTCbox, Korbit)
Standard Deviation Fee	0.0378	0.0185	0.0183

Before the paper discusses the descriptive statistics of fee structure, it is imperative to note that the 'maker fees' will always be less than the 'taker fees' as 'maker' infuses the liquidity in the market through increasing the volume using limit order book, so he is rewarded, whereas the 'taker' withdraws the liquidity from the market, so he pays a premium for it. Following the theoretical explanation of the maker and taker fee structure, the paper finds that the maker fee is less than the taker fee for all the exchanges. The 'maker' mean fee ranges from -0.06% to 0.22%, where the negative fee represents the rebate for the market maker for inducing the liquidity in the market. The exchanges with the minimum maker fee over the five years are HitBTC (-0.1) and OKCoin (-0.01), while almost all other exchanges were having a minimum maker fee of 0%. The maximum maker fee over the five years is 1% charged by Gemini for low volume traders.

Following the statement above of high taker fees, the paper discovers the maximum mean fee of 0.28% by Gemini and 0.23% by Bitterex. The maximum trading fee was recorded over five years by Gemini (1%) from May 2018 to June 2019. Whereas the minimum fee recorded was 0% by CEX.IO, Poloniex, Korbit, Kraken, BTCbox, and Bitstamp, as these are exchanges experiencing high trade volumes, so they offer a wide range of promotional schemes from time to time to every type of trader (low volume or high volume). Gemini and Kraken exchange's probable reason for the high trading fee is that the exchanges are among the Top 7 Crypto Exchanges by Coingecko in terms of reputation, security, and liquidity. Upon counting the total fee changes in 5 years of all the exchanges, it was observed that more than 50 % of exchanges changed fee only once or twice, and exchanges such as CEX.IO, Korbit, and Poloniex changed fee 3-5 times.

3.2 Volume and Prices (Data from Kaiko)

The other part of the data is directly acquired by Kaiko, a real-time and historical cryptocurrency data provider. The company provided daily data for all the trading pairs traded in the 15 exchanges. The data comprises the numerous cryptocurrency pairs traded in the respective exchanges with open price, high price, low price, close price, and volume traded for the corresponding date. The preliminary step for the analysis part was to merge the trading fee data with the data from Kaiko and form a full-sized panel data set. In total, the data set had approximately 2,300 different crypto pairs across all the 15 exchanges. After removing the cryptocurrency pairs with fewer (discussed below) observations, the paper studies 2,120 crypto pairs across all the 15 exchanges.

Some filters are implemented onto the primary panel data from Kaiko to generate the final data set. The measures applied (discussed below) are to reduce the market microstructure concerns and eliminate the outliers' impact on the entire data set. The following are the measures:

1. The cryptocurrency pairs with less than ten observations are disregarded to reduce the less traded pairs' noise, as the less traded and inexpensive pairs tend to have high volatility and do not portray the accurate picture of the crypto market.
2. The observations below five percentiles and above 95 percentiles are also excluded from the primary data frame to safeguard the study from outliers driving out the results.

To analyze the relationship between the change in trading fee, trading volume, returns, and volatility, we need to compute specific parameters such as return, log volatility, lagged volume, the total fee for low volume traders, the total fee for median volume traders, the total fee for high volume traders and total fee for average volume traders. The following are some computations:

i. Return is approximated by:

$$r_t = \log\left[\frac{Cl_t}{Cl_{t-1}}\right]$$

where Cl_t = closing price of Crypto pair at time ‘t’

Cl_{t-1} = closing price of Crypto pair at time ‘t-1’

ii. Volatility is approximated by: (square of the log returns)

$$v_t = \left[\log\left[\frac{Cl_t}{Cl_{t-1}}\right]\right]^2$$

iii. Total Low Trading Volume Fee = Maker’s Low Trading Volume Fee + Taker’s Low Trading Volume Fee

iv. Total Median Trading Volume Fee = Maker’s Median Trading Volume Fee + Taker’s Median Trading Volume Fee

v. Total High Trading Volume Fee = Maker’s High Trading Volume Fee + Taker’s High Trading Volume Fee

Out of all the 15 exchanges, the most prominent exchanges in volume trading and the number of crypto pairs trading are HitBTC, Bittrex, Binance, Poloniex, and Bitfinex. The massive volume on these exchanges is backed by the minimum trading fee charges, promotional schemes, and the security level. HitBTC has had the minimum trading fee among all exchanges

(Maker= -0.1% and Taker= 0%), and Poloniex introduced a lot of promotional schemes (5 times fee change) and even scrapped off the taker fee under the promotional scheme for some time. After the fall of Mt.Gox in 2014 that handled almost 70% of the crypto trade volume, the investors became quite vigilant about the security, reputation, and liquidity condition of the exchange, that is the one of the primary reason for Binance to become one of most used and trust-worthy exchange. The following table presents some crypto-pairs with their descriptive statistics.

Table 3: Descriptive Statistics various Crypto-Pairs

No.	Particulars	*Mean Volatility*	*Mean log return*	*Standard deviation in log return*
	Five Most Traded Cryptocurrencies			
1	**BTC/USD**	0.0190	-0.0012	0.1378
2	**ETH/BTC**	0.0157	0.0006	0.1253
3	**LTC/BTC**	0.0063	0.0001	0.0794
4	**XRP/BTC**	0.0370	0.0005	0.1505
5	**LTC/USD**	0.0370	-0.0014	0.1925
	Five Medium Traded Cryptocurrencies			
1	**PXC/ETH**	0.006	0.00004	0.077
2	**TRUE/BTC**	0.0319	-0.0009	0.179
3	**WOC/ETH**	0.0428	0.00005	0.207
4	**REBL/ETH**	0.0114	0.0001	0.107
5	**MFT/BTC**	0.0032	-0.0118	0.056
	Five Least Traded Cryptocurrencies			
1	**RBT/BTC**	0.00003	0	0.008
2	**UBTC/BTC**	0.136	0.184	0.369
3	**MOL/ETH**	0.0009	-0.008	0.032
4	**REP/JPY**	0.097	-0.218	0.239
5	**ZEC/BCH**	0.126	-0.087	0.365

Note: The table represents the mean volatility, mean of log return and standard deviation of returns of the five most traded pairs, medium traded pairs and least traded pairs according to the observations in the dataset.

Table 3 illustrates the most traded, medium traded, and least traded crypto-pairs in the market. The table shows that the least traded cryptocurrency pairs have the highest mean volatility, followed by the medium traded and most traded crypto pairs. Least traded pairs have higher volatility because they are uncertain and inexpensive pairs, so even a small change in volume can significantly affect the volatility. Comparatively, the most traded cryptocurrencies are relatively less volatility, as they have persistent volume. The mean returns column demonstrates that the lowest traded cryptocurrencies generated the least returns, followed by the medium traded and high traded crypto pairs. Thus, the least traded pairs prove that the consistent volume affects the returns, as we observe lower returns for less traded pairs. High standard deviation in returns also demonstrates a huge fluctuation in the returns because of the low trading volume. We observe less standard deviation in the medium and high traded pairs comparatively because of their persistent volume.

The paper sheds some additional light on the five most traded cryptocurrency pairs (based on observations), which are btc/usd (Bitcoin/USD), eth/btc (Ethereum/Bitcoin), ltc/btc (Litecoin/Bitcoin), ltc/usd (Litecoin/USD), and xrp/btc (Ripple/Bitcoin). Out of the five most traded crypto pairs, two crypto pairs are fiat-crypto pairs, and three are crypto to crypto pairs. Table-3 shows that out of all the top five crypto pairs, the most volatile pair is eth/usd, followed by the xrp/btc and ltc/usd. The eth/btc has the highest mean of daily log returns out of all the data that is being analyzed, followed by xrp/btc and ltc/btc. It can be clearly observed that the least expensive cryptocurrency pair have more mean volatility and mean log returns, such as eth/usd, xrp/btc and ltc/btc. The most standard deviation in returns is observed in the eth/usd, ltc/usd and xrp/btc. Upon glancing on all the traded crypto pairs, the paper explores that all the less known and less expensive crypto pairs have extreme volatility. However, the recognized crypto pairs, which are expensive, such as Bitcoin pairs, have comparatively less volatility. The paper unveils that high-volume trading pairs have superior returns compared to the ones with low trading volume. This finding is analogous to the famous paper by Maheshwari and Dhankar (2017), who finds a direct relationship between the high trading volume stocks and superior returns in the equity market.

Table 4: Illustrates the Minimum, Mean, Median and Maximum Trading Volume of the Five Most Traded Cryptocurrency pairs across all the exchanges.

No.	Exchnage	btc/usd				eth/btc				ltc/btc			
		Min	Mean	Median	Max	Min	Mean	Median	Max	Min	Mean	Median	Max
1	ACX	–	–	–	–	–	–	–	–	–	–	–	–
2	Btcbox	–	–	–	–	–	–	–	–	–	–	–	–
3	Bitfinex	943.4	28592.9	20402	274470	0	69477.4	43074.2	1308330	64.2	49573.4	18275.3	556693
4	Binance	–	–	–	–	550.6	173016	139730	695154	2022.5	158865	144964	988366
5	Bitforex	–	–	–	–	–	–	–	–	–	–	–	–
6	Bitstamp	–	–	–	–	–	–	–	–	–	–	–	–
7	Bittrex	5.1	441.4	126.4	4784.1	594.1	32305.5	14575.2	343626	226.8	59032.4	17188.6	1138890
8	Bit-Z	na	na	na	na	211.2	123755	805898	805898	72	182185	141555	793736
9	CEX.IO	77.9	998.8	621	8252.5	374.7	5304.	3489.9	11823.5	8.2	340.8	248.389	1906.3
10	Gemini	34.3	4575.5	2954.1	49679.8	7.1	12276.5	3208.1	297444	na	na	na	na
11	HitBTC	833.2	9528.9	7634.7	65549.1	6729.9	66630.7	52424.7	309392	1129.1	18736	10553.8	285682
12	Kraken	0	17562985	1723.5	25602300000	1	102835	38968.8	1781420	147.2	23994.5	8070.3	335256
13	Korbit	–	–	–	–	–	–	–	–	–	–	–	–
14	Okcoin	6.8	6370.5	2348.4	81207.7	0	342.7	203.536	2089.6	0	1413	1133	5095.9
15	Poloniex	–	–	–	–	144.7	433379	202908	4997130	108.3	103853	18762.9	4946720

No.	Exchnage	xrp/btc				ltc/usd			
		Min	Mean	Median	Max	Min	Mean	Median	Max
1	ACX	–	–	–	–	–	–	–	–
2	Btcbox	–	–	–	–	–	–	–	–
3	Bitfinex	704900	13610800	7651760	167756000	1.2	150011	70170.1	2452310
4	Binance	708183	69633100	47718900	667334000	–	–	–	–
5	Bitforex	–	–	–	–	–	–	–	–
6	Bitstamp	194547	6151119	3013780	91341200	1893.8	47188.1	30915.4	413120
7	Bittrex	22805.7	20587772	4945415	434441000	86.8	2782.4	1086.8	14943.1
8	Bit-Z	–	–	–	–	–	–	–	–
9	CEX.IO	26137.8	571046	249538	6284570	125.4	814.1	612.9	3417.3
10	Gemini	–	–	–	–	–	–	–	–
11	HitBTC	213835	28167654	24761078	103140000	0.02	45.7	10.2	921.4
12	Kraken	70738.8	8639717	4159080	256716000	0	13089.9	3080.6	200962
13	Korbit	–	–	–	–	–	–	–	–
14	Okcoin	5.998	16172.3	13155.7	54707.4	0.1	87331.3	32735.8	1257240
15	Poloniex	12111.7	44086916	10574638	3624420000	–	–	–	–

Tables 4 demonstrates the exchange wise volume traded descriptive statistics of the five most traded cryptocurrency pairs in the market. Table 4 highlights the minimum, maximum, mean, and median volume for cryptocurrency pairs in respective exchanges. It can be observed that Kraken exchange recorded the highest trading volume for btc/usd and eth/btc among all the exchanges and experienced a massive trading volume of 26 billion (btc/usd) and 1.8 million (eth/btc) approximately. The enormous trading volume is quite evident as to why Kraken is one of the top exchanges, according to the Coingecko report. Bittrex has also experienced a significant volume of 1.1 million for ltc/btc pair, while Poloniex, an exchange known for its promotional schemes, saw an enormous 3.6 billion volume for xrp/btc pair.

Table 5: Illustrates the Standard Deviation of Trading Volume of the Five Most Traded Cryptocurrency pairs across all the exchanges.

No.	Exchange	Standard Deviation				
		BTC/USD	*ETH/BTC*	*LTC/BTC*	*XRP/BTC*	*LTC/USD*
1	**ACX**	-	-	-	-	-
2	**Btcbox**	-	-	-	-	-
3	**Bitfinex**	27215.3	189198	74458.1	72387300	221699
4	**Binance**	-	-	98693.1	65720900	-
5	**Bitforex**	-	-	-	-	-
6	**Bitstamp**	9596.64	35225	20554.6	9845700	49005.7
7	**Bittrex**	749.684	4730.29	97748.5	1279090	3455.13
8	**Bit-Z**	-	-	137410	-	-
9	**CEX.IO**	967.671	5197.02	320.196	885352	654.056
10	**Gemini**	4872.7	81433.6	-	-	-
11	**HitBTC**	8658.11	68012.2	27133.7	21321900	112102
12	**Kraken**	670502000	57031.2	40115.7	16593400	20686.3
13	**Korbit**	-	-	-	-	-
14	**Okcoin**	8741.29	1032.84	1152.68	94521.7	152589
15	**Poloniex**	-	-	256789	161986000	-

Table 5 shows that the inexpensive pairs see high standard deviations except in one pair btc/usd in Kraken exchange. For xrp/btc an inexpensive crypto pair, one can easily observe the massive standard deviation in almost every exchange volume. We also observe substantial amount of standard deviation in ltc/btc and eth/btc as well, throughout all the exchanges.

4. Methodology

Chapter 4 describes the statistical model utilized in the paper to come up with empirical results. The paper aims to analyze the causal effect of the volume on returns and volatility in the cryptocurrency market. This section introduces the comprehensive statistical tool employed in this study, i.e., the Instrumental variable estimation approach. The estimation approach will help study the actual causal relationship between the endogenous independent variable (Volume) and the dependent variable (Volatility and Returns) using the trading fee and lagged volume as exogenous instruments. This section also demonstrates one of the IV estimation methods, i.e., the Two-Stage Least Square or 2SLS/ TSLS model, which provides unbiased regression estimates in the presence of endogenous regressors.

4.1 Instrumental Variable Estimation

Instrumental variable estimation or IV estimation is applied to examine the causal relationship between the dependent variable and the endogenous[1] regressor, i.e. when the independent variable is correlated with the error term. IV estimation is the replacement for the Ordinary Least Square (OLS) and ANOVA, which gives us biased results when the independent variable is correlated with the error term. Hence, a valid instrument(s) is used that impacts the explanatory variable and helps to explain the causal relationship between the dependent variable and the independent variable. The two requirements for IV estimation are: (i) the instrument must be correlated with the endogenous variable (ii) the instrument does not correlate with the error term in the explanatory equation; otherwise, it will suffer from the preceding endogeneity problem again. The instrument helps to discern the precise relationship between the dependent variable and the independent variable.

Bascle (2008) also discusses how endogeneity arises and how to control it. The paper discusses the importance of endogeneity and ways to address this using the Instrumental variable approach. The Two-Stage Least Square (TSLS) method of estimation is used under IV

[1] Endogenous variables in statistical models are those variables which tend to get affected by other variable with which they have relationship with. In other words, it means they correlate with another variable (can be positively or negatively correlated).

estimation. The TSLS or 2SLS is a statistical technique that is used for the structural equations. It is an extension of standard OLS regression, which facilitates the discovery of precise estimates. The following section elucidates the 2SLS model explained by Takashi Yamano in his advanced econometrics article.

4.1.a. Two-Stage Least Square Model

Consider the following model:

$$y_1 = \beta_0 + \alpha_1 y_2 + \beta_1 x_1 + \beta_2 x_2 + .. + \beta_k x_k + u \qquad (1)$$

In equation (1), y_2 is the endogenous variable, $(x_1, \dots, x_k)$ are exogenous variable. If there are m instruments, then $z = (1, x_1, \dots, x_{kw}, z_1, \dots, z_m)$ are correlated with y_2. The reduced form of y_2 can be deduced (where all the exogenous variables along the instruments are presented):

$$y_2 = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + .. + \delta_k x_k + \delta_{k+1} z_1 + .. + \delta_{k+m} z_m + \varepsilon \qquad (2)$$

$$y_2 = \hat{y}_2 + \varepsilon. \qquad (3)$$

In equation (3), $\hat{y}_2$ is the linear estimation of y_2 computed through normal OLS regression, which consists of all the exogenous variables and the instruments. $\hat{y}_2$ possess the exogenous variables and the instruments uncorrelated with the error term (u), in equation (1). However, the error term in equation (1) correlates with the error term in equation (2), which implies while estimating y_2 with the exogenous variable, there are two parts in the equation, one correlated with the error term, u and the other one is not. Hence, estimated y_2 can be written in terms of Z (Z is n*k matrix, which consists of k-1 independent variable and one instrument) in the following way:

$$\hat{y}_2 = Z\hat{\delta} = Z(Z'Z)^{-1}Z'y_2$$

In 2SLS we replace $y_2$with $\hat{y}_2$[2], but y_2 is treated as the variable in X (X is n*k matrix that consists of (k-1) independent variables and one endogenous variable) and demonstrate X with the Z (set of instruments):

$$\hat{X} = Z\hat{\Pi} = Z(Z'Z)^{-1}Z'X = P_z X$$

[2] The reason for the replacement of x_1 with $\hat{x}_1$ is discussed further in the paper.

where $\widehat{\Pi}$ is a (k+m-1) * k matrix with coefficients that appears as following:

$$\widehat{\Pi} = \begin{bmatrix} \delta_1 & 1 & 0 & 0 \\ \delta_2 & 0 & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \delta_{k+m-1} & 0 & 0 & 1 \end{bmatrix}$$

Thus, $y_2$in X must be exhibited as linear projection and additional independent variable in X ought to be articulated by itself. $P_z = Z(Z'Z)^{-1}Z'$ is a n * n idempotent and symmetric matrix. $\hat{X}$ is used as instrument for X and IV estimation is applied:

$$\begin{aligned} \hat{\beta}_{2SLS} &= (\hat{X}'X)^{-1}\hat{X}'Y \\ &= (X'P_z X)^{-1}X'P_z Y \\ &= (X'Z(Z'Z)^{-1}Z'X)^{-1}X'Z(Z'Z)^{-1}Z'Y \end{aligned}$$

It can be represented as:

$$\hat{\beta}_{2SLS} = (\hat{X}'\hat{X})^{-1}\hat{X}'Y$$

The 2SLS estimator is termed as the Two-stage least square estimation method because the endogenous variable is first regressed on the instruments and all the exogenous variables, and the dependent variable is regressed on the fitted endogenous variables and the other exogenous independent variables. The general IV model can be computed using two-stage least squares estimator:

1. First Stage Regression

Compute the normal OLS for each endogenous variable such as y_{2i} on the respective instrument $(Z_{1i}, \ldots\ldots, Z_{mi})$, all the exogenous variables $(x_{1i}, \ldots\ldots, x_{ri})$ and a constant and calculate the fitted values $(\hat{y}_{1i}, \ldots\ldots, (\hat{y}_{ki})$.

2. Second Stage Regression

Regress the dependent variable on the fitted values or the predicted values of the endogenous variables, exogenous terms and constant utilising the OLS. After the regression 2SLS provides with model coefficients such as $\hat{\beta}_0^{2SLS}, \ldots, \hat{\beta}_{k+r}^{2SLS}$.

The paper primarily analyses the impact of volume on volatility and returns through the trading fee. The instrumental variable is an appropriate approach to document the trading volume's effect on the returns and the volatility through trading fees as an instrument. The paper uses the using two-stage least squares method of the IV estimator. The paper demonstrates the following IV regressions using the 2SLS method:

1.a. Regress Log Returns on Log volume (Trading volume) with Trading Fee as instrumental variable. (One instrument)

$$Log\ Return_{i,\tau} = \beta_0 + \beta_1 Log\ Volume + v_{i,\tau}$$

i. **First Stage Regression**:

$$Log\ \widehat{Volume} = \gamma_0 + \gamma_1 Trading\ Fee + \varepsilon_{i,\tau}$$

ii. **Second Stage Regression:**

$$Log\ Return_{i,\tau} = \beta_0 + \beta_1 Log\ \widehat{Volume} + v_{i,\tau}$$

1.b. Regress Log Returns on Log volume (Trading volume) with Trading Fee and Lagged Volume as instrumental variable. (Two Instruments)

$$Log\ Return_{i,\tau} = \beta_0 + \beta_1 Log\ Volume + v_{i,\tau}$$

iii. **First Stage Regression**:

$$Log\ \widehat{Volume} = \gamma_0 + \gamma_1 Trading\ Fee + \gamma_2 Lagged\ Volume + \varepsilon_{i,\tau}$$

iv. **Second Stage Regression:**

$$Log\ Return_{i,\tau} = \beta_0 + \beta_1 Log\ \widehat{Volume} + v_{i,\tau}$$

2.a. Regress Log Volatility on Log volume (Trading volume) with Trading Fee and Lagged Volume as instrumental variable. (One instrument)

$$Log\ Volatility_{i,\tau} = \beta_0 + \beta_1 Log\ Volume + v_{i,\tau}$$

i. First Stage Regression:

$$Log\ \widehat{Volume} = \gamma_0 + \gamma_1 Trading\ Fee + \varepsilon_{i,\tau}$$

ii. Second Stage Regression:

$$Log\ Volatility_{i,\tau} = \beta_0 + \beta_1 Log\ \widehat{Volume} + v_{i,\tau}$$

2.b. Regress Log Volatility on Log volume (Trading volume) with Trading Fee and Lagged Volume as instrumental variable. (Two Instruments)

$$Log\ Volatility_{i,\tau} = \beta_0 + \beta_1 Log\ Volume + v_{i,\tau}$$

i. First Stage Regression:

$$Log\,\widehat{Volume} = \gamma_0 + \gamma_1 Trading\ Fee + \gamma_2 Lagged\ Volume + \varepsilon_{i,\tau}$$

ii. Second Stage Regression:

$$Log\ Volatility_{i,\tau} = \beta_0 + \beta_1 Log\,\widehat{Volume} + \upsilon_{i,\tau}$$

The paper analyses the relationship between the dependent and independent variables with all the maker and taker's fee structure. Furthermore, the trading fee also incorporates the total fee, which summarizes the maker and taker fee at a different trading volume level. The log volume regression of the instruments. The two crucial aspects to keep in mind while analyzing IV estimation results are:

1. 2SLS can give negative R-square sometimes as the actual values of the endogenous variable are taken into the consideration and not the instruments to calculate the Model Sum of Squares (MSS). Model's residuals are calibrated on the basis of the set of the regressors that fits the model (not the instruments). The constant-only model of unexplained variable is not included in the 2SLS model, however it does have coefficient for the constant, so Residual Sum of Squares (RSS) can be more than the Total Sum of Squares (TSS). So, when RSS > TSS, the MSS and R2 is negative.

2. The standard error yielded in the second stage is not accurate because $\hat{x}_1$is itself an estimate used in the regression. The second stage regression cannot yield valid standard errors using the estimated variable, i.e., the $\hat{x}_1$in this case. The IV estimation in Python or any other package in R or Stata automatically corrects for it.

The 2SLS IV estimation method is considered less accurate than OLS because of the following two reasons:

a. The fitted variable ($\hat{x}_1$) has less sample variation than the x_1.
b. The new fitted variable is correlated to higher extent with all the exogenous regressors than the x_1.

5. Results

Chapter 5 presents the results of this study. The main aim is to analyze the causal relationship between the volume and the returns and the volatility in the presence of the trading fee (as an instrument) in the cryptocurrency market. Before implementing and testing the primary 2SLS model, the paper studies the general relationship between all the variables, as mentioned above, to develop a broad understanding.

1. The first section (5.1) presents and interprets the Pearson correlation coefficients.
2. The second section (5.2) shows and describes the Ordinary Least Squares (OLS) regressions results.
3. The third section (5.3) portray the Durbin-Wu-Hausman test, which is the independent variable's endogeneity test.
4. The fourth section (5.4) illustrates the Instrumental Variable estimation's main results through the 2SLS method.

5.1 Pearson Correlation

The first step to study the causal relationship between volume, returns, and volatility is to have a general understanding of the correlation with each other. The Pearson correlation coefficient, along with the p-value, have been computed for this purpose. The paper examines the correlation between the

1. log volume and different types of trading fee,
2. log volatility and different types of trading fee,
3. Log return and different types of trading fee,
4. between log volume, log returns and log volatility.

The Pearson correlation coefficient is given by:

$$\rho_{x,y} = \frac{cov(X,Y)}{\sigma_x \sigma_y}$$

where:

cov is covariance.

σ_x is standard deviation

σ_y is standard deviation

Table 6: Illustrates the Pearson Correlation Coefficient

No.	Particulars	Log Volume	Log Returns	Log Volatility
		Coefficient	*Coefficient*	*Coefficient*
1	**Total Low Volume Fee**	0.0784*	-0.0050*	-0.1472*
2	**Total Median Volume Fee**	0.1212*	-0.0042*	-0.1138*
3	**Total High-Volume Fee**	0.1367*	-0.0026*	-0.0617*
4	**Total Average Volume Fee**	0.1171*	-0.0042*	-0.1146*
5	**Low Volume Fee (Maker)**	0.1370*	-0.0065	-0.1596*
6	**Median Volume Fee (Maker)**	0.1689*	-0.0056*	-0.1283*
7	**High Volume Fee (Maker)**	0.1821*	-0.0050*	-0.0996*
8	**Average Volume Fee (Maker)**	0.1666*	-0.0059*	-0.1335*
9	**Low Volume Fee (Taker)**	-0.0342*	-0.0017	-0.1090*
10	**Median Volume Fee (Taker)**	0.0041*	-0.0007	-0.0675*
11	**High Volume Fee (Taker)**	0.03479*	0.0020*	0.0147*
12	**Average Volume Fee (Taker)**	0.0002	-0.0001	-0.0622*

Note: * represents the statistical significance at the 0.05 level.

Table 6 represents a comprehensive Pearson coefficient correlation, along with the statistical significance. The paper computed the correlation coefficient between the various fee structure (such as Total fees, Makers' Fee, and Takers' Fee) and pertinent variables such as Log volume, Log returns, and Log volatility.

5.1.1 Correlation between Log Volume and different Trading Volume Fee categories

Several studies in the empirical literature suggest a negative correlation between the trading fee and trading volume. Barclay (1998) and Kramer (1999) illustrate in their studies that the trading fee (documented as a bid-ask spread) has a significant negative impact on the trading volume. However, unlike the equity market, the cryptocurrency market is exposed to several other risk factors. Cryptocurrencies volume is assumed to be based on the positive economic news regarding the cryptocurrency legalization and the google web searches proposed by the Kristoufek (2013) and Polasik et al. (2015) in their papers.

Upon careful observation of table 6, it can be noticed that magnitude of the correlation coefficient increases with the increase in trading volume fee, which suggests that high volume traders are less concerned with the trading fee than the low volume traders. The correlation between total low volume fee and log volume is 0.07 (p-value < 0.001), whereas, for a total high-volume fee and log volume, it is 0.13 (p-value < 0.001), which is almost two times greater than the former. The correlation coefficient's difference of magnitude can be understood by the lower fee charged to the high-volume traders than the low volume traders. The relatively strong positive correlation between trading volume and trading fee can be observed in makers' fee structure may be because they are sometimes entitled to a good rebate or charged 0% as the trading fee, so even if the exchange removes the rebate or charge 0%, the makers are not affected that much as they are not paying anything. Moreover, if the exchange increases the trading fee, the makers tend to pass on that fee to the takers, thereby saving themselves from the fees' burden. However, maker/taker has the same pattern of increase in the magnitude of correlation as the trading volume changes (the correlation coefficient increases as more volume is traded). Nevertheless, we also witness a negative correlation between log volume and takers' low volume fee (coefficient = -0.0342, p-value < 0.001), which indicates small retail buyers are indeed affected by the trading fee. That is why a weak positive correlation is observed with low trading fees and log volume.

5.1.2 Correlation between Log Returns and different Trading Volume Fee categories

Negative correlation coefficients are observed between the log returns and almost all the trading fee structures. The negative relation between the returns and the trading fee was anticipated. As the trading fee decreases, we expect the volume to increase among the makers and takers, thereby increasing the returns given the limited supply and increased demand principle and many other factors. The total fee structure and makers' fee structure illustrates a higher magnitude of negative correlation with log returns than the takers' fee structure, which denotes that the maker’s return decline relatively more than taker’s return as the trading fee goes up. At the same time, we witness that the correlation coefficient for the total low volume trader is -0.0050, whereas for the total high-volume trader is -0.0026, which indicates that the low volume traders returns get more affected when the trading fee is increased as they will not be able to afford or willing to pay a higher price to acquire cryptocurrency. It can be justified by the taker's low volume fee, which shows the negative correlation coefficient between volume and trading fee. We also see a positive correlation between log-returns and trading fees for high volume takers (coefficient = 0.0020, p-value= -0.045), suggesting that high volume takers trade even in high fee scenario, thereby driving the returns.

5.1.3 Correlation between Log Volatility and different Trading Volume Fee categories

The paper witnesses a relatively stronger negative correlation between the log volatility and trading fee than log-returns and trading fees. This stronger negative relation implies the trading fee has relatively more impact on the log volatility than on the log returns. The negative correlation articulates that an increase in the trading fee decreases the volatility in the crypto market. People tend to refrain from excessive buying and selling during high fees, which brings down the market volatility. Table 6 shows that the makers and total fee structures tend to have a higher negative correlation with the log volatility, which specifies that makers affect the volatility more than the takers. It can be perceived from the table that the low volume fee traders affect the market volatility relatively more than the high-volume fee traders as the high-volume traders tend to be uniform in providing with the liquidity into the market, thereby inducing less volatility in the market. However, again, we notice that high volume fees in takers cases do not affect the volatility when the fee is increased, probably because they are not much affected by the trading fee.

5.1.4 Correlation between Log Volatility and different Trading Volume Fee categories

This section discusses the correlation between the main three essential variables in our study: (1) Log volume, (2) Log returns, (3) Log volatility. Table 7 demonstrates a very low (0.01) but a positive and significant relationship between log volume and log returns.

Table 7: Illustrates Pearson Correlation between Log Volume, Log Return and Log Volatility

No.	Particulars	Log volume	Log Return	Log volatility
		Coef.	*Coef.*	*Coef.*
1	**Log Volume**	1	0.01*	0.0547*
2	**Log Return**	0.01*	1	0.0609*
3	**Log Volatility**	0.0547*	0.0547*	1

Note: * represents the statistical significance of the p-value at alpha level of 0.05.

Most empirical studies, such as Maheshwari and Dhankar (2017), portrayed that high-volume stocks showed better returns than those traded less. Another interesting paper by Gebka and Wohar, 2013, shows a positive relationship between the volume and returns. The studies above in the equity market have been consistent with the cryptocurrency market finding in this paper. Moreover, cryptocurrency has no underlying asset, so we expect volume to drive the returns even more, but in reality, returns are affected by many factors (such as government policies, state of the economy, trading fees) other than volume.

A positive correlation can also be seen between log volume and log volatility. The empirical research by Bouri et al. (2019) concluded the relationship between log volume and log volatility but only for three cryptocurrencies out of the seven cryptocurrencies studied. Negative relationship illustrates illiquidity of the asset, risk premium, and lack of information. However, a positive relation in the domain of cryptocurrency sums up perfectly as it suggests that the market is uncertain about the future values, as the buyers/ investors have a 50-50 split view for the value of an asset, which is precisely the scenario in the crypto market. Log volatility and log return also have a positive correlation, suggesting that in the uncertain market such as the cryptocurrency market, trading volume tends to go up, thereby driving the volatility, which drives the log return up.

5.2 Ordinary Least Squares (OLS)

The study runs various regressions to examine the statistical relationship between the dependent variable (such as returns and volatility) and the independent variable (such as fee and volume) and their economic significance. This section explores various regression models, such as regressing log volume on different fee structures (Total Fee, Makers' Fee, and Takers' Fee), log volatility on different fee structures, and log return on different fee structures to determine the first-hand effect of the trading fee on the variables above. For the primary variable, such as log volume and log volatility, the paper scrutinizes additional regression models by incorporating the lagged dependent variable as the regressor. The inclusion of the lagged dependent variable in the regression possibly takes out much variance and reduces the other estimates, such as betas. To provide the unbiased estimated, which are not based on the past values, the paper considers the lagged dependent variable to control for autocorrelation. As the paper examines the daily observations, it makes more sense to use the lagged dependent variable or lagged volume variable to correct the excessive variance by the previous day's values. The paper discusses the regressions in the following order: (1) Log volume on different trading fee, (2) Log volume on different trading fee and lagged volume,

(3) Log volatility on different trading fee, (4) Log volatility on different trading fee and lagged volatility, (5) Log return on different trading fee, (6) Log volatility on log volume, (7) Log volatility on log volume and lagged volatility, (8) Log return on log volume.

5.2.1 Log Volume on different Trading Fee

The following table 8 demonstrates the results computed from the following regression:

$$Log\ volume_{i,\tau} = \alpha + \beta_1 Trading\ fee_i + \epsilon_{i,\tau}$$

The table illustrates a positive relationship between the volume and fee structures. According to the empirical literature, a negative affiliation was expected. The probable reason for the positive relationship can be that positive economic news drives the trading fee, and irrespective of the increase in the fee, the volume goes up. Given the fixed supply of cryptocurrency and more demand, people anticipate an decent appreciation that outshines the high trading fee.

Table 8: OLS Regression of Log Volume on different Fee Structures

No.	Fees (Independent Variable)	α	β (Fee)	R-Squared	Economic Significance
1	**Total Low Volume Fee**	10.4711*** (0.006)	1.3852*** (0.018)	0.006	0.2864
2	**Total Median Volume Fee**	10.3954*** (0.006)	2.4396*** (0.021)	0.015	0.4424
3	**Total High-Volume Fee**	10.3954*** 0.005	2.6212*** (0.020)	0.019	0.4987
4	**Total Average Volume Fee**	10.3371*** (0.006)	2.3380*** (0.020)	0.014	0.4273
5	**Low Trading Volume Fee (Taker)**	-6.1522*** (0.009)	-3.5197*** (0.048)	0.010	-0.2754
6	**Median Trading Volume Fee (Taker)**	10.7864*** (0.010)	0.2444*** (0.061)	0	0.01502
7	**High Trading Volume Fee (Taker)**	10.6038*** (0.008)	1.7917*** (0.053)	0.001	0.1266
8	**Average Trading Volume Fee (Taker)**	10.8211*** (0.010)	0.0137 (0.061)	0.000	0.0008
9	**Low Trading Volume Fee (Maker)**	10.5212*** (0.004)	3.6288*** (0.027)	0.019	0.5001
10	**Median Trading Volume Fee (Maker)**	10.5546*** (0.004)	4.7920*** (0.029)	0.029	0.6165
11	**High Trading Volume Fee (Maker)**	10.6137*** (0.004)	5.1407*** (0.029)	0.033	0.6646
12	**Average Trading Volume Fee (Maker)**	10.5391*** (0.004)	4.7351*** (0.004)	0.028	0.6079

Note: This table presents results of OLS regressions of log volume on the level of fees and lagged log volume ($Log\ volume_{i,\tau} = \alpha + \beta_1 Trading\ fee + \epsilon_{i,\tau}$). Each row presents results using a different measure of trading fees. Economic significance is computed as the change in the dependent variable for a one-standard-deviation increase in the fee. Standard errors are presented in the parenthesis ** represents p-value < 10%, ** represents p-value < 5%, and *** represents P-value < 1%*. The sample period is from 2015 to 2020.

The other main probable reason for the positive relation is marked in the study by Harris (2013) and Angel et al. (2006), which states that the intricate Maker and Taker Fee causes inaccuracy in the market and the actual make/taker fee is not disclosed in the quotation. This kind of actual transaction cost concealment misleads the traders. One might trade on the market outlook and get deceived by the undisclosed trading fee. Cryptocurrency is a highly unpredictable financial asset class, so many reasons can be deduced from the given empirical research.

Secondly, we observe a more substantial positive relation between makers trading fees and the volume traded than the relationship with takers' fees and total fees. This relationship is primarily due to two possible reasons: first, the makers are given a rebate or mostly charged zero percent, so even if the fee is increased, it is very low. The second probable reason is that the makers tend to pass on the high trading fee load to the takers by quoting the higher price in the limit order book. That is why we observe the stronger positive relationship between high trading volume and log volume. In line with the reasoning above, we can also comprehend that the low taker fee is negatively related to the volume, having a coefficient of -3.5197 with a p-value of 0.010. This result explicitly indicates that the low volume traders are affected by the high trading fee passed on by the exchanges and the makers.

The magnitude of the coefficient increases from low volume trading fees to high volume trading fees in all the types of fees, which tells us that the impact or the burden of the fee is less for the high-volume traders, possibly because they pay relatively less than low volume traders. In table 8, we find that the beta coefficients statistically significant for all the fee structures except the average fee taker.

5.2.2 Log volume on different Trading Fee and Lagged Volume

Table 9 shows results that are computed on the basis of the following regression:

$$Log\ volume_{i,\tau} = \alpha + \beta_1 Trading\ fee + \beta_2 Lagged\ volume + \epsilon_{i,\tau}$$

The equation is similar to the table 8 equation, except for one thing, i.e., the lagged volume, which is the lagged dependent variable. The lagged dependent variable is added to correct the autocorrelation and reduce the variance by the past values. This variable gives us less significant estimates (such as betas). Upon a careful comparison between table 8 (without lagged dependent variable) and table 9 (with lagged dependent variable), it can be concluded that the theory of dependent lagged volume is appropriate, as our coefficients in the lagged model (table 9) are relatively less than the model without a lagged variable (table 8). Also, the R-square of the model improves significantly upon the introduction of a lagged dependent variable.

The beta coefficients for the model without the lagged dependent variable seem to have a massive variance from the past values, as we notice the reduction in beta coefficients by ten times (in some cases 20 times) upon using the lagged volume model. The lagged dependent variable model changes the signs for almost all types of takers' fees except high volume traders of takers. It illustrates that the low volume, medium volume, and average volume takers are affected by the increase in the trading fee, where the high-volume takers and Makers are not much affected by the increase in trading volume. When it comes to magnitude from low volume fee to high volume fee, it increases as presented in the earlier model. Table 9 undoubtedly marks the importance of using the lagged dependent variable in the regression model.

Table 9: OLS Regression of Log volume on different Fee Structures and Lagged Volume

	Particulars	α	β (Fee)	β (Lagged Volume)	R-Squared	Economic Significance
1	**Total Low Volume Fee**	1.0263*** (0.005)	0.0746*** (0.007)	0.9053*** (0.000)	0.854	0.0154
2	**Total Median Volume Fee**	1.0170*** (0.005)	0.1736*** (0.008)	0.9046*** (0.000)	0.854	0.0314
3	**Total High-Volume Fee**	1.0232*** (0.004)	0.2099*** (0.008)	0.9041*** (0.000)	0.854	0.0399
4	**Total Average Volume Fee**	1.0178*** (0.005)	0.1644*** (0.008)	0.9046*** (0.000)	0.854	0.0300
5	**Low Trading Volume Fee (Taker)**	1.0853*** (0.005)	-0.245*** (0.018)	0.9054*** (0.000)	0.854	-0.0190
6	**Median Trading Volume Fee (Taker)**	1.0524*** (0.006)	-0.071*** (0.023)	0.9056*** (0.000)	0.854	-0.0044

7	**High Trading Volume Fee (Taker)**	1.0237*** (0.005)	0.1556*** (0.020)	0.9055*** (0.000)	0.854	0.0109
8	**Average Trading Volume Fee (Taker)**	1.0542*** (0.006)	-0.085*** (0.023)	0.9056*** (0.000)	0.854	-0.0052
9	**Low Trading Volume Fee (Maker)**	1.0347*** (0.004)	0.2495*** (0.011)	0.9043*** (0.000)	0.854	0.0343
10	**Median Trading Volume Fee (Maker)**	1.0442*** (0.004)	0.3667*** (0.011)	0.9034*** (0.000)	0.854	0.0471
11	**High Trading Volume Fee (Maker)**	1.0530*** (0.004)	0.4133*** (0.011)	0.9030*** (0.000)	0.854	0.0534
12	**Average Trading Volume Fee (Maker)**	1.0424*** (0.004)	0.3583*** (0.011)	0.9035*** (0.000)	0.854	0.0460

Note: This table presents results of OLS regressions of log volume on the level of fees and lagged log volume ($Log\ volume_{i,\tau} = \alpha + \beta_1 Trading\ fee + \beta_2 Lagged\ volume + \epsilon_{i,\tau}$). Each row presents results using a different measure of trading fees. Economic significance is computed as the change in the dependent variable for a one-standard-deviation increase in the fee. Standard errors are presented in the parenthesis ** represents p-value < 10%, ** represents p-value < 5%, and *** represents P-value < 1%*. The sample period is from 2015 to 2020.

5.2.3 Log Volatility on different Trading Fees

Table 10 shows results that are computed on the basis of the following regression:

$$Log\ volatility_{i,\tau} = \alpha + \beta_1 Trading\ fee + \epsilon_{i,\tau}$$

The results in table 10 show a negative relation between the trading fee and volatility for all types of fee structures. The results are consistent with our anticipation because as the fee increases, the volume is expected to decrease, bringing down the volatility. Mainly, in the cryptocurrency domain, an unpredictable asset without any underlying base, an increase in volume tends to increase volatility.

The maker and taker fee structure's coefficients remain somewhat similar to their subcategories, except the taker's high trading fee, which has a statistically significant beta coefficient of 0.736 with 0.0531 economic significance. It implies that the high-volume takers are not affected much by the fee change. Upon analyzing the maker's subcategories, taker, and total fee structures, we see a commonality in them, i.e., we observe the beta coefficients decrease as the trading volume increases. It denotes that the low volume traders are more affected by an increase in fee, and as they reduce the trading volume, the volatility also gets reduced. Hence, a negative relationship between the trading fee and log volatility is justified and anticipated. The small beta coefficients of volume for high volume makers fee and total high-volume fee suggests that these big traders tend to be calm in the case of fee increase and does not cause much volatility in case of a fee increase or decrease, maybe because they are already given a lower fee and more rebates comparatively to the takers.

Table 10: OLS Regression of Log Volatility on different Fee Structures

No.	Particulars	α	β (Fee)	R-Squared	Economic Significance
1	**Total Low Volume Fee**	-6.2822*** (0.005)	-1.8347*** (0.014)	0.019	-0.3785
2	**Total Median Volume Fee**	-6.4287*** (0.005)	-1.5835*** (0.016)	0.011	-0.2872
3	**Total High-Volume Fee**	-6.6367*** (0.004)	-0.7760*** (0.015)	0.003	-0.1491
4	**Total Average Volume Fee**	-6.4262*** (0.005)	-1.5833*** (0.016)	0.011	-0.2895
5	**Low Trading Volume Fee (Taker)**	-6.1522*** (0.007)	-3.5197*** (0.038)	0.010	-0.2754
6	**Median Trading Volume Fee (Taker)**	-6.3694*** (0.008)	-2.5992*** (0.048)	0.003	-0.1611
7	**High Trading Volume Fee (Taker)**	-6.8575*** (0.006)	0.7363*** (0.041)	0	0.0531
8	**Average Trading Volume Fee (Taker)**	-6.4116*** (0.008)	-2.3639*** (0.047)	0.003	-0.1477
9	**Low Trading Volume Fee (Maker)**	-6.4965*** (0.004)	-3.0201*** (0.021)	0.022	-0.4132
10	**Median Trading Volume Fee (Maker)**	-6.6112*** (0.003)	-2.5655*** (0.023)	0.014	-0.3287
11	**High Trading Volume Fee (Maker)**	-6.6805*** (0.003)	-1.9432*** (0.023)	0.008	-0.2512
12	**Average Trading Volume Fee (Maker)**	-6.5930 *** (0.003)	-2.6803 *** (0.023)	0.015	-0.3423

Note: This table presents results of OLS regressions of log volume on the level of fees and lagged log volume ($Log\ volatility_{i,\tau} = \alpha + \beta_1 Trading\ fee + \epsilon_{i,\tau}$). Each row presents results using a different measure of trading fees. Economic significance is computed as the change in the dependent variable for a one-standard-deviation increase in the fee. Standard errors are presented in the parenthesis ** represents p-value < 10%, ** represents p-value < 5%, and *** represents P-value < 1%*. The sample period is from 2015 to 2020.

5.2.4 Log Volatility on different Trading Fee and Lagged Volatility

Table 11 shows results that are computed on the basis of the following regression:

$$Log\ volatility_{i,\tau} = \alpha + \beta_1 Trading\ fee + \beta_2 Lagged\ volatility + \epsilon_{i,\tau}$$

The following table illustrates that the negative relation between the trading fee and log volatility persists, even after the inclusion of the lagged dependent variable in the model. Although we witness the significant reduction in the beta coefficients of this model as compared to the beta coefficients without the lagged dependent variable model (table 10), it still manages to stay negative, unlike in the case of the regression of log volume on trading fee, where the beta coefficients signs changed for takers in the lagged dependent variable model. The magnitude of the beta coefficient of fee and economic significance decreases as the trading volume fee goes up, signifying that the low volume traders are ones who are affected much rather than high volume traders. In other words, it implies low volume traders are more likely to act robustly relatively to the high-volume traders in the event of increase or decrease of the fee. It can be seen that the high-volume taker fee has a positive relationship with volatility, which implies that the trading fee has a less impact on the high-volume takers.

Table 11: OLS Regression of Log Volatility on different Fee Structures

No.	Particulars	α	β (Fee)	β (Lagged Volatility)	R-Squared	Economic Significance
1	**Total Low Volume Fee**	-4.617***	-1.346***	0.2651***	0.088	-0.2778
		(0.008)	(0.014)	(0.001)		
2	**Total Median Volume Fee**	-4.687***	-1.152***	0.2710***	0.084	-0.2090
		(0.008)	(0.016)	(0.001)		
3	**Total High-Volume Fee**	-4.800***	-0.560***	0.2767***	0.080	-0.1076
		(0.008)	(0.015)	(0.001)		
4	**Total Average Volume Fee**	-4.686***	-1.152***	0.2709***	0.084	-0.2107
		(0.008)	(0.016)	(0.001)		
5	**Low Trading Volume Fee (Taker)**	-4.482***	-2.558***	0.2716***	0.083	-0.2002
		(0.009)	(0.037)	(0.001)		
6	**Median Trading Volume Fee (Taker)**	-4.610***	-1.874***	0.2764***	0.080	-0.1161
		(0.010)	(0.046)	(0.001)		
7	**High Trading Volume Fee (Taker)**	-4.948***	0.5333***	0.2786***	0.078	0.0385
		(0.009)	(0.039)	(0.001)		
8	**Average Trading Volume Fee (Taker)**	-4.638***	-1.703***	0.2768***	0.080	-0.1064
		(0.010)	(0.046)	(0.001)		
9	**Low Trading Volume Fee (Maker)**	-4.792***	-2.225***	0.2624***	0.090	-0.3045
		(0.008)	(0.021)	(0.001)		
10	**Median Trading Volume Fee (Maker)**	-4.836***	-1.874***	0.2685***	0.086	-0.2401
		(0.008)	(0.022)	(0.001)		
11	**High Trading Volume Fee (Maker)**	-4.858***	-1.411***	0.2728***	0.082	-0.1824
		(0.008)	(0.022)	(0.001)		

12	**Average Trading Volume Fee (Maker)**	-4.829*** (0.008)	-1.960*** (0.022)	0.2676*** (0.001)	0.086	-0.2504

Note: This table presents results of OLS regressions of log volume on the level of fees and lagged log volume ($Log\ volatility_{i,\tau} = \alpha + \beta_1 Trading\ fee + \beta_2 Lagged\ volatility + \epsilon_{i,\tau}$). Each row presents results using a different measure of trading fees. Economic significance is computed as the change in the dependent variable for a one-standard-deviation increase in the fee. Standard errors are presented in the parenthesis ** represents p-value < 10%, ** represents p-value < 5%, and *** represents P-value < 1%*. The sample period is from 2015 to 2020.

5.2.5 Log Return on different Trading fee

Table 12 illustrates the results that are computed on the basis of the following regression:

$$Log\ returns_{i,\tau} = \alpha + \beta_1 Trading\ fee + \epsilon_{i,\tau}$$

An essential aspect in this regression would be that the paper will not incorporate the lagged returns or the lagged dependent variable, it is because that the returns have no such persistence as in the case of volume and volatility. It can also be perceived from table 12 that the regression estimates of log returns on trading fee i.e., the betas (fee), are small and significant, unlike volume and volatility regression without their lagged values, which articulates that there is less variance in the beta from the past values.

Table 12: OLS Regression of Log Returns on different Fee Structures

No.	Particulars	α	β (Fee)	R-Squared	Economic Significance
1	**Total Low Volume Fee**	0.0007*** 0	-0.0044*** (0.001)	0.000	-0.0009
2	**Total Median Volume Fee**	0.0004*** 0	-0.0042*** (0.001)	0.000	-0.0007
3	**Total High-Volume Fee**	-2.5e-05* 0	-0.0025*** (0.001)	0.000	-0.0004
4	**Total Average Volume Fee**	0.0004 0	-0.0042*** (0.001)	0.000	-0.0007
5	**Low Trading Volume Fee (Taker)**	0.0003 0	-0.0041* (0.002)	0.000	-0.0003
6	**Median Trading Volume Fee (Taker)**	-0.0001 0	-0.0022 (0.003)	0.000	-0.0001
7	**High Trading Volume Fee (Taker)**	-0.0011*** 0	0.0053** (0.003)	0.000	0.0003
8	**Average Trading Volume Fee (Taker)**	-0.0004 0	-0.0006 (0.003)	0.000	-0.00003
9	**Low Trading Volume Fee (Maker)**	0.0003 0	-0.0086*** (0.001)	0.000	-0.0011

10	**Median Trading Volume Fee (Maker)**	3.264e-06** 0	-0.0079*** (0.001)	0.000	-0.0010
11	**High Trading Volume Fee (Maker)**	-0.0002 0	-0.0071*** (0.001)	0.000	-0.0009
12	**Average Trading Volume Fee (Maker)**	6.079e-05 0	-0.0083*** (0.001)	0.000	-0.0010

Note: This table presents results of OLS regressions of log volume on the level of fees and lagged log volume ($Log\ returns_{i,\tau} = \alpha + \beta_1 Trading\ fee + \epsilon_{i,\tau}$). Each row presents results using a different measure of trading fees. Economic significance is computed as the change in the dependent variable for a one-standard-deviation increase in the fee. Standard errors are presented in the parenthesis ** represents p-value < 10%, ** represents p-value < 5%, and *** represents P-value < 1%*. The sample period is from 2015 to 2020.

The empirical literature has documented a negative relationship between the returns and the trading fee. The high trading fees affect the trading volume to some extent, thereby affecting the market's returns. Upon comparing the magnitude of the low volume trader and high-volume trader's beta coefficients, it can be concluded that the fee impacts the low traders' return more than the high-volume traders. On comparing the average trading fee for the maker and taker, it can be inferred that the increase in trading fees affects the log-returns more in the makers' case than the takers. It is because the makers induce liquidity through limit order into the market, and when the trading fee is increased, their returns get significantly affected.

5.2.6 Log Volatility on Log Volume

The table 13 represents various regressions, out of which we will discuss the following regression model:

$$Log\ volatility_{i,\tau} = \alpha + \beta_1 Log\ volume + \epsilon_{i,\tau}$$

We notice a positive relationship between volatility and volume. The empirical literature also suggests a positive relationship between the two if the financial asset is not sound and has an indeterminate future, and cryptocurrency is one such unpredictable domain. A famous paper by Louhichi, 2011, also finds a strong significant positive relationship between volume and volatility, even after controlling for the intraday patterns' impact. Hence, whenever the volume surges, it is expected that the volatility will also go up in cryptocurrency market.

5.2.7 Log volatility on Log Volume and Lagged Volatility

The table 13 represents various regressions, out of which we will discuss the following regression model:

$$Log\ volatility_{i,\tau} = \alpha + \beta_1 Log\ volume + \beta_2 Lagged\ volatility + \epsilon_{i,\tau}$$

In this regression, lagged volatility is also integrated to control for the mean version, as the volatility tends to mean revert and likely to be persistent. So, if an asset is volatile at day 'd,' it is likely that the asset is volatile at 'd+1' day as well, even if there is consistent volume. In order to control for the scenario above, lagged volatility is integrated into the equation. A paper by Louhichi, 2011 apply a similar principle to control the previous values of intraday trade patterns from driving the regression estimates. A positive relationship can be witnessed yet again, although the beta coefficient magnitude is slightly less than the model not having the lagged volatility, which was expected in this model. Bouri et al. (2019) also finds that volume granger causes 3 cryptocurrencies' volatility out of the seven cryptocurrencies studied in the paper.

5.2.8 Log Returns on Log Volume

The table 13 represents various regressions, out of which we will discuss the following regression model:

$$Log\ returns_{i,\tau} = \alpha + \beta_1 Log\ volume + \epsilon_{i,\tau}$$

The relationship between the log returns and log volume has been subject to various empirical literature contradictions. Researchers such as Maheshwari and Dhankar (2017) and Lee and Swaminathan (2002) illustrate a relation between the volume and returns; on the other hand, researchers such as Lee and Rui, 2002 and Cheng. F. Lee (2000) demonstrates no significant relation between the trading volume and the returns. The paper finds a statistically significant positive relationship between the trading volume and the economic significance is 0.17%. The paper does not examine the log-returns with lagged variables because the returns do not show much persistence than the volume and volatility. It is a critical finding in the cryptocurrency domain that is consistent with the Bouri et al. (2019) results, who finds that volume granger causes the returns in the cryptocurrencies.

Table 13: OLS Regression of Log Volatility on Log Volume/Log Volume and Lagged Volume

No.	Dependent Variable	α	β (Volume)	β (Lagged Volatility)	R-squared	Economic Significance
1	**Log Volatility**	7.1737***	0.0420***	-	0.003	0.1523
		(0.009)	(0.001)			
2	**Log Return**	-0.005***	0.0005***	-	0	0.0017
		(0.001)	(5.08e-05)			
3	**Log Volatility (with Lagged Volatility)**	-5.2803***	0.0360 ***	0.2773 ***	0.080	0.1298
		(0.012)	(0.001)	(0.001)		

Note: This table presents results of various OLS regressions (I) Log volatility on log volume, (II) Log return on log volume (III) Log volatility on log volume and lagged volume, and (IV) Log return on log. Each row presents the dependent variable against which volume is regressed and column (α, β, R-squared, Economic significance) represents the regression estimates. Economic significance is computed as the change in the dependent variable for a one-standard-deviation increase in the fee. Standard errors are presented in the parenthesis ** represents p-value < 10%, ** represents p-value < 5%, and *** represents P-value < 1%*. The sample period is from 2015 to 2020.

5.3 Durbin-Wu-Hausman Test

Durbin-Wu-Hausman test is the classical test of endogeneity that compares the OLS estimates with that of the 2SLS, as it utilizes the fact that the OLS estimates are relatively efficient (discussed below). Suppose a given linear model $y = bX + e$, where y is the dependent variable, and X is the matrix of regressors, b is the coefficients vector, and e is the error term. The b has two coefficients: b_0 and b_1. Under the null hypothesis, the b_1 is the efficient (in terms that it has the smallest asymptotic variance) at least among all the elements in b vector, and the alternate hypothesis is that b_0 is efficient and b_1 is not. If the p-value is less than the threshold assuming alpha as 0.05, then we reject the null hypothesis that b_1 is not efficient and exogenous. Thus, discarding the null hypothesis implies that the independent variable is endogenous for the given instrument in the 2SLS estimation model. The Wu-Hausman Test Statistic is:

$$H = (b_1 - b_0)'(Var(b_0) - Var(b_1))^{\dagger}(b_1 - b_0)$$

where † is the Moore-Penrose pseudoinverse. Under the null the statistic has chi squared distribution, having degrees of freedom as rank of the matrix $Var(b_0) - Var(b_1)$.

This test is computed after the Two-square least square regression (IV estimation method) to check the endogeneity. The test is conducted on IV regression as it investigates the endogeneity of the independent variable. The paper computes the IV-2SLS, i.e., Instrumental variable estimation through Two-Stage Least Square method, on four models:

1. Log return on log volume with the fee structure as an instrumental variable

2. Log return on log volume with the fee structure and lagged volume as an instrumental variable
3. Log volatility on log volume with the fee structure as an instrumental variable
4. Log volatility on log volume with the fee structure and lagged volume as an instrumental variable

The paper divides the results for the Durbin test in two parts (I) Log return on log volume with one instrument (trading fee) and two instruments (trading fee and lagged volume), and (II) Log volatility on log volume with one instrument (trading fee) and two instruments (trading fee and lagged volume).

5.3.1 Log Return on Log Volume (For One Instrument and Two Instruments)

Table 14: Durbin-Wu-Hausman Test Results

Number	Fee (instruments)	One Instrument	Two instruments
		Statistic	***Statistic***
1	**Total Low Volume Fee**	32.0074*	1545.3258*
2	**Total Median Volume Fee**	28.2149*	1547.1026*
3	**Total High-Volume Fee**	15.4409*	1544.7945*
4	**Total Average Volume Fee**	27.6380*	1546.7870*
5	**Low Trading Volume Fee (Taker)**	1.9147	1538.6694*
6	**Median Trading Volume Fee (Taker)**	0.5803	1539.2344*
7	**High Trading Volume Fee (Taker)**	2.7995	1538.9262*
8	**Average Trading Volume Fee (Taker)**	0.0435	1539.1941*
9	**Low Trading Volume Fee (Maker)**	60.1198*	1551.8287*
10	**Median Trading Volume Fee (Maker)**	52.3023*	1308.3776*
11	**High Trading Volume Fee (Maker)**	46.4358*	1307.9638*
12	**Average Trading Volume Fee (Maker)**	55.3923*	1554.1367*

Note: The fee column represents various fee structures, which are the different instruments for the instrumental variable regression of Log return on log volume on one instrument (fee) and two instruments (fee and lagged volume). * represents the statistical significance of the p-value at alpha level of 0.05

Table 14 illustrates the results of the 2SLS for Log returns on log volume using one instrument and two instruments. In the scenario of one instrument, it is observed that log volume is not endogenous in case of takers fee structure (because the p-value exceeds 0.05), whereas in case of two instruments, it is seen log volume appears to be endogenous for all the instrumental variables. For the taker's fee, the p-value exceeds the level of 0.05, which does not reject all the endogenous variables' null hypothesis. According to the results, the taker's fee alone fails to be an efficient instrument to predict the relation between the log returns and log volume. Whereas all the other fee structures having a p-value approximately zero, proves to be reasonably decent instruments.

5.3.2 Log Return on Log Volume (For One Instrument and Two Instruments)

Table 15: Durbin-Wu-Hausman Test Results

Number	Particulars	One Instrument	Two instruments
		Statistic	Statistic
1	**Total Low Volume Fee**	20479.2358*	12824.4862*
2	**Total Median Volume Fee**	13047.2581*	11808.1753*
3	**Total High-Volume Fee**	4330.9120*	11022.8658*
4	**Total Average Volume Fee**	13158.9671*	11817.3016*
5	**Low Trading Volume Fee (Taker)**	10151.2110*	14136.7178*
6	**Median Trading Volume Fee (Taker)**	4040.2247*	12098.9267*
7	**High Trading Volume Fee (Taker)**	155.3151*	10625.2122*
8	**Average Trading Volume Fee (Taker)**	3399.7601*	11823.6660*
9	**Low Trading Volume Fee (Maker)**	25264.7900*	11368.0736*
10	**Median Trading Volume Fee (Maker)**	17314.3609*	11398.9689*
11	**High Trading Volume Fee (Maker)**	11056.2030*	11401.7479*
12	**Average Trading Volume Fee (Maker)**	18587.9296*	11398.3486*

Note: The fee column represents various fee structures, which are the different instruments for the instrumental variable regression of Log volatility on log volume on one instrument (fee) and two instruments (fee and lagged volume). * represents the statistical significance of the p-value at alpha level of 0.05

Table 15 represents the 2SLS regression outcomes of Log volatility on log volume using the one instrument (fee) and two instruments (fee and lagged volume). It can be comprehended that all the trading fee structures are able and efficient predictors of the relationship between the log volatility and log volume. This implies the volume is endogenous for the instruments above.

5.4 Instrumental Variable Estimation

IV estimation is means to accomplish the paper's main aim, i.e., to study the impact of the trading volume on the returns and volatility through the trading fee as an instrument. The paper studies the 2SLS estimation using one instrument (trading fee) and two instruments (trading fee and log volume). The log volume is used as an instrument in the model to reduce the variance of the previous values of the log volume and obtain unbiased estimates. The paper demonstrates the results in four parts: (1) 2SLS estimation of log returns on log volume with trading fee as an instrument (2) 2SLS estimation of log returns on log volume with trading fee and lagged volume as an instrument (3) 2SLS estimation of log volatility on log volume with trading fee as an instrument (4) 2SLS estimation of log volatility on log volume with trading fee and lagged volume as an instrument. Every table has the same respective dependent and independent variable; the only difference is the instrumental variable applied. The paper designs the results table in such a way where the instrument column refers to a specific regression with the different instruments used in the 2SLS process, and the corresponding estimates such as alpha, beta, R-squared and economic significance pertains to the regression using that specific instrument.

5.4.1 2SLS of Log Returns on Log Volume with Trading Fee as an instrument.

Table 16 displays a negative and statistically significant relationship between the log-returns and the trading volume (with trading fee as the only instrument) for the total and makers fee structures. It denotes that with the increase in trading volume with respect to the trading fee, the returns tend to be somewhat constant or drops to a small extent. Nevertheless, a positive relationship between the log returns and trading volume is observed in some categories of the taker's fee such as low volume traders fee (significant at an alpha level of 0.10) and high-volume traders (significant at an alpha level of 0.05), signifying that trading volume drives up the log-returns. The empirical literature demonstrates that the trading volume granger causes the returns. On the other hand, some contradictory studies found no relationship between the two, but no study analyzed the relationship between volume and returns using the change in trading fee as an instrument.

The result suggests a negative impact of volume on the log returns in the presence of the trading fees, which is not a surprise as the returns are affected by numerous other factors, particularly when the assets have an uncertain future. The first primary reason for this negative relationship between the returns and trading volume could be the crash in cryptocurrency around the end of 2017 when the bitcoin touched a high of $20,000, and it tumbled by 85% within one year. Even if the volume increased a bit during that period, the returns did not change much. The second possible explanation for the negative relationship between the returns and trading volume is that the increase in returns through volume could have been neutralized by the trading fee. The third plausible reason for Maker's average fee's negative relationship could be that the makers induce liquidity into the market and deal with decent quantities of financial assets such as cryptocurrency. If there is a high supply of assets or trading volume when the trading fee drops in the market, it will reduce the asset's price, thereby deteriorating the market's returns. The fourth rationale for a negative relationship between the returns and volume may be because the volume may be having an equilibrium level, after which if the volume increases, the price tends to be constant or falls a bit, thereby yielding a bit negative return. However, as discussed, it could likely be one of the numerous reasons. From the table, it can be deduced that the low volume traders (in makers fee and total fees case) have a higher magnitude of a negative relationship between the returns and trading volume than the high-volume traders. It implies that low-volume trader's returns are affected relatively more than the high-volume traders. It may be because of the differential fee charged from them or certain rebates in trading.

Table 16: 2SLS Regression Results for Log Return on Log Volume (One Instrument: Trading fee)

No.	Instruments	α	β (Volume)	R-Squared	Economic Significance
1	**Total Low Volume Fee**	0.0339***	-0.0032***	-0.0054	-0.0115
		(0.0080)	(0.0007)		
2	**Total Median Volume Fee**	0.0183***	-0.0017***	-0.0019	-0.0063
		(0.0050)	(0.0005)		
3	**Total High-Volume Fee**	0.0100**	-0.0010***	-0.0008	-0.0035
		(0.0040)	(0.0004)		
4	**Total Average Volume Fee**	0.0189***	-0.0018***	-0.0020	-0.0065
		(0.0051)	(0.0005)		
5	**Low Trading Volume Fee (Taker)**	-0.0279*	0.0025*	-0.0016	0.0092
		(0.0162)	(0.0015)		
6	**Median Trading Volume Fee (Taker)**	0.0963	-0.0089	-0.0362	-0.0326
		(0.1285)	(0.0119)		
7	**High Trading Volume Fee (Taker)**	-0.0322**	0.0029**	-0.0024	0.0107
		(0.0159)	(0.0015)		
8	**Average Trading Volume Fee (Taker)**	0.4616	-0.0427	-0.7621	-0.1558
		(3.0017)	(0.2773)		
9	**Low Trading Volume Fee (Maker)**	0.0252***	-0.0024***	-0.0032	-0.0086
		(0.0040)	(0.0004)		
10	**Median Trading Volume Fee (Maker)**	0.0165***	-0.0015***	-0.0018	-0.0060

		(0.0031)	(0.0003)		
11	**High Trading Volume Fee (Maker)**	0.0139***	-0.0013***	-0.0014	-0.0051
		(0.0029)	(0.0003)		
12	**Average Trading Volume Fee (Maker)**	0.0186***	-0.0018***	-0.0020	-0.0064
		0.0033	0.0003		

Note: This table presents results of 2SLS regressions of log returns on the log volume, with different trading fee as an instrument. Each row presents results using a different measure of trading fees. Economic significance is computed as the change in the dependent variable for a one-standard-deviation increase in the fee. Standard errors are presented in the parenthesis ** represents p-value < 10%, ** represents p-value < 5%, and *** represents P-value < 1%*. The sample period is from 2015 to 2020.

5.4.2 2SLS of Log Returns on Log Volume with Trading Fee and Lagged Volume as an instrument.

Table 17 demonstrates a negative relationship between returns and the lagged volume. Upon inclusion of lagged volume in the 2SLS model, we note that the coefficients for the taker fee structures also turned out to be negative, but the issue with this model is that the coefficients for all types of fee appear to be negative and the same, which signals that the lagged volume drives the results. Moreover, the negative relationship between the volume and returns could be justified by the Great Crypto Crash, when the traders were conservative, and a little increase in volume did not affect the returns much or even drove it down because of the adverse market sentiments. The results again suggest that there may be an equilibrium level of volume, which, when achieved, stabilizes, or drives the returns in a negative direction to a small extent. The results in Table 17 also shows a significant decrease in the standard errors, which is an obvious implication when the lagged variable is added in the model. The paper also observes traces of negative relationship between the trading volume and returns that is demonstrated in Appendix-A.

Table 17: 2SLS Regression Results for Log Return on Log Volume (Two Instrument: Trading fee and Lagged Volume)

No.	Instruments	α	β (Volume)	R-Squared	Economic Significance
1	**Total Low Volume Fee**	0.0033***	-0.0003***	-0.0002	-0.0012
		(0.0008)	(6.631e-05)		
2	**Total Median Volume Fee**	0.0033***	-0.0003***	-0.0002	-0.0012
		(0.0008)	(6.631e-05)		
3	**Total High-Volume Fee**	0.0033***	-0.0003***	-0.0002	-0.0012

		(0.0008)	(6.631e-05)		
4	**Total Average Volume Fee**	0.0033***	-0.0003***	-0.0002	-0.0012
		(0.0008)	(6.631e-05)		
5	**Low Trading Volume Fee (Taker)**	0.0033***	-0.0003***	-0.0002	-0.0012
		(0.0008)	(6.631e-05)		
6	**Median Trading Volume Fee (Taker)**	0.0033***	-0.0003***	-0.0002	-0.0012
		(0.0008)	(6.631e-05)		
7	**High Trading Volume Fee (Taker)**	0.0033***	-0.0003***	-0.0002	-0.0012
		(0.0008)	(5.502e-05)		
8	**Average Trading Volume Fee (Taker)**	0.0033***	-0.0003***	-0.0002	-0.0012
		(0.0008)	5.502e-05		
9	**Low Trading Volume Fee (Maker)**	0.0033***	-0.0003***	-0.0002	-0.0012
		(0.0008)	5.502e-05		
10	**Median Trading Volume Fee (Maker)**	0.0018***	-0.0002***	-9.843e-05	-0.0007
		0.0006	5.096e-05		
11	**High Trading Volume Fee (Maker)**	0.0018***	-0.0002***	-9.843e-05	-0.0007
		0.0006	5.096e-05		
12	**Average Trading Volume Fee (Maker)**	0.0033***	-0.0003***	-0.0002	-0.0012
		(0.0008)	5.502e-05		

Note: This table presents results of 2SLS regressions of log returns on the log volume, with different trading fee and lagged volume as an instrument. Each row presents results using a different measure of trading fees. Economic significance is computed as the change in the dependent variable for a one-standard-deviation increase in the fee. Standard errors are presented in the parenthesis ** represents p-value < 10%, ** represents p-value < 5%, and *** represents P-value < 1%*. The sample period is from 2015 to 2020.

5.4.3 2SLS of Log Volatility on Log Volume with Trading Fee as an instrument.

Table 18 illustrates a negative relationship between the log volatility and log volume in the total fee and makers fee. In general, when the paper examined the relationship between the log volatility and log volume without the instrument, it was positive and justifiable in case of uncertain asset classes like cryptocurrency. Makers are comparatively stable participants compared to the takers, so if makers increase the trading volume, the volatility will decrease as they have a stable investment system through placing the order in the limit order books and does not suddenly reduce the trading volume on the fee increased. On the other hand, the takers have a positive and significant relationship between trading volume and volatility, implying that the fee decrease encourages them to increase the volume, thereby driving up the market volatility. Nevertheless, when the fee is increased, they are even more conservative in buying as makers tend to extract the fee increase from takers. Therefore, the takers have mixed feelings about market perception, so they are more likely to drive the market volatility, given the fee change. Another possible reason for the negative relationship between volume and volatility could be the variance caused by the volume's past values in the model. The following model that includes the lagged volume as an instrument, can present a better picture of the relationship between the trading volume and volatility.

Table 18: 2SLS Regression Results for Log Volatility on Log Volume (One Instrument: Trading fee)

No.	Instruments	α	β (Volume)	R-Squared	Economic Significance
1	**Total Low Volume Fee**	12.855***	-1.8065***	-5.8033	-6.5505
		(0.3801)	(0.0351)		
2	**Total Median Volume Fee**	1.9565***	-0.8007***	-1.2036	-2.9032
		(0.1167)	(0.0108)		
3	**Total High-Volume Fee**	-2.7493***	-0.3663***	-0.2803	-1.3283
		(0.0759)	(0.0070)		
4	**Total Average Volume Fee**	2.3780***	-0.8396***	-1.3176	-3.0443
		(0.1252)	(0.0116)		
5	**Low Trading Volume Fee (Taker)**	-24.026***	1.5974***	-4.1077	5.7922
		(0.3840)	(0.0355)		
6	**Median Trading Volume Fee (Taker)**	-66.116***	5.4822***	-50.286	19.8739
		(6.6516)	(5.4809)		
7	**High Trading Volume Fee (Taker)**	-10.796***	0.3763***	-0.1869	1.3646
		(0.3173)	(0.0293)		
8	**Average Trading Volume Fee (Taker)**	-45.210***	3.5526 ***	-20.938	12.8818
		(3.0603)	(0.2825)		
9	**Low Trading Volume Fee (Maker)**	3.9906***	-0.9884***	-1.8011	-3.5840
		(0.1187)	(0.0110)		
10	**Median Trading Volume Fee (Maker)**	0.0044***	-0.6205***	-0.7428	-2.2499
		(0.0732)	(0.0067)		
11	**High Trading Volume Fee (Maker)**	-1.9801***	-0.4373***	-0.3874	-1.5858
		(0.0593)	(0.0055)		
12	**Average Trading Volume Fee (Maker)**	0.3865***	-0.6558***	-0.8243	-2.3778
		(0.0760)	(0.0070)		

Note: This table presents results of 2SLS regressions of log volatility on the log volume, with different trading fee. Each row presents results using a different measure of trading fees. Economic significance is computed as the change in the dependent variable for a one-standard-deviation increase in the fee. Standard errors are presented in the parenthesis ** represents p-value < 10%, ** represents p-value < 5%, and *** represents P-value < 1%*. The sample period is from 2015 to 2020.

5.4.4 2SLS of Log Volatility on Log Volume with Trading Fee and Lagged Volume as instrument

The lagged volume as an additional instrument in 2SLS estimation of log volatility on log volume with the trading fee as an instrument yields a positive and statistically significant relationship between the log volatility and log volume illustrated in table 19. The inclusion of lagged volume has corrected the noise/variance from the past log volume values, thereby depicting the plausible beta coefficients that were anticipated. The positive relation concludes that the volume does impact the volatility in the cryptocurrency market in the presence of trading fee. In other words, it can be said that the trading fee affects the volume that further influences the volatility in the market. A small positive r-squared affirm the positive relationship the volume and volatility, but it also indicates that many other factors are not included in the model that also explain volatility.

It can be noticed that the magnitude happens to be more significant for the low volume traders as compared to the high-volume traders, which implies low volume traders tends to increase relatively more volatility in the market than the high-volume traders. It is because the low volume traders are the ones affected more by the increase in trading fees, so they tend to reduce the volume, thereby affecting the volatility more than the high-volume traders. The exact opposite situation happens when the trading fee is reduced. On the other hand, the makers illustrate comparatively less volatility when the volume is increased because of their persistence and systematic way of inducing the liquidity into the markets. The following table provides significant evidence for our central hypothesis of the volume's impact through trading fees on the volatility in the cryptocurrency market. The result is consistent with Lee and Rui, 2002, who finds volume does not granger cause the returns but finds a positive relationship between trading volume and volatility in the stock market.

Table 19: 2SLS Regression Results for Log Volatility on Log Volume (Two Instrument: Trading fee and Lagged Volume)

No.	Instruments	α	β (Volume)	R-Squared	Economic Significance
1	**Total Low Volume Fee**	-6.8048***	0.0080***	0.0010	0.0288
		(0.0113)	(0.0010)		
2	**Total Median Volume Fee**	-6.8000 ***	0.0075***	0.0010	0.0272
		(0.0113)	(0.0010)		
3	**Total High-Volume Fee**	-6.8012 ***	0.0076***	0.0010	0.0276
		(0.0113)	(0.0010)		
4	**Total Average Volume Fee**	-6.8004 ***	0.0076***	0.0010	0.0273
		(0.0113)	(0.0010)		

5	**Low Trading Volume Fee (Taker)**	-6.8151***	0.0089***	0.0011	0.0323
		(0.0113)	(0.0010)		
6	**Median Trading Volume Fee (Taker)**	-6.8091***	0.0084***	0.0011	0.0302
		(0.0113)	(0.0010)		
7	**High Trading Volume Fee (Taker)**	-6.8073***	0.0082***	0.0011	0.0296
		(0.0099)	(0.0009)		
8	**Average Trading Volume Fee (Taker)**	-6.8090***	0.0083***	0.0011	0.0302
		(0.0099)	(0.0008)		
9	**Low Trading Volume Fee (Maker)**	-6.7965***	0.0072***	0.0009	0.0260
		(0.0099)	(0.0009)		
10	**Median Trading Volume Fee (Maker)**	-6.7936***	0.0069***	0.0009	0.0251
		(0.0099)	(0.0009)		
11	**High Trading Volume Fee (Maker)**	-6.7944***	0.0070***	0.0009	0.02540
		(0.0099)	(0.0009)		
12	**Average Trading Volume Fee (Maker)**	-6.7936***	0.0069***	0.0009	0.0251
		(0.0099)	(0.0009)		

Note: This table presents results of 2SLS regressions of log volatility on the log volume, with different trading fee and lagged volume. Each row presents results using a different measure of trading fees. Economic significance is computed as the change in the dependent variable for a one-standard-deviation increase in the fee. Standard errors are presented in the parenthesis ** represents p-value < 10%, ** represents p-value < 5%, and *** represents P-value < 1%*. The sample period is from 2015 to 2020.

6. Robustness Checks

Chapter 6 will validate the robustness of the results that were discussed in chapter 5. One way is to check the results' robustness is to analyze the effect of volume n return and volatility in different samples, using the same approach, i.e., the IV estimation. This section explores the robustness of the results through sampling based on two dimensions, which are as follows:

i. The first type of subsample considered is based on exchange type, i.e., Crypto-only exchange and fiat exchange (discussed below).

ii. The second type of subsamples considered is based on the date. The entire data is divided into 1st January 2015 to 15th December 2017 and 16th December 2017 to 31st May 2020.

The robustness is checked only on the Total Average Trading Fee, Maker's Average Trading Fee, and Taker's Average trading fee. The paper performs IV estimation on all the subsamples through the Two-Stage Least Square method (2SLS). The 2SLS is computed for (1) log returns on log volume and (2) log volatility on log volume, with trading fees and lagged volume as instruments for both. We computed the robustness check for one instrument as well that is illustrated in Appendix-B. The following part discusses the results of the 2SLS performed on the two dimensions of the sampling: (6.1) Sampling on the basis exchange type (6.2) Sampling-based on the date.

6.1 Subsamples based on the Exchange Type

To compute the robustness check, we divide the entire data set into two parts based on the exchange type. There are two types of exchanges in the cryptocurrency domain: The Crypto-only exchange and the fiat exchange. In Crypto-only exchange, only those crypto pairs are traded that cannot be bought with government-regulated currencies such as btc/eth, eth/btc or xrp/btc. On the other hand, the fiat exchanges allow trade in cryptocurrency through the government-regulated currency (such as USD or CAD) and standard Crypto-only pairs such as btc/usd, eth/usd or eth/btc. Out of all the fifteen exchanges, the paper finds five Crypto-only exchanges and ten fiat exchanges. We compute the 2SLS, based on the exchange type. The following are the results of the IV estimation using two instruments on the sub-samples:

6.1.1 2SLS of Log-Returns on Log Volume with Trading fee and Lagged Volume.

Table 20 illustrates that the negative and insignificant relationship between the log returns and log volume while using trading fees and lagged volume as instruments for both types of exchanges, i.e., Fiat and Crypto-only exchange. This result is in accordance with the results yielded in the result section 5 (table 17) when the 2SLS method was implemented on the entire data results. However, the negative relation between the returns and volume in the robustness checking does not appear to be significant, as in the entire data.

Table 20: 2SLS of Log Return on Log Volume of subsample

No.	Particulars	α		β (Volume)		R-Squared		Economic Significance	
		Fiat Exch.	Crypto Exch.	Fiat Exch.	Crypto Exch.	Fiat Exch.	Crypto Exch.	Fiat Exch.	Crypto Exch.
1	**Total Average Volume Fee**	0.0015 (0.0009)	-0.0002 (0.0014)	-8.11e-05 (8.513e-05)	-0.0002 (0.0001)	-3.708e-05	-0.0002	-0.0003	-0.0002
2	**Average Trading Volume Fee (Taker)**	0.0015 (0.0009)	-0.0002 (0.0014)	-8.024e-05 (8.512e-05)	-0.0002 (0.0001)	-3.667e-05	-0.0002	-0.0003	-0.0002
3	**Average Trading Volume Fee (Maker)**	0.0015 (0.0009)	-0.0002 (0.0014)	-8.141e-05 (8.513e-05)	-0.0001 (0.0001)	-3.724e-05	-0.0002	-0.0003	-0.0002
	Total Observation	**655461**	**312308**						

Note: This table presents results of 2SLS regressions of log returns on the log volume, with different trading fee and lagged volume. Each row presents results using a different measure of trading fees. The crypto exch. stands for Crypto-only exchange. Economic significance is computed as the change in the dependent variable for a one-standard-deviation increase in the fee. Standard errors are presented in the parenthesis ** represents p-value < 10%, ** represents p-value < 5%, and *** represents P-value < 1%*. The sample period is from 2015 to 2020.

The results demonstrate that the fiat exchanges have got less negative coefficient for the volume than the Crypto-only exchange. It implies that the volume has a lesser degree of a negative relationship with fiat exchanges' returns as they experience high market participation because of its openness to accept the government's currency, which is readily available with retail or prominent institutional participants. Whereas, in crypto-only exchange, you need to have existing crypto pairs (transferred from the fiat exchange) to trade, which is a complicated and inconvenient way. Maybe due to high market participation, we observe a less negative relation between the fiat exchanges' returns and volume and greater negative relation between the crypto-only exchanges' returns and volume.

6.1.2 2SLS of Log-Returns on Log Volatility with Trading Fee and Lagged Volume.

Table 21 presents a positive and statistically significant relationship between volume and volatility with the trading fee and lagged volume as instruments. A positive relationship between the two suggests that as the volume surges, the volatility tends to increase. Cryptocurrency is an uncertain financial asset, so it is expected that the volume will affect its volatility. The outcome presented in table 21 for the robustness check is consistent with the outcome of the entire database presented in table 19.

Table 21: 2SLS of Log Volatility on Log Volume of subsample

No.	Particulars	α		β (Volume)		R-Squared		Economic Significance	
		Fiat Exch.	Crypto Exch.	Fiat Exch.	Crypto Exch.	Fiat Exch.	Crypto Exch.	Fiat Exch.	Crypto Exch.
1	**Total Average Volume Fee**	-6.7165*** (0.0124)	-7.7334*** (0.0277)	0.0151*** (0.0011)	0.0583*** (0.0021)	0.0023	0.0117	0.0576	0.2225
2	**Average Trading Volume Fee (Taker)**	-6.7179*** (0.0124)	-7.7313*** (0.0277)	0.0152*** (0.0011)	0.0582*** 0.0021	0.0023	0.0117	0.0581	0.2218
3	**Average Trading Volume Fee (Maker)**	-6.7162*** (0.0124)	-7.7454*** (0.0277)	0.0151*** (0.0011)	0.0593*** (0.0021)	0.0023	0.0118	0.0575	0.2263
	Total Observation	**619973**	**293381**						

Note: This table presents results of 2SLS regressions of log volatility on the log volume, with different trading fee and lagged volume. Each row presents results using a different measure of trading fees. The crypto exch. stands for Crypto-only exchange. Economic significance is computed as the change in the dependent variable for a one-standard-deviation increase in the fee. Standard errors are presented in the parenthesis ** represents p-value < 10%, ** represents p-value < 5%, and *** represents P-value < 1%*. The sample period is from 2015 to 2020.

Upon analyzing the results from table 21, it is observed that the coefficients for volume in Crypto-only exchanges are more significant than that of the fiat exchange. A probable reason behind such difference is that the Crypto-only exchanges encounter fewer traders in comparison to the fiat exchanges, so whenever there is an increase in volume due to change in fee or positive crypto news, traders tend to trade more than they do (because of the complexity), thereby driving the volatility. Whereas, for the fiat exchange, the volume remains relatively stable than the Crypto-only exchange, thereby resisting extreme volatility spike in the exchange. Through bifurcating the exchanges based on fiat and Crypto-only, it can be observed that the magnitude of the volume coefficient in the sub-samples (of both types of exchanges) is greater than those of the main results coefficient.

6.2 Subsamples based on Date

The second way to check the results' robustness is by dividing the entire database based on the date. For this purpose, the paper aims to select a date when the cryptocurrency market goes through the rough phase. The relation between the fee, volume, return, and volatility can be studied in normal and abnormal circumstances. It will help to investigate the consistency in the relationship between the factors mentioned above for the before and after the particular date. We chose December 15, 2017, as the date, when the bitcoin reaches its all-time high to approximately $20,000. As after that date, the bitcoin tumbled severely. This phase was known as the Great Crypto Crash. The Great Crypto Crash led to a massive sell-off in the cryptocurrencies. Post this date, the bitcoin experienced major shocks and saw a decrease in value by 45% in just a week after it touched the record high of $20,000, which fell to $11,000. From January 6 to February 6, 2017, it again saw a steep decline in the value by falling by 65% within a month. By September 2018, the total bitcoin value got declined by 80% from its peak price. The crypto market lost $640 billion, a greater bubble than the Dot-com bubble of 2002 in percentage terms. The MVIS CryptoCompare Digital Assets 10 Index, an index representing the top-performing cryptocurrencies also fell sharply by 80% of its value, demonstrating its impact on the entire cryptocurrency market.

6.2.1 2SLS of Log-Returns on Log Volume with Trading Fee and Lagged Volume.

Table 22 presents the results of IV estimation for the before and after December 15, 2017. We observed a negative and significant relationship between the log returns and log volume with trading fees and lagged volume as instruments for both the periods. The negative relationship between the two of this sub-sample presented in table 20 is consistent with the results derived from the entire set presented in Table 19. The negative relation suggests that the returns tend to fall as the volume increases. It possibly indicates that as the volume reaches the equilibrium level, any additional trading volume either stabilizes the returns or drives it to a negative extent. Apart from the aforementioned reason, there can be many other reasons that affect market returns, particularly in the crypto market, which is highly unpredictable.

Table 22: 2SLS of Log Return on Log Volume of subsample

No.	Particulars	α		β (Volume)		R-Squared		Economic Significance	
		Before	After	Before	After	Before	After	Before	After
1	**Total Average Volume Fee**	0.0037** (0.0018)	0.0031*** (0.0008)	-0.0004** (0.0002)	-0.0003*** (6.875e-05)	-0.0002	-0.0002	-0.001	-0.001
2	**Average Trading Volume Fee (Taker)**	0.0037** (0.0018)	0.0031*** (0.0008)	-0.0004** (0.0002)	-0.0003*** (6.875e-05)	-0.0002	-0.0002	-0.001	-0.001
3	**Average Trading Volume Fee (Maker)**	0.0037** (0.0018)	0.0032*** (0.0008)	-0.0004** (0.0002)	-0.0003*** (6.875e-05)	-0.0002	-0.0002	-0.001	-0.001
	Total Observation	**643871**	**251931**						

Note: This table presents results of 2SLS regressions of log returns on the log volume, with different trading fee and lagged volume. Each row presents results using a different measure of trading fees. The before and after refers to the results for the data before the 15th December 2017 and after represents the results for the data after 15th December 2017. Economic significance is computed as the change in the dependent variable for a one-standard-deviation increase in the fee. Standard errors are presented in the parenthesis ** represents p-value < 10%, ** represents p-value < 5%, and *** represents P-value < 1%*. The sample period is from 2015 to 2020.

6.2.2 2SLS of Log Volatility on Log Volume with Trading Fee and Lagged Volume.

Table 23 illustrates the IV estimation of log volatility on log volume with trading fees and lagged volume as instruments. We observe a positive and significant relationship between the volatility and volume before the crash period and a negative and significant relation after the crash period. The results for the crash period, illustrated in table 23, are parallel with those of the entire database presented in table 19. But the results for after the crash period demonstrate a negative relationship between the volatility and volume in the presence of instruments. The negative relation implies that if the volume decreases, then the volatility will surge. The negative relation is anticipated in a financially sound asset or when the asset market is going through a distressing period. Even with good volume, the sound assets do not show much volatility in the returns, but uncertain asset class volatility increases with a spike in volume.

Table 23:2SLS of Log Volatility on Log Volume of subsample

No.	Particulars	α		β (Volume)		R-Squared		Economic Significance	
		Before	After	Before	After	Before	After	Before	After
1	**Total Average Volume Fee**	-7.169***	-6.682***	0.0692***	-0.013***	0.0137	-0.001	0.2482	-0.048
		(0.0251)	(0.0121)	(0.0021)	(0.0010)				
2	**Average Trading Volume Fee (Taker)**	-7.171***	-6.699***	0.0693***	-0.011***	0.0137	-0.001	0.2487	-0.043
		(0.0251)	(0.0121)	(0.0021)	(0.0010)				
3	**Average Trading Volume Fee (Maker)**	-7.169***	-6.669***	0.0691***	-0.014***	0.0137	-0.001	0.2480	-0.052
		(0.0251)	(0.0121)	(0.0021)	(0.0010)				
	Total Observation	**643871**	**251931**						

Note: This table presents results of 2SLS regressions of log returns on the log volume, with different trading fee and lagged volume. Each row presents results using a different measure of trading fees. The before and after refers to the results for the data before the 15th December 2017 and after represents the results for the data after 15th December 2017. Economic significance is computed as the change in the dependent variable for a one-standard-deviation increase in the fee. Standard errors are presented in the parenthesis ** represents p-value < 10%, ** represents p-value < 5%, and *** represents P-value < 1%*. The sample period is from 2015 to 2020.

Another reason for having a negative relationship between the volume and volatility is when the asset market is going through tough times, as in the cryptocurrency case represented by the after-crash period. The after-crash period includes the Great Crypto Crash, which could be a major reason for this negative relationship. The volume at these distress times falls significantly, so the asset price continuously tumbles, thereby increasing the volatility of returns in the market. This could be a probable reason for the negative relation derived between volume and volatility from the after-crash period. Otherwise, in general, we anticipate a positive relationship between the trading volume and volatility in the cryptocurrency market, which is highly unpredictable domain.

7. Conclusion

The study analyzed the causal relationship between the trading volume, returns, and volatility in the cryptocurrency market. The relationship mentioned above is documented in the cryptocurrency market by Balcilar (2017) and Bouri et al. (2019), yet, it has only been restricted to specific cryptocurrencies. Our research extends to approximately 2,300 crypto pairs traded over 15 reputed exchanges. The study is the first of its kind in the cryptocurrency domain to use the trading fee changes as an instrument for trading volume, thereby analyzing the impact of resultant volume on returns and volatility. To study the comprehensive relationship and capture the oblique causal effect of fee on the trading volume, returns, and volatility, we used the Two-Stage Least square method of the Instrumental Variable estimation approach. The model also comprises the lagged volume as one of the instruments, as it minimizes the variance of the volume's past values, thereby yielding more accurate results and lowering the standard errors.

The paper observed a statistically significant negative relationship between the trading volume and returns. The result contradicts the literature that suggests a positive relationship between the trading volume and returns, such as in the study of Bouri et al. (2019) and Balcilar (2017). However, the model evaluated in the paper includes the trading fee and studies numerous cryptocurrency pairs, so variation in the result is quite possible. The negative affiliation between the trading volume and returns indicates that the increase in return through the surge in volume could have been neutralized by the trading fees. Other potential reasons can also justify the negative relationship between the trading volume and returns. The study discovers a positive and statistically significant relationship between trading volume and volatility. The significant positive relationship is key-findings in this research paper, as it is coherent with the empirical literature. The paper by Lee and Rui, 2002, and Admati and Pfleiderer (1988) also demonstrates a positive relationship between trading volume and volatility. The result is more prominent as the model includes the trading fee factor, which affirms that fee impacts volatility through the trading volume.

To ensure the precision of the paper's results, we performed robust checks. The checks were based on evaluating the 2SLS model on two sub-samples based on the date and exchange type. The sub-sample divided on the basis of exchange, i.e., fiat exchanges and Crypto-only exchanges, yield a negative relationship between the trading volume and returns and a positive relationship between trading volume and volatility. The results for both types of exchanges are consistent with our main results. The data divided based on date (15th December 2017) yields positive relationship trading volume and volatility before the date sample and after the date sample. However, we did not find coherent results for the relationship between the trading volume and volatility for after the date sample with our primary results, as the cryptocurrencies fell severely after that date. Except for one peculiar situation, our robust test confirms the accuracy of our paper's result.

The contradictory result for the relationship between the trading volume and returns opens the avenues for future research, exploring the probable factors that eventually influence the cryptocurrency's returns. Hence, understanding the dynamics of the cryptocurrency markets and exploring the latent aspects can significantly transform the investing pattern and attract more conservative traders in the cryptocurrency market.

Bibliography

- Acharya, V. and L. Pedersen (2005). "Asset pricing with liquidity risk", Journal of Financial Economics 77 (2), 375-410.

- Admati, A. R. and P. Pfleiderer (1988), "A Theory of Intraday Patterns: Volume and Price Variability", Review of Financial Studies 1 (1), 3-40.

- Amihud, Y. and H. Mendelson (2018), "Liquidity and Stock Returns", Financial Analysts Journal 42 (3), 43-48.

- Angel, J. J., L. E. Harris, and C. S. Spatt (2011), "Equity Trading in 21st Century", Quarterly Journal of Finance 1 (1), 1-53.

- Balcilar, M., E. Bouri, R. Gupta, and D. Roubaud (2017), "Can Volume Predict Bitcoin Returns and Volatility? A Quantiles-Based Approach", Economic Modelling 64 (C), 74-81.

- Balduzzi, P., T. C. Green and E. J. Elton (1998). "Economic News and the Yield Curve: Evidence from the U.S. Treasury Market, New York University", Leonard N. Stern School Finance Department Working Paper Series.

- Barclay, M. J., E. Kandel, and L. M. Marx (1998), "The Effect of Transaction Cost on Stock Prices and Trading Volume", Journal of Financial Intermediation 7, 130-150.

- Bascle, G. (2008), "Controlling for endogeneity with instrumental variables in strategic management research", Strategic Organization 6 (3), 285-327

- Blume, L., D. Easley and M. O. Hara (1994). "Market Statistics and Technical Analysis: The Role of Volume', Journal of Finance 49 (1), 153-81.

- Bouoiyour, J. and R. Selmi (2015), “What Does Bitcoin Look Like?”, Annals of Economics and Finance 16 (2), 449-492.

- Bouri, E., C. K. M. Lau, B. Lucey, and D. Roubaud (2019), “Trading volume and the predictability of return and volatility in the cryptocurrency market”, Finance Research Letters 29, 340-346.

- Bouri, E., P. Molnár, G. Azzi, D. Roubaud, and L. I. Hagfors (2017), “On the hedge and safe haven properties of Bitcoin: Is it really more than a diversifier?”, Finance Research Letters 20 (C), 192-198.

- Campbell, J. Y, S. J. Grossman, and J. Wang (1993), “Trading Volume and Serial Correlation in Stock Returns”, Quarterly Journal of Economics 108 (4), 905-939.

- Cepoi, C. O. (2014), “How Trading Costs Affect Liquidity on Bucharest Stock Exchange?”, Procedia Economics and Finance 5, 1153-1158.

- Chao, Y., C. Yao, and M. Ye (2017), “Discrete Pricing and Market Fragmentation: A Tale of Two-Sided Markets”, American Economic Review 107 (5), 196-199.

- Chen, G., M. Firth and O. M. Rui (2005). “The Dynamic Relation Between Stock Returns, Trading Volume, and Volatility, The Financial Review 36 (3), 153-174.

- Chiarella, C., B. Kang, C. S. Nikitopoulos (2015). “The Return–Volatility Relation in Commodity Futures Markets”, Journal of Futures Markets 36 (2), 127-152.

- Ciaian, P., M. Rajcaniova, and A. Kancs (2016), “The economics of BitCoin price formation”, Applied Economics 48 (19), 1799-1815.

- Constantinides, G. M (1986), “Capital Market Equilibrium with Transaction Costs”, Journal of Political Economy 94 (4), 842-862.

- Fleming, M. J. and E. M. Remolona (1999), “Price Formation and Liquidity in the U.S. Treasury Market: The Response to Public Information”, The Journal of Finance 54(5), 1901-1915.
- Comerton-Forde, C. C., V. Grégoire, and Z. Zhong, “Inverted fee structures, tick size, and market quality”, Journal of Financial Economics 134 (1), 141-164.
- Foster, F. D, and S. Viswanathan (1993). “Variations in Trading Volume, Return Volatility, and Trading Costs: Evidence on Recent Price Formation Models”, The Journal of Finance 48 (1), 187-211.
- Foucault, T., O. Kadan, and E. Kandel, “Liquidity Cycles and Make/Take Fees in Electronic Markets”, The Journal of Finance 68 (1), 299-341.
- Gebka, B. and M. Wohar (2013). “Causality between trading volume and returns: Evidence from quantile regressions”, International Review of Economics and Finance 27 (C), 144-159.
- Glaser, F., K. Zimmermann, M. Haferkorn, M. C. Weber, and M. Siering (2014), “Bitcoin - Asset or Currency? Revealing Users' Hidden Intentions”, ECIS 2014.
- Hanley, B. (2015), “The False Premises and Promises of Bitcoin”, Cornell University (8).
- Harris, L. (2013), “Maker-Taker Pricing Effects on Market Quotations”, USC Marshall School of Business.
- Harris, M. and A. Raviv (2015). “Differences of Opinion Make a Horse Race”, The Review of Financial Studies 6 (3), 473-506.
- Hayes, A. (2016), “Cryptocurrency Value Formation: An Empirical Analysis Leading to a Cost of Production Model for Valuing Bitcoin”, Telematics and Informatics.

- Kadlec, G. B. and J. J. McConnell (1994), "The Effect of Market Segmentation and Illiquidity on Asset Prices: Evidence from Exchange Listings", The Journal of Finance 49 (2), 611-636.

- Karpoff, J. M. (1987). "The Relation Between Price Changes and Trading Volume: A Survey", Journal of financial and Quantitative Analysis 22 (1), 109-126.

- Kouwenberg, R., T. Vorst, and W. M. D. Monique (2000), "Options and Earnings Announcements: An Empirical Study of Volatility, Trading Volume, Open Interest and Liquidation", European Financial Management 6 (2), 149-171.

- Kramer, C. (1999), "Noise trading, transaction costs, and the relationship of stock returns and trading volume", International Review of Economics and Finance 8 (4), 343-362.

- Kristoufek, L. (2013), "BitCoin meets Google Trends and Wikipedia: Quantifying the relationship between phenomena of the Internet era", Scientific Reports 3, Article number: 3415.

- Kristoufek, L. (2015), "What Are the Main Drivers of the Bitcoin Price? Evidence from Wavelet Coherence Analysis", PLoS ONE 10 (4).

- Lee, B. S. and O. M. Rui (2002), "The dynamic relationship between stock returns and trading volume: Domestic and cross-country evidence", Journal of Banking and Finance 26 (1), 51-78.

- Lee, C. F and O. M. Rui (2000). "Does Trading Volume Contain Information to Predict Stock Returns? Evidence from China's Stock Markets", Review of Quantitative Finance and Accounting 14, 341-360.

- Lee, C. M. C and B. Swaminathan (2002). "Price Momentum and Trading Volume", The Journal of Finance 55 (5), 2017-2069.

- Llorente, G., R. Michaely, G. Saar, and J. Wang (2002), "Dynamic Volume-Return Relation of Individual Stocks", Review of Financial Studies 15 (4), 1005-1047.

- Louhichi, W. (2011), "What drives the volume-volatility relationship on Euronext Paris?", International Review of Financial Analysis 20 (4), 200-206.

- Maheshwari, S. and R. S. Dhankar (2017). "Profitability of Volume-based Momentum and Contrarian Strategies in the Indian Stock Market", Global Business Review 18 (4), 974-992.

- Michaely, R. and J. L. Vila (1996), "Trading Volume with Private Valuation: Evidence from The Ex-Dividend Day", The Review of Financial Studies 9 (2), 471-509.

- Polasik, Michal, A. I. Piotrowska, T. P. Wisniewski, R. Kotkowski, and G. Lightfoot (2015), "Price Fluctuations and the Use of Bitcoin: An Empirical Inquiry", International Journal of Electronic Commerce 20 (1), 9-49.

- Sarwar, G. (2007). "The informational role of option trading volume in the S&P 500 futures options markets", Applied Financial Economics 16 (14), 1197-1210.

- Todorova, N. and M. Souček (2014). "The impact of trading volume, number of trades and overnight returns on forecasting the daily realized range, Economic Modelling 36(C), 332-340.

- Vaynos, D. (1998), "Transaction Costs and Asset Prices: A Dynamic Equilibrium Model", Review of Financial Studies 11(1), 1-58.

- Yermack, D. (2014), "Is Bitcoin a Real Currency? An economic appraisal", New York University (NYU) - Stern School of Business.

Additional Web Sources

- 7 Best Cryptocurrency Exchanges in 2020, Coin Gecko, retrieved July 15, 2020 from https://www.coingecko.com/buzz/seven-most-reliable-cryptocurrency-exchanges

- Bitwise Asset Management Report, retrieved July 15, 2020 from https://www.sec.gov/comments/sr-nysearca-2019-01/srnysearca201901-5164833-183434.pdf

- Bitwise: Nearly 95% of reported bitcoin trading volume are artificially created by unregulated exchanges, Blackrock, retrieved July 15, 2020 from https://www.theblockcrypto.com/linked/16958/bitwise-nearly-95-of-reported-bitcoin-trading-volume-are-artificially-created-by-unregulated-exchanges

- Majority of bitcoin trading is a hoax, new study finds, CNBC, retrieved July 15, 2020 from https://www.cnbc.com/2019/03/22/majority-of-bitcoin-trading-is-a-hoax-new-study-finds.html

- Popper, N. Book (2015), New York Times (Review), retrieved July 15, 2020 from https://www.nytimes.com/2015/07/05/books/review/nathaniel-poppers-digital-gold-looks-at-bitcoin.html

- Takashi Yamano. Instrumental Variable Estimation, retrieved July 17, 2020 from http://www3.grips.ac.jp/~yamanota/Lecture Note 8 to 10 2SLS & others.pdf

Appendix-A

Upon observing the negative relationship between the trading volume and returns, we observed the data graphically to check the visual relationship of the variables mentioned above. The paper presents the graphical representation of returns and volume for the five most traded cryptocurrency pairs.

1. BTC/USD

Figure 1: Time Series of Log Returns and Log Volume for BTC/USD

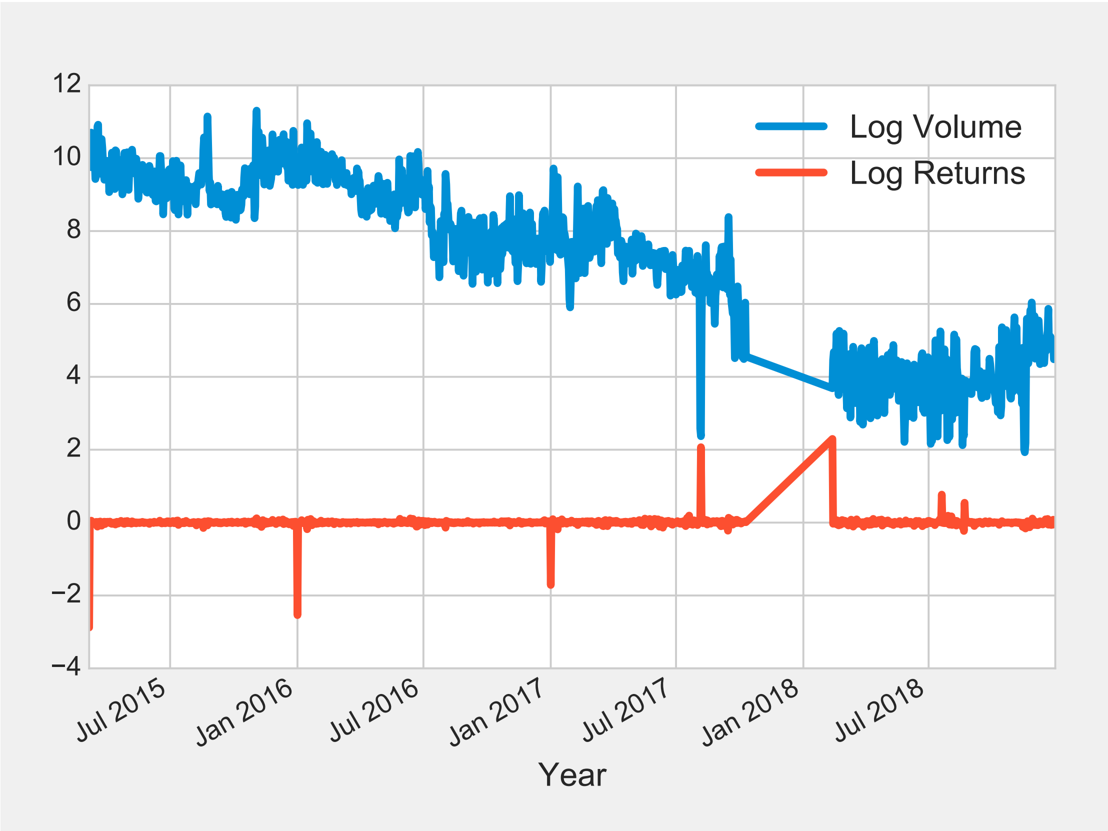

Figure 1 illustrates the time series of log volume and log returns for the BTC/USD pair traded in the Okcoin exchange. We discover traces of the negative relationship between the log returns and log volume. The first evidence is observed around August 2017, when the BTC/USD volume drops, the returns surges. The second significant evidence is witnessed around October 2017, and we see a significant negative relation between trading volume and returns for approximately five straight months. We observed similar traces for BTC/USD in other exchanges as well.

2. ETH/BTC

Figure 2: Time Series of Log Returns and Log Volume for ETH/BTC

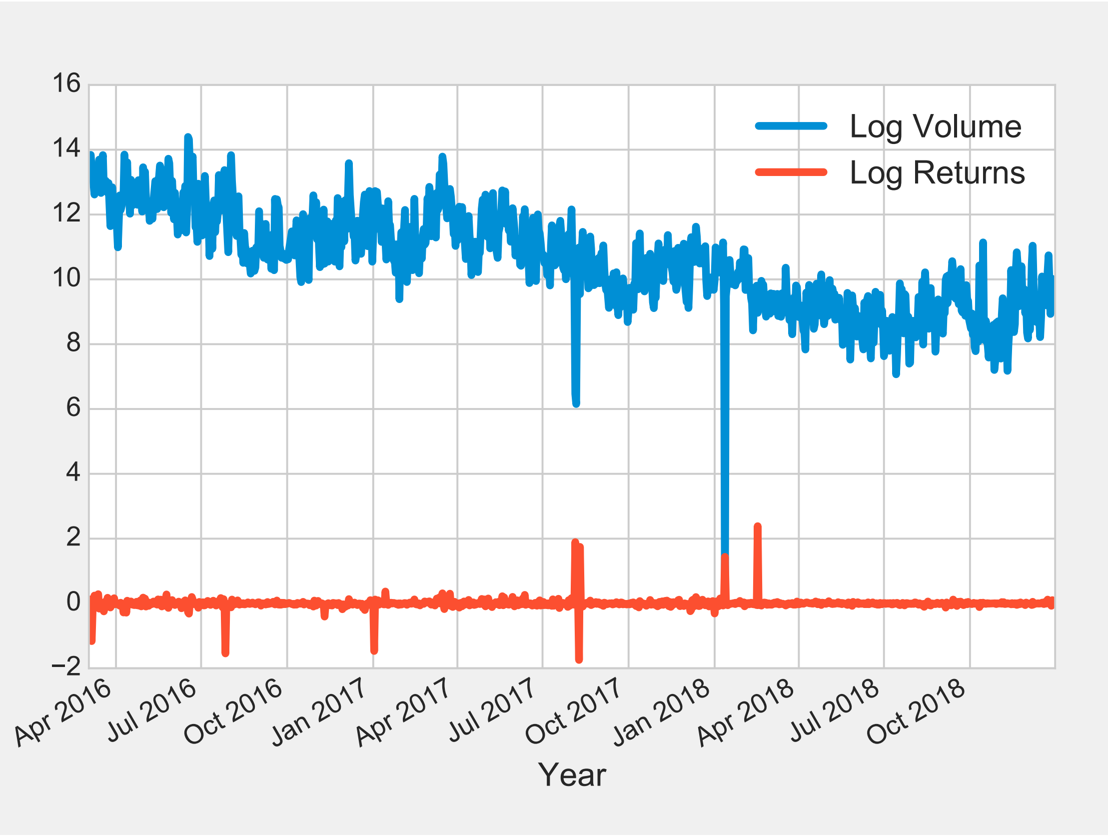

Figure 2 portrays the time series of log volume and log returns for ETH/BTC pair traded in the Kraken exchange. Again, we observe evidence of a negative relationship between volume and returns for ETH/BTC pair. The first indication can be observed in August 2017, when ETH/BTC volumes shrink, the return increases. The second proof for a negative relationship can be observed around January 2018, when the volume tumbles significantly, and the returns shoot up. A similar relationship has been observed for this particular crypto pairs across other exchanges.

3. <u>LTC/BTC</u>

Figure 3: Time Series of Log Returns and Log Volume for LTC/BTC

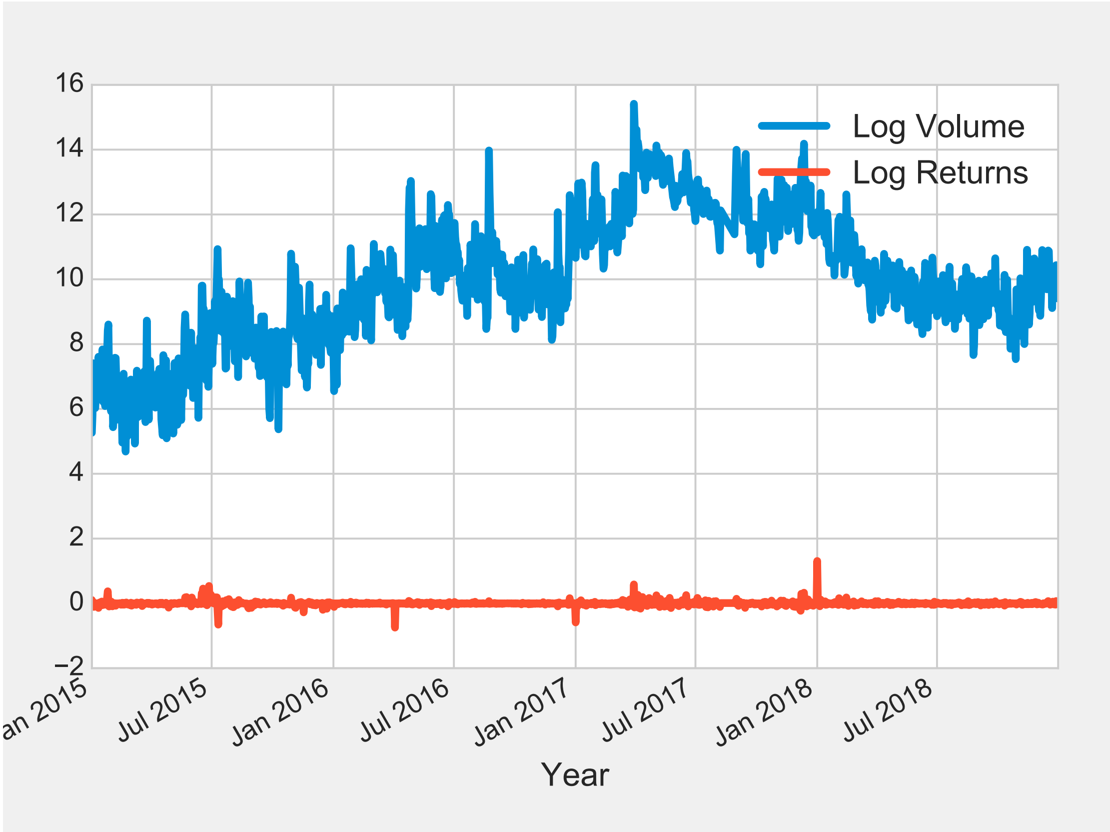

Figure 3 illustrates the time series of log returns and log volume for LTC/BTC traded in the Poloniex exchange. The graphical representation for LTC/BTC does not show any traces of a negative relationship between trading volume and returns, except one around January 2018. However, the figure supports two beliefs: (1) First, the increase in volume does not have much impact on the returns, as the returns tend to be consistent; (2) Second, the change in volume does not affect the inexpensive cryptocurrency pairs much, as we do not observe significant pattern between the trading volume and returns in case of LTC/BTC. Being an inexpensive cryptocurrency, LTC/BTC experienced a better volume than Bitcoin and Ethereum.

4. XRP/BTC

Figure 4: Time Series of Log Returns and Log Volume for XRP/BTC

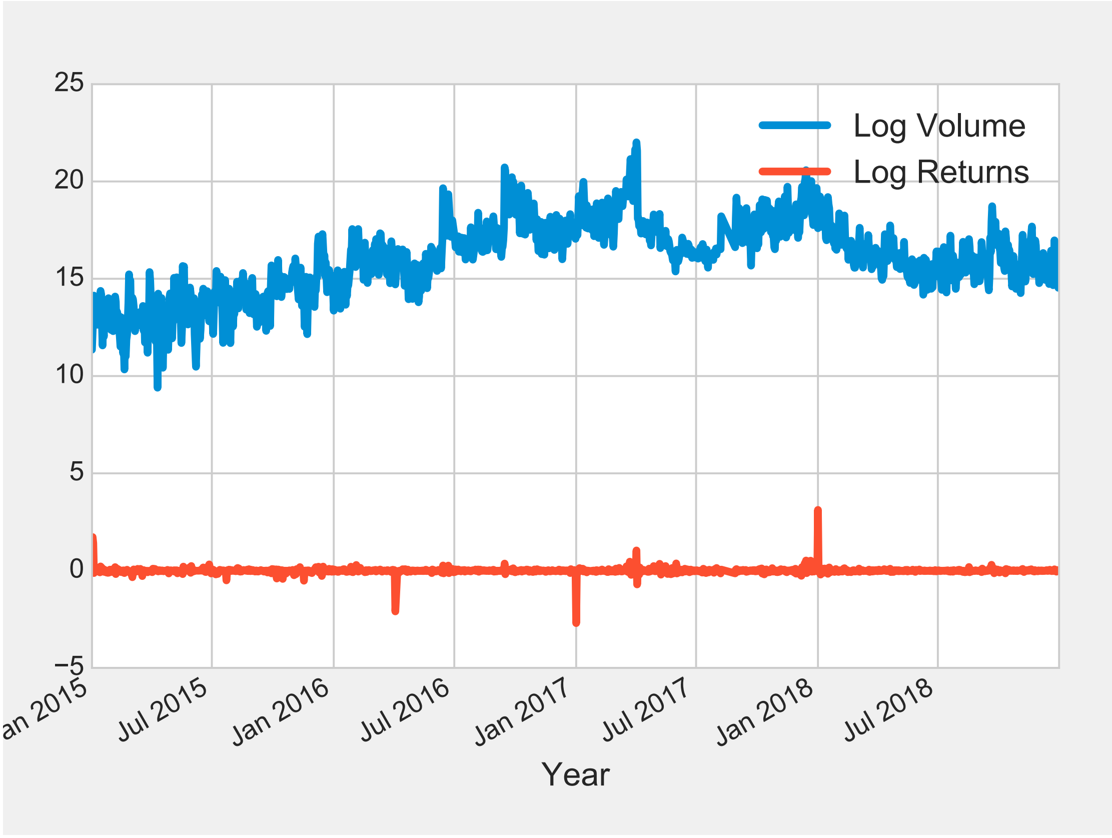

Figure 4 illustrates the time series of log returns and log volume for XRP/BTC traded in the Poloniex exchange. This graphical representation also does not tell much about the relationship between the log volume and log returns. We observe that returns are less volatile than the volume same as in the case of LTC/BTC. The inexpensive currency as XRP/BTC observed the highest trading volume among the top five traded crypto exchanges.

5. LTC/USD

Figure 5: Time Series of Log Returns and Log Volume for XRP/BTC

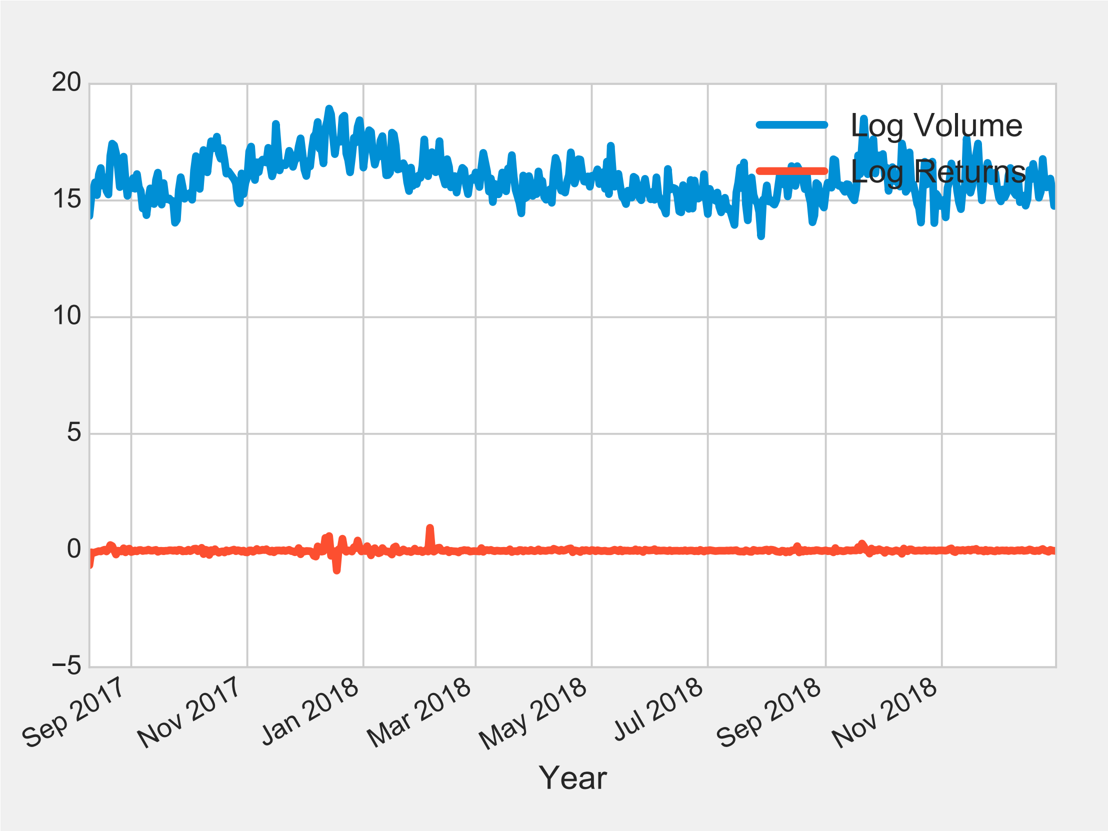

Figure 5 shows the time series of log returns and log volume for LTC/USD traded in the Bitfinex exchange. We cannot infer any significant relationship between volume and return from the graphical representation. However, we do notice high trading volume and returns are comparatively persistent than other pairs except XRP/BTC.

The graphical representation provided us with some clue about the negative relationship between the trading volume and volatility for crypto-pairs such as BTC/USD and ETH/BTC but not for other three top traded crypto-pairs.

Appendix-B

Appendix-B is an extension of the robustness check chapter that illustrates and explains the robustness checks along two dimensions. The section explores the similar IV estimation's 2SLS approach, which is modified and tests only one instrument instead of two instruments, i.e., the trading fee only. We will demonstrate the 2SLS for both the subsamples based on the date and exchange type for log return on log volume and log volatility on log volume. The following is the order:

1. **Sampling based on the Exchange Type (Fiat Exchange type and Crypto-only exchange type)**
   a. **2SLS of Log-Returns on Log Volume with Trading Fee as an instrument.**

Table 24 illustrates the negative and significant relationship between the log returns and log volume with trading fees as instruments for Fiat exchange. This result is in accordance with the results yielded in the result section (table 16). However, we notice a positive relationship between volume and returns for the Total Average volume fee and Maker's average fee (insignificant as well) in Crypto-only exchange type.

Table 24: 2SLS of Log Return on Log Volume of subsample

No.	Particulars	α		β (Volume)		R-Squared		Economic Significance	
		Fiat Exch.	Crypto Exch.	Fiat Exch.	Crypto Exch.	Fiat Exch.	Crypto Exch.	Fiat Exch.	Crypto Exch.
1	**Total Average Volume Fee**	0.0133*** (0.0036)	-0.0030 0.0041	-0.0012*** (0.0003)	0.00008 (0.0003)	-0.0011	7.379e-05	-0.0047	0.0003
2	**Average Trading Volume Fee (Taker)**	0.0111 *** (0.0036)	0.0035 (0.0032)	-0.0010 *** (0.0003)	-0.0005* (0.0003)	-0.0008	-0.0006	-0.0039	-0.0017
3	**Average Trading Volume Fee (Maker)**	0.0142*** (0.0036)	-0.012** (0.0058)	-0.0013*** (0.0003)	0.0009* (0.0005)	-0.0012	0.0003	-0.0051	0.0034
	Total Observation	**655461**	**312308**						

Note: This table presents results of 2SLS regressions of log returns on the log volume, with different trading fee as instrument. Each row presents results using a different measure of trading fees. The crypto exch. stands for Crypto-only exchange. Economic significance is computed as the change in the dependent variable for a one-standard-deviation increase in the fee. Standard errors are presented in the parenthesis ** represents p-value < 10%, ** represents p-value < 5%, and *** represents P-value < 1%*. The sample period is from 2015 to 2020.

The positive relationship between the volume and returns is anticipated according to the literature as well. The positive relationship for Maker's average trading fee suggests that an increase in makers volume will drive the returns up. However, we do not observe such a relationship with any of the other estimates. Most of our results presented in table 24 are consistent with the study's primary results, which ensures our main results' accuracy.

b. 2SLS of Log-Volatility on Log Volume with Trading Fee as an instrument.

Table 25 illustrates that the negative and significant relationship between the log returns and log volume in the presence of trading fees as instruments for Fiat exchange and Crypto-only exchange (except for Total Average volume fee in Crypto-only exchange). We observed a positive relationship between volume and volatility for the Taker's average fee in the main result presented in table 18, but in the following table 25, we notice a negative relationship between them for both the exchange type. The probable reason for a negative relationship between the volume and volatility in most cases could be due to the variance caused by the volume's past values as we observed the positive relationship between the volume and the volatility after the inclusion of the lagged volume. However, we found most of the results consistent with our main results presented in table 18.

Table 25: 2SLS of Log Volatility on Log Volume of subsample

No.	Particulars	α		β (Volume)		R-Squared		Economic Significance	
		Fiat Exch.	Crypto Exch.	Fiat Exch.	Crypto Exch.	Fiat Exch.	Crypto Exch.	Fiat Exch.	Crypto Exch.
1	Total Average Volume Fee	0.1696*** (0.0730)	-3.356*** (0.0728)	-0.660*** (0.0071)	0.3033*** (0.0059)	-0.924	-0.2615	-2.520	-1.1566
2	Average Trading Volume Fee (Taker)	-0.268*** (0.0716)	-4.160*** (0.0566)	-0.617*** (0.0070)	-0.236*** (0.0045)	-0.815	-0.1754	-2.356	-0.9032
3	Average Trading Volume Fee (Maker)	0.3644*** (0.0743)	-2.073*** (0.1143)	-0.679*** (0.0073)	-0.409*** (0.0093)	-0.975	-0.4324	-2.593	-1.5611
	Total Observation	**619973**	**293381**						

Note: This table presents results of 2SLS regressions of log volatility on the log volume, with different trading fee as instrument. Each row presents results using a different measure of trading fees. The crypto exch. stands for Crypto-only exchange. Economic significance is computed as the change in the dependent variable for a one-standard-deviation increase in the fee. Standard errors are presented in the parenthesis ** represents p-value < 10%, ** represents p-value < 5%, and *** represents P-value < 1%*. The sample period is from 2015 to 2020.

2. <u>Sampling based on the Date (Opted Date- 15th December 2017)</u>

a. 2SLS of Log-Returns on Log Volume with Trading Fee as an instrument.

Table 26 illustrates the IV estimation of log returns on log volume with trading fees. We observe a positive relationship between volume and returns in the sub-sample for most of the estimates except the Total Average volume fee and Maker's Average trading fee. The positive relationship is consistent with the empirical literature but not with our main results. The probable reason for this discretion is that period before the 15th December 2017 was smooth, i.e., there was no significant market breakdown in cryptocurrency, but cryptocurrency suffered a significant breakdown after that date and experienced a Great Crypto Crash in 2018. This table marks an important aspect as to why we found the negative relationship between the volume and returns. The inconsistent results are only due to the data's bifurcation into two different periods, i.e., without market shocks and with market shocks.

Table 26: 2SLS of Log Return on Log Volume of subsample

No.	Particulars	α		β (Volume)		R-Squared		Economic Significance	
		Before	After	Before	After	Before	After	Before	After
1	**Total Average Volume Fee**	-0.0206* (0.0106)	0.0310*** (0.0045)	0.0019* (0.0010)	-0.002*** (0.0045)	-0.0002	-0.005	0.0068	-0.010
2	**Average Trading Volume Fee (Taker)**	- 0.0267** (0.0122)	-0.0470** (0.0209)	0.0025** (0.0011)	0.0043** (0.0019)	-0.0007	-0.0072	0.0089	0.0156
3	**Average Trading Volume Fee (Maker)**	-0.0184* (0.0103)	0.0267*** (0.0033)	0.0017* 0.0010	-0.002*** (0.0003)	-9.1e-05	-0.0039	0.0061	-0.009
	Total Observation	**643871**	**251931**						

Note: This table presents results of 2SLS regressions of log return on the log volume, with different trading fee as instrument. Each row presents results using a different measure of trading fees. The crypto exch. stands for Crypto-only exchange. Economic significance is computed as the change in the dependent variable for a one-standard-deviation increase in the fee. Standard errors are presented in the parenthesis ** represents p-value < 10%, ** represents p-value < 5%, and *** represents P-value < 1%*. The sample period is from 2015 to 2020.

b. 2SLS of Log-Returns on Log Volume with Trading Fee as an instrument.

Table 26 presents the IV estimation of log volatility on log volume with trading fees. We observe a positive relationship between volume and volatility for most of the beta coefficients before and after the period. The results of robust checks are consistent with the literature, and our main results are presented in Table 18 as we expect the volatility to increase with the trading volume. However, we find a negative relationship between volume and volatility for Taker's Average Fee and Maker's Average Fee for before the date data, which could be because of the smooth financial period where the increase in volume also does not intensify the volatility. Expect the two scenarios mentioned above, we find the subsample result to be coherent with our main results.

Table 27: 2SLS of Log Volatility on Log Volume of subsample

No.	Particulars	α		β (Volume)		R-Squared		Economic Significance	
		Before	After	Before	After	Before	After	Before	After
1	**Total Average Volume Fee**	-4.226*** (0.1369)	5.6782*** (0.1487)	0.2085*** (0.0129)	1.1446*** (0.0136)	-0.128	-2.444	-0.748	-4.1614
2	**Average Trading Volume Fee (Taker)**	-4.033*** (0.1639)	-33.13*** (0.8804)	-0.226*** (0.0154)	2.4071*** (0.0807)	-0.145	-10.20	-0.814	8.7515
3	**Average Trading Volume Fee (Maker)**	-4.296*** (0.1305)	2.1943*** (0.0790)	-0.202*** (0.0123)	-0.8258 (0.0072)	-0.122	-1.290	-0.725	-3.0022
	Total Observation	**643871**	**251931**						

Note: This table presents results of 2SLS regressions of log volatility on the log volume, with different trading fee as instrument. Each row presents results using a different measure of trading fees. The crypto exch. stands for Crypto-only exchange. Economic significance is computed as the change in the dependent variable for a one-standard-deviation increase in the fee. Standard errors are presented in the parenthesis ** represents p-value < 10%, ** represents p-value < 5%, and *** represents P-value < 1%*. The sample period is from 2015 to 2020.