Non-fundamental Uncertainty in Crypto-assets on Public Blockchains

Redouane Elkamhi* Sedat Ersoy†

May 27, 2020

Abstract

We extend Kyle (1985) by introducing transparency in trade orders and insiders with single/multiple accounts. We find that splitting orders across time without fully revealing private information requires multiple accounts. Therefore, insiders with a single account trade more aggressively. This heterogeneity in trade behaviors leads to non-monotonicity between volume and fundamental value. The model delivers testable empirical predictions. For instance, crypto-asset prices with higher absolute total order flows are less informative in the short-term and more likely to have price reversals in the long-term. Those in exchanges with larger volatility of liquidity shocks are more informative in the short-term.

JEL classification: D82, G14.

Keywords: Blockchain, non-monotonicity, price discovery, strategic trading, transparency.

*Joseph L. Rotman School of Management, University of Toronto, 105 St. George Street, Toronto, Ontario, M5S 3E6. Redouane.Elkamhi@rotman.utoronto.ca.

†Department of Economics, University of Toronto, 150 St. George Street, Toronto M5S 3G7 Canada; Corresponding author: sedat.ersoy@mail.utoronto.ca.

1. Introduction

One of the main distinctions between the public blockchain system and the traditional financial system is the transparency that the public blockchain provides in that everyone can observe the holdings and transactions of each public address.[1] This difference raises the following question: Can transparency and privacy go hand in hand in the blockchain technology, especially for informed investors? This concern is of interest to the literature because the visibility of both current holdings as well as all past transactions might reveal their private information. Therefore, such transparency can (i) disincentivize informed investors to trade more often, and (ii) decrease information production. The latter is due to a better technology which induces traders to extract others' information from order flow rather than producing their own information, as discussed in Farboodi and Veldkamp (2019).[2] The analysis of the former effect is the main purpose of this study.

We show that the analysis of the incentive to trade by informed investors is important since information asymmetry is one of the main reasons why investors trade and, revealing such information through the blockchain technology can hamper their trading activities and directly their private information. As a result, this concern can affect both market liquidity and price discovery. While we focus on public blockchains, our results are still valid for any platforms (i) with similar transparency regarding past transactions and (ii) having extra cost to split the transactions across time and/or across different accounts.

It is worth highlighting that this privacy has been a main concern and a subject of discussion among the blockchain community. For example, Buterin (2016) discusses methods to keep traders' information private. They range from low to high-tech methods, as stated by Vitalik Buterin, co-founder of Ethereum:[3] *"The simplest strategies have to do with making maximum use of separate one-time accounts for each operation, so that a single individ-*

[1] We refer the readers to Catalini and Gans (2016) and, more recently, Biais, Bisiere, Bouvard, and Casamatta (2019) for a detailed discussion of public blockchains.

[2] Yang and Zhu (2020) is another example of such argument. In their model, the back-runner who is uninformed extracts information from order flow trades without trying to produce new information.

[3] Ethereum is one of the mostly used cryptocurrencies, and its platform is commonly used for start-ups.

ual's activities cannot be linked to each other". As suggested by Buterin, using multiple accounts across transactions can be the easiest method to hide information. This is because even though all the accounts' transactions are permanently stored in public blockchains and traceable by other investors, there is no perfectly accurate algorithm to match these accounts with the users. Therefore, traders can maintain their anonymity. Ron and Shamir (2013) show that although there are significant number of traders using multiple accounts, many of them use a single account.[4] Their empirical evidence demonstrates that investors are heterogeneous in the use of the number of accounts.

We introduce a two-period model, similar to Kyle (1985), where the insider and liquidity traders give market orders in both periods and the market maker clears the market by setting prices each period.[5] Additionally, our model has two main features, that are fundamentally different from traditional markets. First, the transparency makes traders' transaction history visible to the community, implying that individual trades are public to both the market maker and other investors. Such visibility of individual trades would not give extra information to the community in the short-term.[6] However, trades of informed investors would reveal their private information in the long-term unless they use multiple accounts. Second, total transaction fees paid by the trader increases with the frequency of trades. This second feature occurs because verifying a transaction of 1 token or 100 tokens requires a similar amount of computational work. Hence, informed traders who want to smooth their trades across time have to incur much higher costs. For example, consider two similar informed investors in terms of their private information in which one of them trades only in the first period whereas the other investor trades in both periods. The latter investor would pay twice as much money for her transactions as the former investor would pay. Both features constitute important issues for informed investors who prefer trading multiple periods to

[4]More specifically, they show that, for Bitcoin transactions, there are 2,214,186 owners using a single account, and there are 234,015 owners using at most 10 multiple accounts in their transactions. Despite the high variance in the number of multiple accounts, each trader has about 1.5 accounts, on average.

[5]We use the insider and an informed trader/investor interchangeably in this paper.

[6]We use the short-term (long-term) and the first period (second period) interchangeably in this paper.

exploit available information as long as possible because they incur higher transaction fees, compared to trading once. Therefore, smoothing trades across periods in public blockchains comes with the expense of paying higher transaction fees.

In line with the work of Ron and Shamir (2013), which shows that traders vary in the number of use of accounts, we assume that there are two types of potential insiders. We refer to the insider with a single account (multiple accounts) as *s-type* (*m-type*). We also assume that the conditional probability of being s-type insider given the first period total order flow is constant, μ. Put differently, the likelihood of being s-type is independent of the first period total order flow, implying that the magnitude of total order flow would not provide extra information about the type of the insider. We consider this conditional probability, μ as a proxy for the probability of being s-type insider.

Our paper contributes to the literature on Kyle (1985) type models in a number of ways. First, Kyle-type models imply that there is one to one mapping between insider's private information and her trade. Put differently, knowing insider's trade perfectly reveals the underlying private information. However, we show that, in equilibrium, even if insider's trade is known by other market participants in the short-term, her trade does not reveal her private information since her type is not known at this stage. Second, in Kyle-type models, the existence of liquidity traders in the market enables insider to benefit from obscurity since the market maker observes only the sum of trading by insiders and liquidity trading. As a result, the market maker cannot discover the trading by insiders. In addition to this uncertainty, the existence of two types of informed traders with different trading behavior leads the market maker to consider both scenarios in a probabilistic sense, which creates the second layer of uncertainty. We call this layer of uncertainty non-fundamental uncertainty. The reason is that the type of insiders does not affect the fundamental value of crypto-assets, but the valuation of crypto-assets. Third, s-type insiders trade very aggressively and only in the first period.[7]

[7]Kyle-type models do not generate such one-time trading with high aggressiveness. See, for example, Collin-Dufresne and Fos (2016), Yang and Zhu (2020), Brunnermeier (2005), Holden and Subrahmanyam

We first prove that there is a threshold level of cost of using multiple accounts. Above the threshold level, m-type traders do not trade in both periods because benefits of smoothing trades cannot compensate the cost and, hence, act similarly to s-type traders. Below the threshold level, m-type traders can bear the cost and behave differently than s-type traders. We also show that the threshold level increases in both the volatility of the liquidity shock and volatility of the asset, which results in an important observation: Consider similar types of crypto-assets in a market with different volatility of the liquidity shock. M-type traders can (not) smooth their trades in the case where that volatility is significantly higher (lower), implying that similar types of assets would have different trading patterns and, hence, market qualities.

We then characterize the model for the general case with $\mu \in (0, 1)$ and prove that a linear equilibrium always exists and is unique. More specifically, we show that when the cost of using multiple accounts is sufficiently high, both types of investors trade only in the first period. In such specific case, the equilibrium parameters have a closed-form solution. More interestingly, when the cost is sufficiently low, we show that unique equilibrium model parameters are only in the forms of (i) proxy for the likelihood of being the s-type investor, (ii) the volatility of the asset and (iii) the volatility of the liquidity shock.

In the case of sufficiently low cost of using multiple accounts, we analyze the behavior of equilibrium parameters with respect to the proxy for the probability of being the s-type insider. As that proxy increases, the market maker puts more weight to the case that the insider is of s-type. Since s-type insider trades more aggressively, short-term price impact increases, yielding lower trade aggressiveness in the short-term for both investors. In the long-term, we focus on the model parameters for which the investor is of m-type. In this case, trade aggressiveness of m-type investor decreases in that proxy. The reason is as follows: As that proxy increases, m-type considers to trade more in the long-term rather than short-term since short-term price impact becomes higher. However, the market maker takes her

(1992), Admati and Pfleiderer (1988), among others.

strategy into consideration and, hence, set long-term price impact higher when the type of the investor is of m-type. Such a higher price impact in the long-term decreases her trade aggressiveness.

We also explore the effects of non-fundamental uncertainty on the short-term price discovery. We show that the price discovery is a function of both (i) absolute net order flow and (ii) the volatility of the liquidity shock if and only if the second layer of uncertainty is present in the market, i.e., $\mu \in (0, 1)$. Furthermore, our model shows that informativeness of the prices decreases with the absolute short-term net order flow. As that order flow increases, the market maker considers possible values for the fundamental value which are farther apart from each other, leading to lower price prediction power. As a result, our model predicts that large price reversals (price momentum) in crypto-assets follow the high absolute short-term net order flow, and, on average, reversals (momentum) gets larger as that order flow increases. Moreover, our model explains that as the volatility of the liquidity shock increases, price discovery in the short-term increases. This result is, at first, counter-intuitive because more randomness in the uninformed trading improves the price informativeness in the short-term. However, the insight is that the larger the volatility of that decreases the price impact, yielding to increase the incentive of more aggressive trading. Such trading behavior brings benefits to the short-term price discovery.

We then analyze the long-term price discovery when the cost of using multiple accounts are sufficiently low. We show that compared to the market with the s-type investor, the long-term price discovery is higher in the market with the m-type investor. This comparison implies that trading both periods reveals more information than that trading very aggressively in the short-term. We also demonstrate that whether the investor is of s-type or m-type, a higher proxy for the likelihood of being the s-type insider decreases both the short-term and the long-term price discovery. For the market with s-type investor, price discovery in the short-term is the same as that in the long-term, and decreases in that proxy. The intuition for the long-term is that there are two channels affecting the price discovery,

which are short-term and long-term trade aggressiveness of m-type investor. We already show that her aggressiveness in both periods decreases due to a higher proxy level, yielding lower price discovery in the long-term.

We finally argue that due to the difference in aggressive trading across both types of insiders, even relatively low volume in the short-term could be a sign of high absolute fundamental value. Similarly, relatively high volume in the short-term could be associated with a low absolute fundamental value. Hence, our model concludes that volume versus fundamental value relation is non-monotonic, suggesting that volume cannot have predictive power for future returns of crypto-assets. Similar to our result, Balcilar, Bouri, Gupta, and Roubaud (2017) show that the relation between the return and volume in Bitcoin has both nonlinearity and structural breaks. Our finding, thus, provides a theoretical explanation of their results.

The remainder of this section provides a review of the related literature and the relevant features of the public blockchain technology. Section 2 provides the description of the model. Section 3 presents both the characterization of the equilibrium and the main results. Section 4 analyzes the price discoveries. In Section 5, we present model predictions on crypto-asset market. Section 6 concludes. Additional discussion and all proofs are in the appendix.

1.1. Related Literature

This paper is related to the fast-growing literature on blockchain. In this literature, benefits of blockchain in corporate governance (Yermack, 2017), benefits of transparency on supply chain (Chod, Trichakis, Tsoukalas, Aspegren, and Weber, 2020), and welfare destroying consequences (Cong and He, 2019) are discussed in great detail. Easley, O'Hara, and Basu (2019) show that the blockchain cannot fully eliminate transaction fee. When it does, the blockchain is not viable in the long-term. More related studies to our paper consider the consequences of FinTech on market quality. More specifically, Farboodi and Veldkamp (2019) argue that improvement in the productivity of information processing decreases the

production of information, yet increases price informativeness and decreases market liquidity. Zhu (2018) shows that introduction of big data including consumer transactions increases price efficiency due to lower information acquisition cost.

We are not the only one considering the benefits of using multiple accounts in financial markets. Malinova and Park (2017) investigate the possible market designs for the blockchain technology and its effects on investor welfare, investor trading behavior, and trading cost. They show that the most transparent setting results in the highest investor welfare. Moreover, in the absence of full transparency, multiple-ID yields weakly higher welfare since investors who are hit by liquidity shock are more likely to trade. Their study is based on different degree of transparency settings and single or multiple IDs used in their trades. In contrast to their work where trading activities triggered by liquidity shocks, the main reason to trade in our paper is information asymmetry. We also have novel implications on (i) both short-term and long-term price discovery, and (ii) the prediction power of volume on crypto-asset returns, which are completely new to this paper.

High price volatility is a main concern for the crypto-currencies in which many studies attempt to uncover the underlying reasons. Zimmerman (2020) rationalizes high price volatility by showing that speculative trading crowds out monetary usage of the currency and hence traders with investment purpose face higher riskiness. The reason could also be due to passive monetary policy in Proof-of-Work environment (Saleh, 2019), pump and dump schemes (Li, Shin, and Wang, 2019), suspicious trading (Gandal, Hamrick, Moore, and Oberman, 2018). Aside from these works, the underlying reason could be unrelated to fundamentals (Biais, Bisiere, Bouvard, Casamatta, and Menkveld, 2018). Our paper contributes to this literature by showing that the reason could be because of the presence of different types of traders. Having both informed traders with a single account and multiple accounts leads to additional uncertainty to set the short-term prices. Such non-fundamental uncertainty yields an absolute total order flow terms and the volatility of the liquidity shock to show up in the short-term price discovery. As a result, price informativeness can be quite

inefficient in the time of high absolute total order flow or low volatile of the liquidity shock, yielding prices highly volatile in the short-term.

Our paper is also closely related to Huddart, Hughes, and Levine (2001). The authors argue that dissimulation of first period trade, i.e., adding a random part to insider's first period strategy, known by only herself, makes it possible to exploit her initial private information in the second period. This is important to note because of two reasons. First, the public disclosure in their setup is a similar mechanism as transparency in our settings. Second, their dissimulation technique plays a similar role as using multiple accounts across the periods in our setting. However, such dissimulation in public blockchains is not sufficient to hide their private information. Even if the insider uses this method rather than using multiple accounts, her second period trade reveals her private information perfectly. The reason is that her private information would be a linear function of first period total order flow and her second period trades, which are observable to the market maker at the time of setting the second period price.[8] As a result, dissimulation of trades by adding noise to the first-period trade is not a substitute for using multiple accounts in the blockchain environment.[9]

1.2. Relevant Features of the Public Blockchain Technology

In this section, we discuss the features on the Blockchain system, making its environment substantially different than the traditional financial markets. This discussion will guide our analysis on the ability of both s-type and m-type insider to hide their information while trading.

[8]Formally, $v = \frac{x_2 - \beta_y y_1}{\beta_2}$, where β_y and β_2 are common knowledge and the trade amount in the second period, x_2, and net total order in the first period are observable to the market maker as well.

[9]It is important to note that in a mixed strategy equilibrium, the insider cannot dissimulate her second period strategy, $x_2 = \beta_2 v + \beta_y y_1 + \varepsilon_2$, otherwise she would not be indifferent for each realization of ε_2.

1.2.1. Costs of Using Multiple Accounts

The nature of accounts and process of transactions in this setting are significantly different. In this environment, every account owner has one public key and one private key. The public key is publicly shared and is required if someone else sends money to the owner's account. The private key provides privacy for your account and is required if the owner sends money/coin/token to someone else. Once both the public and the private keys are known by someone else, then they have control over this account. Therefore, keeping the private key safe is a crucial step to keep your saving safe.[10]

Moreover, opening an extra account on an exchange is time consuming process. Having more than one accounts can also increase the possibility of forgetting private key because these keys are very complicated to memorize. Therefore, they have been a concern for many users in the blockchain community.[11] More specifically, each account has a private key to complete transactions and if one forgets this key, there is no way to have this account back.

Lastly but most importantly, transaction fees generally do not depend on the amount of token/unit transferred, but more on the frequency of transactions. Even if an extremely cautious insider can minimize the cost associated with the above cases, (e.g. time consuming and the possibility to forget), she has to bear the higher cost if she trades more frequently. For example, for the case of Ethereum transactions, Figure 1 shows that transaction fees do not linearly increase in the amount of token transferred, rather it declines. As a result, if the insider splits her orders across time or across multiple accounts in the same or different periods, she bears much higher transaction fees.

In line with the above discussion, we consider the cost of having more than one account and the cost of using multiple accounts as the total cost which is incurred by the insider with multiple accounts for each additional trade, and, denote it as c. Although there is

[10] https://www.mcclatchydc.com/news/nation-world/national/national-security/article210865704.html

[11] See examples: https://www.wsj.com/articles/good-news-you-are-a-bitcoin-millionaire-bad-news-you-forgot-your-password-1513701480, https://www.wired.com/story/i-forgot-my-pin-an-epic-tale-of-losing-dollar30000-in-bitcoin/. Recently, QuadrigaCX, one of the largest crypto exchange in Canada, has lost $190 Million: https://www.coindesk.com/quadriga-creditor-protection-filing.

substantial heterogeneity on c across crypto-assets on different exchanges, we do not impose any restrictions. As a result, the cost c can take even zero value for a specific crypto-asset on a specific exchange.

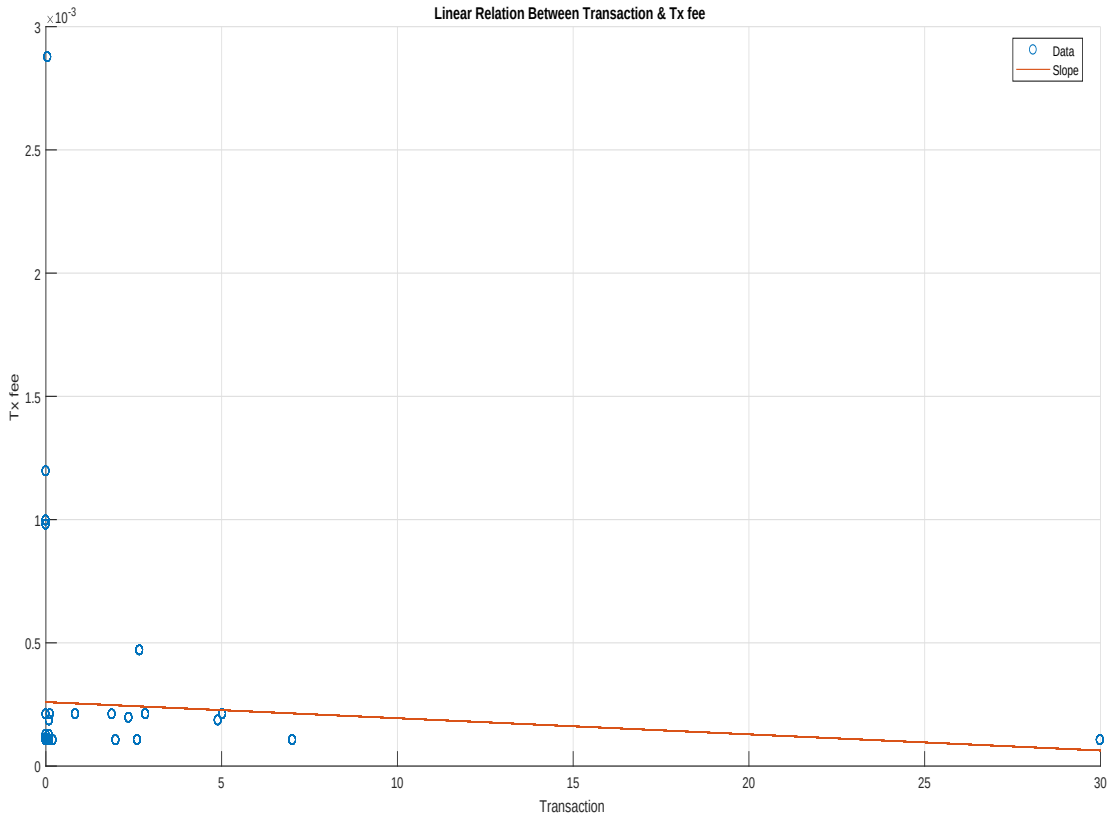

Fig. 1. This figure plots the relation between Ether transactions and corresponding transaction fees in Ether. 1257 transactions with positive amount, made on 23 February, 2019 are considered for this plotting. The source of the data is https://etherscan.io.

1.2.2. Transparency

This technology by its nature provides the feature that every single transaction is recorded on the block and becomes public. Such public information does not only include the current transaction of the relevant parties in the form of a public key but also provides the transaction

histories of the parties. Such histories would not reveal information to others if those parties behave non-strategically. However, when the investor has private information, she would like to act strategically by using this information in multiple periods. This strategy may not be materialized due to the possibility that her information would be revealed by others via public transactions before she trades in multiple periods.

2. The Model

2.1. Setup

In this paper, we model the market with an insider, market maker and liquidity traders. Their past transactions and holdings are visible to other market participants via public blockchain technology. We focus on pure linear strategies.

Asset. There is a single risky crypto-asset (asset, henceforth) which pays a liquidation value v at the end of the second period, where $\tilde{v} \sim N(0, \sigma_v^2)$ with $\sigma_v > 0$.[12] In our setting, the liquidation value of the crypto-asset can be transactional benefits on the platform (Sockin and Xiong 2018; Biais et al. 2018), likelihood of its adoption as a global currency (Makarov and Schoar 2019), benefits due to computing power (Bhambhwani, Delikouras, and Korniotis 2019; Pagnotta and Buraschi 2018).

Market Participants. There are three types of market participants: (i) $N \geq 1$ liquidity traders(i), a risk-neutral insider, and (iii) a risk-neutral market maker (MM, henceforth).

Liquidity Traders. We assume that there is infinitely many uninformed traders who might be hit by liquidity shock in each period due to exogenous reasons.[13] The ones who do not get the shock remain outside the market. The number of the liquidity traders in each period, N, is common knowledge. Moreover, since there is a continuum of them, it is quite rare for

[12]It is common to refer to crypto as crypto-asset. For instance, Mark Carney, Governor of the Bank of England, referred to crypto as crypto-assets throughout his speech on 2 March, 2018 about the future of money.

[13]O'Hara (1995) and Brunnermeier (2001) give a detailed discussion of possible reasons for liquidity traders to trade.

these traders to be hit by liquidity shock in both periods. Therefore, we assume that liquidity traders in period 1 are different than those in period 2.[14] In our setup, the liquidity traders, both in the first and the second period, place market orders, in total, u_1 and u_2, which are independently and normally distributed with mean 0 and variance σ_u^2, with $\sigma_u > 0$.

Insider (her). The insider might have either a single account or multiple accounts, denoted by s-type and m-type, respectively. She observes the value of the project v before she trades in the first period.[15] Her information set at $t = 1$ would be $I_1^i = \{v\}$. After trading has finished in the first period, individual trades become public along with the trading price, p_1. Additionally, observable individual trades would reveal liquidity trading to the insider since it is equal to trading difference between total order, y_1 and her first period trade, x_1. Thus, her information set at $t = 2$ would be $I_2^i = \{x_1, u_1, p_1, v\}$. The insider's objective function is to maximize her expected profit given her available information, shown explicitly in subsection 2.3.

Market Maker. There is a competitive risk-neutral market maker who observes market orders and set the market prices each period. Due to the competitive market feature, MM sets $p_t = E(\tilde{v} \mid I_t^{mm})$, where I_t^{mm} denotes all available information that MM can process at time t. Given the total order flow at $t = 1$, y_1, the conditional probability of being s-type (m-type) insider equals μ $(1 - \mu)$. Assuming that trading strategies at $t = 1$ are linear for both s-type and m-type insiders, $\beta_{1,s}v$ and $\beta_{1,m}v$, respectively, where $\beta_{1,s}(\beta_{1,m})$ is the trade aggressiveness of s-type (m-type) insider at $t = 1$, the distribution of total order flow at $t = 1$ when the insider is s-type is given by

$$\tilde{y}_{1,s} \sim N(0, \beta_{1,s}^2\sigma_v^2 + \sigma_u^2).$$

[14] We get a similar result by assuming that informed investors trade more often than liquidity traders, which is consistent with the literature on institutional traders. For example, Van Kervel and Menkveld (2019) show that institutional orders consist of many child orders following the parent order.

[15] Here, the insider is the one who might have private information about the project, e.g. corporate insiders, foundation members or some members of the token ecosystem.

Similarly, the distribution of total order flow at $t = 1$ when the insider is m-type is given by

$$\tilde{y}_{1,m} \sim N(0, \beta_{1,m}^2 \sigma_v^2 + \sigma_u^2).$$

Then μ can be denoted as

$$P(\tilde{y}_{1,s} = y_1 \mid y_1) = \mu.$$

As the population of s-type insider increases, the conditional probability of being s-type insider given the total order flow would increase as well, hence μ can be seen as the proxy for the population of s-type insiders.

Timing

$t = 1$

(i) The insider privately observes the true value of the project, and decides to trade in both periods or only in the first period.[16]

(ii) The liquidity traders and the insider place their market orders, i.e., u_1 and x_1, respectively, without knowing the market price, p_1.

(iii) MM observes market orders without knowing their identities, and sets the market price in the first period, p_1.

$t = 2$

(i) The liquidity traders and the insider observe the market price, p_1, and place their market orders, i.e., u_2 and x_2, respectively, without knowing the market price, p_2.

(ii) MM observes (a) market orders without knowing their identities, and (b) accounts used in both periods, and then sets the market price in the second period, p_2.

(iii) Asset payoff $\tilde{v}$ is realized, and traders get paid.

The main differences in our setting compared to Kyle-type models, (i) MM can see individual market orders, and (ii) the holdings and past transactions of the traders are observable as well. These two features have consequences on the equilibrium strategies.

[16]The reason why she might want to trade only once is due to having a single account in the platform.

Strategies. We conjecture the following linear strategies for s-type and m-type insider, respectively:

$$x_{1,s} = \beta_{1,s} v, \quad x_{2,s} = 0,$$
$$x_{1,m} = \beta_{1,m} v, \quad x_{2,m} = \beta_2 v + \beta_y y_1$$

where subscript s and m stands for s-type insider and m-type insider, respectively. β's are trade aggressiveness, and y_1 is the total order flow at $t = 1$. We explain the reason why second period trade of s-type is 0 below in subsection 2.2.

2.2. Trading Behaviour of Insiders

s-type. If the insider has a single account and would like to trade in both periods, her private information would be perfectly revealed by MM if the market maker keeps track of account activities, that is, whether the trader in the second period trades in the first period as well. This is because liquidity traders in both periods are different and thus, any repeated trader across periods has to be the insider.[17] Therefore, MM can identify the s-type along with her private information, v. In such a case, MM would set the price as $p_2 = v$, and s-type cannot benefit from her information, even she would lose for a certainty due to the transaction cost. This analysis leads to the following result.

Lemma 2.1. *The insider with a single account only trades once.*

m-type. If the insider has multiple accounts and trades in both periods, the method discussed above would not work to identify the insider since m-type uses different accounts across periods.[18]

[17] The complexity of such tracking is linear in number of total traders, $N + 1$. Hence, if the computerized system uses this method, time complexity would be $O(N)$. Indeed, the average number of daily transactions in the year of 2019 for Bitcoin is above 500,000, and for Ether is above 600,000. Despite the huge number of daily transactions in these platforms, keeping tracking of repeated traders across periods would not be computationally hard problem.

[18] This relies on an implicit assumption so that m-type insider is not identified before p_2 is set. The details

Having multiple accounts could be one of the easiest ways among other technical methods described in the report of Buterin (2016) to keep information hidden for a longer period. The reason to focus on having multiple accounts among other technical ways is the simplicity of modeling such a decision. Therefore, having multiple accounts represents the insider's effort to hide her private information whereas associated costs represent the cost of insider's investment on technological knowledge / human capital. Thus, multiple accounts in our model represent the more secure way to hide identity at the expense of higher cost.

2.3. Equilibrium Derivation

Equilibrium Definition: A linear equilibrium of the setting is given by a strategy profile

$$\left\{x^*_{1,s}(v), x^*_{1,m}(v), x^*_2\left(v, y_1\right), p^*_1(y_1), p^*_2(y_1, y_2)\right\}$$

such that

$$x^*_2 \in \operatorname{argmax} E\left[x_2(\tilde{v} - p_2) - c\mathbb{1}_{\{x_2 \neq 0\}} | \tilde{v} = v, y_1\right], \tag{1}$$

$$x^*_{1,m} \in \operatorname{argmax} E\left[x_1(\tilde{v} - p_1) + x^*_2(\tilde{v} - p_2) - c | \tilde{v} = v\right], \tag{2}$$

$$x^*_{1,s} \in \operatorname{argmax} E\left[x_1(\tilde{v} - p_1) | \tilde{v} = v\right], \tag{3}$$

$$p_1 = E(\tilde{v}|I^{mm}_1),$$

$$p_2 = E(\tilde{v}|I^{mm}_2),$$

where $\mathbb{1}_{\{x_2 \neq 0\}}$ is an indicator function of trading whose value is 1 if there is a non-zero trade by the insider in the second period and 0 otherwise. c is the cost of using multiple accounts, which is incurred at $t = 2$.

Market maker's decision. MM observes individual trades at each time and sets semi-strong efficient prices. MM knows that there would be exactly N liquidity traders in each

are in Appendix A. It is, however, important to note that with a slightly varied information structure, the model would not need such an implicit assumption and preserves main results we have in this paper.

period and, hence, the number of trades in the second period reveals the type of the insider in the first period. If there are $N+1$ traders at $t=2$, MM concludes that the insider is of m-type. Otherwise, MM knows that there would not be an insider at $t=2$. Although the type of the insider is revealed at $t=2$, MM has no possibility to figure out her type before setting first period price, p_1. However, MM has a belief μ about her being s-type insider conditional on total order flow.

Since identities within the period and across the periods are not identified, total order flows as well as the number of trades in the second period are sufficient statistics to set the prices.

Price setting at $t=1$.

When MM observes the total order flow at $t=1$, y_1, MM knows that total order flow is in the form of $y_1 = \beta_{1,s}\tilde{v} + \tilde{u}_1$ ($y_1 = \beta_{1,m}\tilde{v} + \tilde{u}_1$) with probability μ $(1-\mu)$. Then, the pricing decision at $t=1$ would be a linear function of total order flow, y_1.

$$\begin{aligned} p_1 &= E(\tilde{v} \mid y_1) \\ &= P(y_1 = \tilde{y}_{1,s} \mid y_1)E(\tilde{v} \mid y_1 = \tilde{y}_{1,s}) + P(\tilde{y}_1 = \tilde{y}_{1,m} \mid y_1)E(\tilde{v} \mid y_1 = \tilde{y}_{1,m}) \\ &= \mu\frac{Cov(\tilde{v}, \tilde{y}_{1,s})}{Var(\tilde{y}_{1,s})}y_1 + (1-\mu)\frac{Cov(\tilde{v}, \tilde{y}_{1,m})}{Var(\tilde{y}_{1,m})}y_1 = \mu\frac{\beta_{1,s}}{\beta_{1,s}^2 + \frac{\sigma_u^2}{\sigma_v^2}}y_1 + (1-\mu)\frac{\beta_{1,m}}{\beta_{1,m}^2 + \frac{\sigma_u^2}{\sigma_v^2}}y_1, \end{aligned}$$

where the last line exploits the projection theorem. The last equality implies that the price impact in the first period, λ_1, can be written as follows:

$$\lambda_1 \equiv \mu\frac{\beta_{1,s}}{\beta_{1,s}^2 + \frac{\sigma_u^2}{\sigma_v^2}} + (1-\mu)\frac{\beta_{1,m}}{\beta_{1,m}^2 + \frac{\sigma_u^2}{\sigma_v^2}}. \tag{4}$$

Price setting at $t=2$.

Once MM observes the number of trades at $t=2$, ambiguity about the type of the insider would disappear. Therefore, we need to consider two cases where the number of trades at $t=2$ is $N+1$ and N. Let's denote the number of trades at $t=2$ as $\#_2$.

Case 1. $\#_2 = N+1$. Since the insider is of m-type, total orders in the first and the second

period are in the form of $y_1 = \beta_{1,m}\tilde{v} + \tilde{u}_1$, and $y_2 = (\beta_2 + \beta_{1,m}\beta_y)\tilde{v} + \beta_y\tilde{u}_1 + \tilde{u}_2$, respectively. Hence, by using the projection theorem we have

$$p_2 = E(\tilde{v} \mid y_1, y_2) = \frac{\beta_{1,m} - \beta_2\beta_y}{\beta_{1,m}^2 + \beta_2^2 + \frac{\sigma_u^2}{\sigma_v^2}} y_1 + \frac{\beta_2}{\beta_{1,m}^2 + \beta_2^2 + \frac{\sigma_u^2}{\sigma_v^2}} y_2.$$

Hence, we have

$$\gamma_1 \equiv \frac{\beta_{1,m} - \beta_2\beta_y}{\beta_{1,m}^2 + \beta_2^2 + \frac{\sigma_u^2}{\sigma_v^2}}, \tag{5}$$

$$\lambda_2 \equiv \frac{\beta_2}{\beta_{1,m}^2 + \beta_2^2 + \frac{\sigma_u^2}{\sigma_v^2}}. \tag{6}$$

Case2. $\#_2 = N$. In this case, the insider is of s-type and does not trade in the second period. The total order flow at $t = 2$ is then completely noisy and, therefore, the price would not depend on y_2.

$$p_2 = E(\tilde{v} \mid \tilde{y}_{1,s} = y_1) = \frac{\beta_{1,s}}{\beta_{1,s}^2 + \frac{\sigma_u^2}{\sigma_v^2}} y_1.$$

Hence, we have

$$\Gamma_1 \equiv \frac{\beta_{1,s}}{\beta_{1,s}^2 + \frac{\sigma_u^2}{\sigma_v^2}}. \tag{7}$$

The pricing decision at $t = 2$ can be written as follows:

$$p_2 = \begin{cases} \gamma_1 y_1 + \lambda_2 y_2, & \text{if} \quad \#_2 = N + 1 \\ \Gamma_1 y_1, & \text{if} \quad \#_2 = N \end{cases}$$

The decision of m-type insider in period 2. The m-type insider first decides on being present in the market or not, and then transaction amount. Our conjecture, which is proven in Proposition 1, is that the m-type insider trades in the second period if the cost of using multiple accounts is sufficiently low. If the cost is sufficiently high, she prefers to not trade and her optimal trades in period 1 and period 2 would be the same as optimal trades of the s-type insider. As a result, for the remainder of this section, we assume that the cost is sufficiently low. Her second period problem, then, leads to the following first order condition

(FOC):

$$x_2^* = \frac{v}{2\lambda_2} - \frac{\gamma_1}{2\lambda_2} y_1, \tag{8}$$

and the second order condition (SOC) is $\lambda_2 > 0$. FOC implies the following identifications:

$$\beta_2 = \frac{1}{2\lambda_2}, \tag{9}$$

$$\beta_y = -\frac{\gamma_1}{2\lambda_2}. \tag{10}$$

Then, the expected profit of m-type insider in the second period, given private information levels and total order flow, y_1, can be expressed as

$$E[\Pi_{2,m} \mid \tilde{v} = v, y_1] = \frac{(v - \gamma_1 y_1)^2}{4\lambda_2} - c.$$

The decision of m-type insider in period 1. Her first period problem is expressed by (2) or equivalently by the following:

$$x_{1,m}^* \in \operatorname{argmax} E\left[x_{1,m}(\tilde{v} - p_1) + \frac{(v - \gamma_1 y_1)^2}{4\lambda_2} - c | \tilde{v} = v \right]. \tag{11}$$

Equivalently, it can be expressed as follows:

$$x_{1,m}^* \in \operatorname{argmax} x_{1,m} v - \lambda_1 x_{1,m}^2 + \frac{\gamma_1^2 x_1^2 - 2v\gamma_1 x_1}{4\lambda_2} + \frac{v^2 + \gamma_1 \sigma_u^2}{4\lambda_2} - c, \tag{12}$$

which gives us the following FOC:

$$x_{1,m}^* = \frac{2\lambda_2 - \gamma_1}{4\lambda_1\lambda_2 - \gamma_1^2} v, \tag{13}$$

and SOC is $4\lambda_1\lambda_2 - \gamma_1^2 > 0$. FOC implies the following identification:

$$\beta_{1,m} = \frac{2\lambda_2 - \gamma_1}{4\lambda_1\lambda_2 - \gamma_1^2}. \tag{14}$$

The expected profit of m-type insider after observing a private signal is given by equation (15) below. Her profit has four components. The first component is due to the information advantage of v in period 1. However, her information advantage generates a second component in the second period as well because this information is not fully revealed to prices at the end of the first period. The third component is generated in the second period due to liquidity trading in the first period. Note that liquidity trading in the first period provides profits to the insider in the second period, but not in the first period. The reason for this effect is that she knows the total order flow at $t = 1$, and hence liquidity trading at $t = 1$. Having observed that, at the end of first period, she knows the degree of noise trading in p_1 and, hence, takes the action in the second period accordingly, bringing her an extra profit in the second period. The last component is the cost of using multiple accounts.

$$\begin{aligned} E[\Pi_m \mid \tilde{v} = v] &= \left(\beta_{1,m} - \beta_{1,m}^2\lambda_1 + \frac{\beta_{1,m}^2\gamma_1^2 - 2\gamma_1\beta_{1,m} + 1}{4\lambda_2}\right) v^2 + \frac{\gamma_1^2\sigma_u^2}{4\lambda_2} - c \\ &= \underbrace{\underbrace{\left(\frac{\lambda_2 + \lambda_1 - \gamma_1}{4\lambda_1\lambda_2 - \gamma_1^2} - \frac{1}{4\lambda_2}\right) v^2}_{\text{due to } v \text{ at } t=1}}_{\text{generated at } t=1} + \underbrace{\underbrace{\underbrace{\frac{v^2}{4\lambda_2}}_{\text{due to } v \text{ at } t=2} + \underbrace{\frac{\gamma_1^2\sigma_u^2}{4\lambda_2}}_{\text{due to liquidity trading at } t=1} \underbrace{-c}_{\text{extra cost}}}_{\text{generated at } t=2}} \end{aligned} \tag{15}$$

Additionally, by taking the expectation of equation (15) over v, the ex-ante expected profit of m-type insider can be written as

$$E[\Pi_m] = \left(\frac{\lambda_2 + \lambda_1 - \gamma_1}{4\lambda_1\lambda_2 - \gamma_1^2}\right)\sigma_v^2 + \frac{\gamma_1^2\sigma_u^2}{4\lambda_2} - c. \tag{16}$$

The decision of s-type insider in period 1. Her first period problem, expressed in (3), results in SOC, i.e., $\lambda_1 > 0$ as well as $x_{1,s} = \frac{v}{2\lambda_1}$, which yields her trade aggressiveness:

$$\beta_{1,s} = \frac{1}{2\lambda_1}. \tag{17}$$

Similarly, we can express her interim and ex-ante expected total profit, respectively, as follows:

$$E[\Pi_s \mid \tilde{v} = v] = \frac{v^2}{4\lambda_1}, \tag{18}$$

$$E[\Pi_s] = \frac{\sigma_v^2}{4\lambda_1}. \tag{19}$$

Characterization of the equilibrium is determined by the equations (4, 5, 6, 7, 9, 10, 14, 17) along with the following SOC's:

$$\lambda_1 > 0, \quad \lambda_2 > 0, \quad 4\lambda_1\lambda_2 - \gamma_1^2 > 0.$$

Since this system of equations is non-linear there might be more than one solution or even it might not exist.[19] To prove the existence and uniqueness of this system, we first reduce it to one polynomial equation with only one unknown. We show in the next section that for any value of $\mu \in [0, 1]$, the equilibrium always exists and is unique. Before starting the reduction of the system, let's denote $g \equiv \frac{\lambda_2}{\gamma_1}$, and $h \equiv \frac{\lambda_1}{\gamma_1}$, where both are only functions of μ. In the next section, we show that the coefficients of a polynomial equation, A_i's, are only functions of μ, and hence it proves our argument that g and h depend only on the proxy for the likelihood of being the s-type insider, μ.

3. Main Results

Our main interest is to analyze how mixed population including insiders who use a single account and multiple accounts affect asset price decision, market liquidity, and price discoveries. In such a mixed population, we show that the cost of using multiple accounts has a critical impact on m-type insiders. This is because in the case of high cost, m-type insiders would not prefer to trade in both periods although they have multiple accounts and

[19] As we discussed in the next section, there are 11 roots who might be potentially the solution. We eliminate 10 of these roots and prove the remaining one always exists.

would not smooth her trades. Thus, m-type insiders behave as if they are s-type insiders. The following proposition formalizes this argument.

Proposition 1. *If the cost of using multiple accounts is sufficiently high, i.e.*

$$c > \bar{c} = \frac{(64g^6 - 16g^4 + 2g + 1)\sqrt{2g-1}}{16g^3(8g^3 - 4g^2 + 1)\sqrt{2g+1}}\sigma_u\sigma_v,$$

where g solves the 11th order polynomial $f(x) = \sum_{i=0}^{11} A_i x^i$, where the coefficients of A_i's are given in Appendix B.2.1, both s-type and m-type insiders trade only in the first period. Otherwise, only m-type insiders trade in both periods.

Proof. See Appendix B.1. □

Proposition 1 implies that the cost of using multiple accounts, c, switches the trading behavior of m-type insiders, leading to different equilibrium and hence different market qualities. We prove that the equilibrium exists and is unique for any value of $\mu \in [0,1]$ in both cases whether the cost is high or low. The following propositions state this argument.

Proposition 2. *(Existence and Uniqueness) Suppose that $\mu \in (0,1)$. If the cost of using multiple accounts is sufficiently low, i.e. $c \leq \bar{c}$ then a linear pure strategy equilibrium is characterized by*

$$\beta_{1,m} = \sqrt{\frac{2g-1}{2g(2h-1)}}\frac{\sigma_u}{\sigma_v}, \qquad \gamma_1 = \frac{\sqrt{2g(2h-1)(2g-1)}}{4gh-1}\frac{\sigma_v}{\sigma_u},$$

$$\lambda_1 = h\gamma_1, \qquad \lambda_2 = g\gamma_1, \qquad \Gamma_1 = \frac{\beta_{1,s}}{\beta_{1,s}^2 + \frac{\sigma_u^2}{\sigma_v^2}},$$

$$\beta_{1,s} = \frac{1}{2\lambda_1}, \qquad \beta_2 = \frac{1}{2\lambda_2}, \qquad \beta_y = -\frac{1}{2g},$$

along with the following SOC's:

$$h\gamma_1 > 0, \quad g\gamma_1 > 0, \quad 4gh - 1 > 0,$$

where $h = 2g^2 - g + \frac{1}{4g}$, *and* g *solves the 11th order polynomial:*

$$f(x) = \sum_{i=0}^{11} A_i x^i$$

where the coefficients of A_i*'s are given in Appendix B.2.1.*
If the cost of using multiple accounts is sufficiently high, i.e., $c > \bar{c}$*, then linear pure strategy equilibrium parameters take the following closed form:*

$$\beta_{1,s} = \beta_{1,m} = \frac{\sigma_u}{\sigma_v}, \qquad \lambda_1 = \frac{\sigma_v}{2\sigma_u},$$

Proof. See Appendix B.2. □

It is crucial to uncover the possible benefits and costs of having a low cost of using multiple accounts, hence we compare the equilibrium parameters in the case of low cost with those of the case of high cost. Let superscript H (L) denotes the equilibrium parameters in the case of high (low) cost. The next proposition states this comparison.

Proposition 3. *(Comparison of the equilibrium parameters)*

$$\beta_{1,s}^H < \beta_{1,s}^L, \qquad \beta_{1,m}^H > \beta_{1,m}^L,$$

$$\lambda_1^H > \lambda_1^L, \qquad 0 = \lambda_2^H < \lambda_2^L.$$

Proof. See Appendix B.3. □

The intuition of the Proposition 3 is as follows: When the cost is high, m-type insider would prefer to not trade in the period 2, hence she would increase trade aggressiveness in the first period. Since MM knows that both types would trade very aggressively, the price impact increases, which yields less aggressive trading for s-type insider. Lastly, since there would be no informed trader in the second period, the price impact would decrease to 0.

It is clear that in the case of high cost, the liquidity traders in the second period would benefit from the market with infinite liquidity, whereas those in the first period face very high price impact. Thus, this observation concludes that the timing of the liquidity shock is crucial from the perspective of the liquidity traders.

We, now, focus on the case of low cost and analyze the behavior of the equilibrium parameters with respect to the proxy for the likelihood of being the s-type insider, μ. Note that as shown in Proposition 2, the equilibrium is mainly described by the unique value of g which is, in turn, determined by μ. We investigate the relation between g and μ. This is important since such a relation allows us to generate empirical implications on the cryptocurrency market as a function of μ. The following lemma characterizes this relation.

Lemma 3.1. *The unique solution of g that pins down the equilibrium defined in Proposition 2 is increasing in the proxy for the likelihood of being the s-type insider, i.e., $\frac{\partial g}{\partial \mu} > 0$.*

Proof. See Appendix B.4. □

This relation allows us to set comparative statics in a tractable way.

Proposition 4. *(Comparative Statics) Suppose that the cost of using multiple accounts is sufficiently low, i.e., $c \leq \bar{c}$, then we have*

$$\frac{\partial \beta_{1,s}}{\partial \mu} < 0, \qquad \frac{\partial \beta_{1,m}}{\partial \mu} < 0, \qquad \frac{\partial \beta_2}{\partial \mu} < 0,$$

$$\frac{\partial \lambda_1}{\partial \mu} > 0, \qquad \frac{\partial \lambda_2}{\partial \mu} > 0.$$

Proof. This follows from Sturm's theorem. □

The intuition for the first period model parameters is that as the market maker assigns the higher conditional probability to the case that the insider is of s-type given the total order flow at $t = 1$, the price impact in the first period would be higher since the single account owner trades more aggressively than the multiple accounts owner, i.e., $\frac{\partial \lambda_1}{\partial \mu} > 0$. As

a result of such a higher price impact in the first period, both types of investors would lower their trade aggressiveness, i.e., $\frac{\partial \beta_{1,m}}{\partial \mu} < 0$, and $\frac{\partial \beta_{1,s}}{\partial \mu} < 0$.

The intuition for the price impact at $t = 2$ is that if the trader is of m-type, higher μ decreases her first period trade so that she would plan to compensate such decline in her trade at $t = 1$ by trading more at $t = 2$. As a result, having higher aggressive trading in period 2 leads to higher price impact at $t = 2$, i.e., $\frac{\partial \lambda_2}{\partial \mu} > 0$. There are two opposing effects for the second period trade aggressiveness for the m-type insider. (i) On the one hand, trading less in period 1 due to higher price impact (since μ is higher) pushes the m-type insider to trade more in period 2. (ii) On the other hand, once MM realizes that the insider is of m-type, MM sets the price impact higher since the m-type insider has an appetite for trading more in period 2. In equilibrium, the second effect dominates the first one since the higher μ leads to less trade aggressiveness in period 2 for the m-type insider, i.e., $\frac{\partial \beta_2}{\partial \mu} < 0$.

4. Price Discoveries

It is important to analyze price discoveries separately because it has implications on the crypto-currency market, which may help us to understand the excessive volatility of these currencies and severe price reversals. Price discoveries are thought of as proxies to show how well prices reflect the underlying the asset value i.e., $\Sigma_t = Var(\tilde{v} \mid y_t)$. A lower Σ_t indicates more price discovery or a more informative price at t about v.[20] We first derive the price discovery in the first period (short-term), $\Sigma_1 = Var(\tilde{v} \mid y_1)$.

$$\begin{aligned}
\Sigma_1 &= Var(\tilde{v} \mid y_1) = E(\tilde{v}^2 \mid y_1) - E^2(\tilde{v} \mid y_1) \\
&= P(\tilde{y}_{1,s} = y_1 \mid y_1) E(\tilde{v}^2 \mid \tilde{y}_{1,s} = y_1) + P(\tilde{y}_{1,m} = y_1 \mid y_1) E(\tilde{v}^2 \mid \tilde{y}_{1,m} = y_1) \\
&\quad - \left(P(\tilde{y}_{1,s} = y_1 \mid y_1) E(\tilde{v} \mid \tilde{y}_{1,s} = y_1) + P(\tilde{y}_{1,m} = y_1 \mid y_1) E(\tilde{v} \mid \tilde{y}_{1,m} = y_1)\right)^2 \\
&= \mu(\sigma_{1,s}^2 + \mu_{1,s}^2) + (1-\mu)(\sigma_{1,m}^2 + \mu_{1,m}^2) - \left(\mu\mu_{1,s} + (1-\mu)\mu_{1,m}\right)^2, \qquad (20)
\end{aligned}$$

[20] We use the terms price informativeness and price discovery interchangeably.

where $\mu_{1,s} = \frac{\beta_{1,s}}{\beta_{1,s}^2+\frac{\sigma_u^2}{\sigma_v^2}} y_1$, $\mu_{1,m} = \frac{\beta_{1,m}}{\beta_{1,m}^2+\frac{\sigma_u^2}{\sigma_v^2}} y_1$, $\sigma_{1,s}^2 = \frac{\sigma_u^2}{\beta_{1,s}^2+\frac{\sigma_u^2}{\sigma_v^2}}$, and $\sigma_{1,m}^2 = \frac{\sigma_u^2}{\beta_{1,m}^2+\frac{\sigma_u^2}{\sigma_v^2}}$.

We show that Σ_1 can be written as follows:

$$\begin{aligned}\Sigma_1 &= \mu\sigma_v^2 B_1(\mu) + (1-\mu)\sigma_v^2 B_2(\mu) \\ &\quad + \frac{y_1^2\sigma_v^2}{\sigma_u^2}\left[\mu B_3(\mu) + (1-\mu)B_4(\mu) + \mu^2 B_5(\mu) + (1-\mu)^2 B_6(\mu) + 2\mu(1-\mu)B_7(\mu)\right] \\ &= \sigma_v^2 \tilde{B}_1(\mu) + \frac{\sigma_v^2 \tilde{B}_2(\mu)}{\sigma_u^2} y_1^2, \end{aligned} \tag{21}$$

where the coefficients of B_i's are given in Appendix B.5. Above derivations lead to the following result.

Proposition 5. *Assume that $c < \bar{c}$.*

(i) Σ_1 is a function of both the first period total order flow, y_1, and the volatility of the liquidity shock, σ_u, if and only if $\mu \in (0,1)$.

Suppose that $\mu \in (0,1)$.

(ii) $\frac{\partial \Sigma_1}{\partial |y_1|} > 0$.

(iii) $\frac{\partial \Sigma_1}{\partial \sigma_u} < 0$.

(iv) $\frac{\partial^2 \Sigma_1}{\partial |y_1| \partial \sigma_u} < 0$.

(v) $\frac{\partial^2 \Sigma_1}{\partial |y_1| \partial \sigma_v} > 0$.

Proof. See Appendix B.6. □

Proposition 5 shows important channels through which non-fundamental uncertainty affects the first period price discovery. (i) implies that both the first period net order flow and the volatility of the liquidity shock have an impact on short-term price discovery if and only if non-fundamental uncertainty is present in the market, i.e., $\mu \in (0,1)$. (ii) demonstrates that the short-term price discovery decreases in absolute short-term total order flow when $\mu \in (0,1)$. The intuition is the following. When short-term individual trades are observable to the market maker, MM considers two cases whether the insider is of s-type or m-type in a probabilistic sense. Hence as the value of absolute total order flow increases, MM considers

possible values for the value of the asset which are further apart from each other. As a result of considering more dispersed values for the asset, pricing error gets larger. (iii) shows that the higher the volatility of the liquidity shock, the higher the short-term price discovery if and only if non-fundamental uncertainty is present in the market. The intuition is that higher volatility of the liquidity shock decreases the price impact. As a result of lower price impact, insiders have an incentive to trade more aggressively, yielding higher price discovery in the short-term. (iv) shows that the effect of the first period absolute total order flow on the first period price discovery decreases as the volatility of liquidity shocks increases. Finally, (v) shows that the effect of the first period absolute total order flow on the first period price discovery increases as the asset volatility increases.

Figure 2 provides further insights into the effects of model parameters on the relation between information revelation in the short-term and the first period total order flow. We plot the first period price discovery against the first period total order flow for a set of parameters. In all panels, σ_u, σ_v and μ are set to 1, 1, 0.5, respectively, unless it is set on the x-axis.

Top panels illustrate the effects of μ on that relation. In the top-left panel, we focus on the extreme values of μ. First and most importantly, the larger the amount of absolute total order flow, the less informative price in the first period when $\mu \in (0, 1)$. Second, when there is no non-fundamental uncertainty in the model, i.e., μ equals 0 or 1), total order flow has no effect on the price discovery. Third, the price informativeness in the first period when the population has only s-type insiders is higher than that when the population has only m-type insiders. This is because very aggressive trading by s-type insiders reveals more information in the first period. In the top-right panel, we focus on the interior values of μ. That panel illustrates that, for sufficiently low values of y_1, the higher μ leads to higher price discovery in the first period.

The bottom-left panel illustrates the effects of asset volatility on that relation. First, the

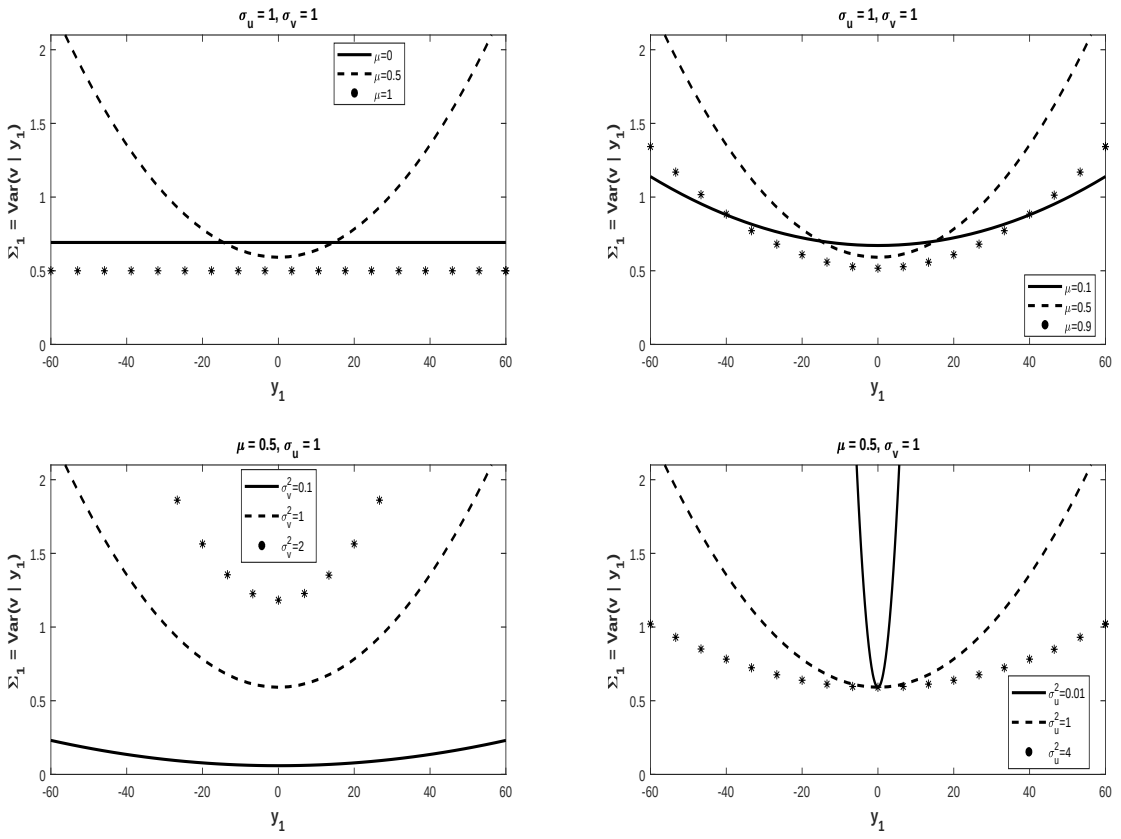

Fig. 2. This figure plots the first period price discovery against the first period total order flow, y_1, for several parameter specifications.

higher the asset volatility, the lower the price discovery. Intuitively, the higher asset volatility creates more information asymmetry between the insider and other market participants. Second and more importantly, as the asset volatility increases, PD_1 becomes more sensitive to y_1. As a result, high volatile assets are more likely to have severe price reversals (momentum) in the long-term if their short-term total order flow is high enough.

Finally, the bottom-right panel shows the effects of liquidity volatility on that relation. First, the higher the volatility of the liquidity shock, the higher the price informativeness. Surprisingly, the volatility of that helps to increase price informativeness. Second, as the volatility of that increases, price discovery curve becomes flatter. As a result, price discovery in the first period is less sensitive to the first period total order flow for higher values of

liquidity volatility.

We now derive the price discovery in the second period (long-term), $\Sigma_2 = Var(\tilde{v} \mid y_1, y_2)$. When there are N traders in the second period and, hence, the insider is of s-type, we have

$$\Sigma_2^s = Var(\tilde{v} \mid \tilde{y}_{1,s} = y_1, \tilde{u}_2 = y_2) = Var(\tilde{v} \mid \tilde{y}_{1,s} = y_1) = \frac{\sigma_u^2}{\beta_{1,s}^2 + \frac{\sigma_u^2}{\sigma_v^2}}.$$

Similarly, in the opposite case where the insider is of m-type, we have

$$\Sigma_2^m = Var(\tilde{v} \mid \tilde{y}_{1,m} = y_1, \tilde{y}_{2,m} = y_2) = \frac{\sigma_u^2}{\beta_2^2 + \beta_{1,m}^2 + \frac{\sigma_u^2}{\sigma_v^2}}.$$

Proposition 6. *Assume that* $c < \bar{c}$.

(i) $\Sigma_2^s > \Sigma_2^m$,

(ii) $\frac{\partial \Sigma_2^s}{\partial \mu} > 0$ *and* $\frac{\partial \Sigma_2^m}{\partial \mu} > 0$.

Proof. For part (i), it is sufficient to show that $\beta_{1,m}^2 + \beta_2^2 > \beta_{1,s}^2$, which directly follows from Sturm's theorem. For part (ii), it is sufficient to show that $\frac{\partial \beta_{1,s}}{\partial \mu} < 0, \frac{\partial \beta_{1,m}}{\partial \mu} < 0, \frac{\partial \beta_2}{\partial \mu} < 0$, which directly follow from Proposition 4. □

Proposition 6 shows that (i) long-term prices reflect more information if the informed investor is of m-type. This is because in the case of multiple accounts, even though the price discovery is lower in the short-term, trading by the insider continues in period 2 and such trading helps to reveal extra information. As a result, the total revelation exceeds the one with a single account in period 2. (ii) As μ increases, the market maker assigns more probability for being s-type and, hence, increases price impact. As a result, both types of insiders trade less aggressively due to low market depth and reveal less information, lowering the price discovery for both cases.

5. Implications

In line with the discussion we have in the previous section, we begin with the predictions regarding the price discoveries and then we move to the model prediction on volume. In the following predictions, we assume that population has both types of insiders, i.e., $\mu \in (0, 1)$, and the cost of using multiple accounts are sufficiently low, i.e., $c < \bar{c}$.

5.1. Price Discovery

In this subsection, we present both novel testable predictions and predictions that provide a theoretical explanation of the results documented in the literature.

Prediction 1: All else equal, short-term prices are less informative in the time of the extremely large absolute total order flow.

Prediction 2: All else equal, crypto-asset prices in exchanges with larger volatility of liquidity shocks are more informative in the short-term.

Prediction 3: All else equal, the short-term price discovery becomes less sensitive to short-term absolute total order flows in exchanges with a higher volatility of the liquidity shock.

Prediction 4: All else equal, the short-term price discovery becomes more sensitive to short-term absolute total order flows for crypto-assets with a higher asset volatility.

Predictions 1-4 follow from (ii)-(v) of Proposition 5, respectively. Prediction 1 can explain the initial high price of crypto-assets, followed by large price reversals (momentum) in the long-term. This finding stems from two facts. First, short-term prices are extremely less informative. Second, non-fundamental uncertainty disappears in the long-term because the second period price discovery does not depend on order flows. Therefore, prices would be set more efficiently. Consistent with this prediction, Li et al. (2019) find that Bitcoin prices peak within minutes, followed by quick reversals. Prediction 3(4) implies that these price reversals would be more severe in exchanges with a lower volatility of the liquidity shock (for crypto-

assets with a higher asset volatility). These predictions 3 and 4 as well as prediction 2 show that the effect of absolute total order flows on short-term price discovery is heterogeneous across both crypto-assets and exchanges. This finding provides a theoretical explanation of the results documented in Brandvold, Molnár, Vagstad, and Valstad (2015), showing that price discoveries vary across exchanges.

Prediction 5: All else equal, when absolute net order flows are sufficiently low, short-term crypto-asset prices are more informative about v in exchanges with higher μ.

This prediction follows from top-right panel of Figure 2. The reason is that higher likelihood of being the s-type insider given the total order flow, μ, leads to more aggressive trading in the first period.

As a result of predictions 1-5, we can conclude that μ, short-term absolute total order flows and the volatility of the liquidity shock are the main factors to determine the price discovery on exchanges. Therefore, these factors could be an explanation to have low price discovery and hence more volatile prices in the long-term. Furthermore, our model also predicts that, all else equal, because of varying these parameters across crypto-assets, price informativeness across crypto-assets in the same exchange can differ as well.

Prediction 6: All else equal, the long-term prices are more (less) informative if the insider is of m-type (s-type).

Prediction 7: All else equal, the long-term prices would be more informative as the likelihood of being the s-type insider given the total order flow decreases.

Predictions 6-7 follow from (i)-(ii) of Proposition 6, respectively. These predictions imply that the higher likelihood of being the s-type insider improves the price discovery whereas a realization that the insider is of s-type harms the price discovery.

5.2. Volume

Prediction 8: Given a fixed liquidity trading at $t = 1$, volume versus fundamental value relation is non-monotonic in the short-term.

Volume in the first period can be expressed as $volume_1 = \beta_1|v| + |u_1|$. Since trade aggressiveness for the s-type insider is higher than that for the m-type insider, all else equal, more volume in the first period does not have to be associated with high value of $|v|$. For example, a large volume of the insider trading could be due to the s-type insider with a relatively low v rather than the m-type insider with a relatively high v, and hence high volume could be misleading about the fundamental value.[21] Consistent with this, Balcilar et al. (2017) show that return-volume relation in Bitcoin has both nonlinearity and structural breaks.

6. Conclusion

Blockchain technology brings us new features that allow financial economist to analyze cases that have not been available before. For instance, in the public blockchain, the most common form of this technology, it is common to observe current and past holdings of investors. Since such transparency leads to information privacy concerns for investors, they try to keep their information safe as long as possible by different methods. The commonly used method is to rely on multiple accounts for their transactions.

In this paper, we extend Kyle (1985) in the public blockchain environment by introducing the transparency on individual level orders as well as past holdings. Following Ron and Shamir (2013), we assume that the population has two types of traders: a trader with a single account and a trader with multiple accounts. Unlike Kyle (1985) where the insider exploits her private information over multiple periods, in our model, the s-type insider prefers to not trade in the second period, and hence trades more aggressively in the first period.

When the cost of using multiple accounts is sufficiently high, m-type insiders cannot bear that cost, and hence trade only in the first period. The aggressive trading by both types of insiders improves (harms) the price informativeness in the short-term (long-term). When

[21] The relation between volume and expected returns in the short-term is well-documented in the traditional market. See, for example, Llorente, Michaely, Saar, and Wang (2002), Conrad, Hameed, and Niden (1994), Gervais, Kaniel, and Mingelgrin (2001), and Kaniel, Ozoguz, and Starks (2012).

the cost of that is sufficiently low, m-type insiders behave differently than s-type insiders, and smooth their trades across both periods. This heterogeneity in trade behaviors creates a second layer of uncertainty when the market maker sets short-term prices. In contrast to Kyle (1985), this non-fundamental uncertainty makes the short-term price discovery sensitive to the volatility of the liquidity shock, and short-term total order flows. As a result of this sensitiveness, we show that crypto-asset prices have lower informativeness as the short-term absolute total order flow increases, $|y_1|$. Therefore, such high $|y_1|$ should not mislead the market participants in the short-term, since price informativeness is quite low in the time of high $|y_1|$. We also present the benefit of uninformed trading in the crypto-asset market by showing that more randomness in uninformed trading, i.e., the higher volatility of the liquidity shock, leads to more informative prices in the short-term. This benefit gets larger as the short-term total order flow increases.

Due to the difference in aggressive trading, even relatively low volume in the short-term could be a sign of high absolute fundamental value. Similarly, relatively high volume in the short-term could be associated with low absolute fundamental value. Our model concludes that volume versus fundamental value relation is non-monotonic, and analysts in the crypto-asset market should be cautious when considering the level of volume as an explanatory variable for the expected return.

Appendix A. Discussion of Our Implicit Assumption

When the m-type insider uses multiple accounts, there is a way for the market maker to reveal insider's information although it is a computationally hard problem. The method is the following. First, MM knows the time when the m-type insider is present because of the fixed number of liquidity traders, N. Second, the linear strategies of the m-type is known and therefore, MM or computerized system can calculate $x_1/\beta_{1,m}$ for $N+1$ traders at $t=1$ and $(x_2-\beta_y y_1)/\beta_2$ for $N+1$ traders at $t=2$. Notice that both calculations are equal to v for the m-type insider. Therefore, keeping track of the transactions of all traders across periods would reveal both accounts of the m-type insider. These operations have a time complexity of the order of N^2 $[O(N^2)]$, and hence a very costly process. To give some sense how costly it could be for the case of average daily Bitcoin transactions, if we assume that revealing the s-type insider takes one second by using the method with complexity of the order of N $[O(N)]$, revealing the m-type insider takes, on average, more than 5 days. Therefore, we assume that MM cannot identify the m-type even though MM knows her existence in the market. However, such possibility of revelation of the m-type insider would disappear if we modify the model in the following way: Assume that the insider has a noisy signal in the first period but perfect information before the second period. Our main results still hold in such a setting.

Appendix B. Proofs

B.1. Proof of Proposition 1

When the cost of using multiple accounts is equal to the threshold level, $\bar{c}$, the m-type insider is indifferent between trading in both periods and trading only in the first period. When she trades in both periods, her ex-ante expected profit becomes $E[\Pi_m]$ as in equation (16). When she trades only in the first period, her ex-ante expected profit becomes $\sigma_u\sigma_v/2$.

Equality of these expressions pins down the threshold level, $\bar{c}$ in terms of $\lambda_1, \lambda_2, \gamma_1, \sigma_u$ and σ_v as follows:

$$\bar{c} = \left(\frac{\lambda_2 + \lambda_1 - \gamma_1}{4\lambda_1\lambda_2 - \gamma_1^2} \right) \sigma_v^2 + \frac{\gamma_1^2 \sigma_u^2}{4\lambda_2} + \frac{\sigma_u \sigma_v}{2}.$$

From equation (B.7), we know that γ_1 is a function of only g, σ_u and σ_v. From the definition of g and h, λ_1 and λ_2 can be expressed in terms of only g, σ_u and σ_v, as well. We insert these three variables, $\gamma_1, \lambda_1, \lambda_2$, into above equation, which completes the proof.

B.2. Proof of Proposition 2

For the case of sufficiently high cost of using multiple accounts, insiders trade only in the first period, and hence first period model parameters would be the same as one-period Kyle (1985) model. In equilibrium, we have

$$\beta_1 = \frac{\sigma_u}{\sigma_v}, \qquad \lambda_1 = \frac{\sigma_v}{2\sigma_u}, \qquad \lambda_2 = 0.$$

Hence, their ex-ante expected profit becomes $\sigma_u \sigma_v / 2$. For the remainder of this proof, we need to consider the sufficiently low cost of using multiple accounts. For this purpose, we first characterize the equilibrium with one polynomial equation and then attempt to prove the existence and uniqueness of the model.

B.2.1. Characterization

Once we express that $\beta_{1,m}$, and γ_1 in terms of our model parameters, the other variables are determined straightforwardly. To do so, we have 5 steps to prove our statement.

Step1. Find the expressions for $\beta_{1,m}$ and γ_1 in terms of g and h.

Step2. Show that $h = 2g^2 - g + \frac{1}{4g}$.

Step3. Find the equation characterizing the parameter μ.

Step4. Show that $g > \frac{1}{2}$, $h > \frac{1}{2}$, and $\gamma_1 > 0$.

Step1. First, by inserting equation (9) into (6), we have:

$$\beta_2^2 = \beta_{1,m}^2 + \frac{\sigma_u^2}{\sigma_v^2}. \tag{B.1}$$

By dividing equation (5) by (6) and using (10), we have:

$$\beta_{1,m} = -\beta_2 \beta_y. \tag{B.2}$$

By inserting equation (B.1-B.2) into (5), we can express γ_1 as follows:

$$\gamma_1 = \frac{\beta_{1,m}}{\beta_{1,m}^2 + \frac{\sigma_u^2}{\sigma_v^2}}, \tag{B.3}$$

alternatively, we also have:

$$\beta_{1,m}\gamma_1 = \frac{\beta_{1,m}^2}{\beta_{1,m}^2 + \frac{\sigma_u^2}{\sigma_v^2}}. \tag{B.4}$$

Second, by multiplying equation (14) with (γ_1), and using the definition of g and h we have:

$$\beta_{1,m}\gamma_1 = \frac{2g-1}{4gh-1}. \tag{B.5}$$

After solving $\beta_{1,m}$ by using equation (B.4) and (B.5), and inserting it into equation (B.5), we get the expressions of both $\beta_{1,m}$ and γ_1 as follows:

$$\beta_{1,m} = \sqrt{\frac{2g-1}{2g(2h-1)}}\frac{\sigma_u}{\sigma_v}, \tag{B.6}$$

$$\gamma_1 = \frac{\sqrt{2g(2h-1)(2g-1)}}{4gh-1}\frac{\sigma_v}{\sigma_u}. \tag{B.7}$$

The above equations are valid when $g > \frac{1}{2}$ and $h > \frac{1}{2}$. However, these conditions are satisfied in equilibrium as shown below.

Now, we can derive the SOC's in terms of g and h. The first two SOC's are directly

determined by definition of g and h. For the third SOC, by assuming $\gamma_1 > 0$, we can divide this SOC by γ_1^2, and get the desired result. Again, we prove, later on, that this assumption is true in equilibrium.

Step2. Given equation $(B.1)$ and $\beta_2 = \frac{1}{2\lambda_2} = \frac{1}{2g\gamma_1}$, where γ_1 is given by equation $(B.7)$, we have

$$\frac{1}{4}\frac{(4gh-1)^2}{2g^3(2h-1)(2g-1)} = \frac{4gh-1}{2g(2h-1)}.$$

Further simplification shows that

$$h = 2g^2 - g + \frac{1}{4g}. \tag{B.8}$$

Step3. So far, we have not used the equation $(B.6)$. It is used to characterize g. First, by dividing equation $(B.6)$ by γ_1, we have

$$h = \frac{\mu\beta_{1,s}}{\beta_{1,s}^2 + \frac{\sigma_u^2}{\sigma_v^2}}\frac{1}{\gamma_1} + 1 - \mu.$$

By using the expression $\beta_{1,s} = \frac{1}{2\lambda_1} = \frac{1}{2h\gamma_1}$, we alternatively have

$$h - 1 + \mu = \frac{2\mu h}{1 + 4\lambda_1^2\frac{\sigma_u^2}{\sigma_v^2}}.$$

We can also express λ_1 in terms of g and h by using equation $(h\gamma_1)$. After manipulation we get the following:

$$2h^4 + (2\mu - 3)h^3 + (1-\mu)h^2 - 2(2\mu - 1)hg^3(2g-1) - 2(1-\mu)g^3(2g-1) = 0.$$

First notice that $g = \frac{1}{2}$ is a solution to this system regardless of the value of μ. However, as shown in the next step, in equilibrium, g is higher than $\frac{1}{2}$. Hence, we can factor out $(g - \frac{1}{2})$ from this equation as well as multiplying with $128g^4$, after inserting the value of h. Then,

the equation can be written as follows:

$$F(\mu, g) = \sum_{i=0}^{11} A_i g^i = 0, \tag{B.9}$$

where the coefficients are given as follows:

$$\begin{aligned}
A_0 &= -1, \\
A_1 &= (4 - 4\mu), \\
A_2 &= 16, \\
A_3 &= -(72 - 48\mu), \\
A_4 &= -(64\mu + 32), \\
A_5 &= (480 - 192\mu), \\
A_6 &= -384(1 - \mu), \\
A_7 &= -(1408 - 256\mu), \\
A_8 &= (2048 - 512\mu), \\
A_9 &= 512, \\
A_{10} &= -3072, \\
A_{11} &= 2048.
\end{aligned}$$

This completes the step 3.

Step4. We, now, turn our attention to show that the assumption we state earlier on g being higher than $\frac{1}{2}$. The third SOC becomes $4g^2(2g-1) > 0$ after we insert h into it. This condition directly shows us that $g > \frac{1}{2}$. Now, to show $h > \frac{1}{2}$ holds, it is sufficient to show that $h - \frac{1}{2} > 0$ for any $g > \frac{1}{2}$. $h - \frac{1}{2} = 2g^2 - g + \frac{1}{4g} - \frac{1}{2} = \frac{(2g-1)^2(2g+1)}{4g} > 0$ for any $g > \frac{1}{2}$. We are left to show that $\gamma_1 > 0$. Since g is positive, from the second SOC, $g\gamma_1 > 0$, we

directly conclude that $\gamma_1 > 0$.

B.2.2. Existence

Since $g > \frac{1}{2}$ in equilibrium, any candidate g should satisfy this condition. We argue that there always exists a real solution in the interval $(0.90, 0.95)$. Let's define $F(\mu, x) = \sum_{i=0}^{11} A_i x^i$, which is given by equation $(B.9)$. $F(\mu, 0.90) = -\frac{6974352\mu}{390625} - \frac{24575681}{48828125}$, which is negative for any $\mu \in [0, 1]$. Similarly, $F(\mu, 0.95) = \frac{4228090380089}{100000000000} - \frac{2098098351\mu}{50000000}$, which is positive for any $\mu \in [0, 1]$. By using the mean value theorem (MVT), there exists at least one real solution to this system in the interval $(0.90, 0.95)$.

B.2.3. Uniqueness

First, note that for the remaining parts, all the functions that we applied Sturm's theorem are square-free. Showing the function f is square-free is then straightforward. It is sufficient to show that the greatest common divisor of f and f' is 0. Before attempting to prove the uniqueness, we present the following lemmata in which we use repeatedly.

Lemma B.1. *Consider the function*

$$f(x) = -32x^5 + 16x^4 + 16x^3 - 8x^2 + 1 = 8x^2(2x^2 - 1)(1 - 2x) + 1.$$

Then there exists unique real solution $x^\star$ in the interval $[0.5, 0.9]$ such that $f(x) < 0$ for any $x \in (x^\star, 0.9)$, and $f(x) > 0$ for any $x \in (0.5, x^\star)$.

Proof. We need to show four facts: i.) and ii.) the existence and the uniqueness of such $x^\star$, iii.) and iv.) f always takes negative [positive] values in $(x^\star, 0.9)$ $[(0.5, x^\star)]$.

i.) Clearly, $f(0.5) = 1 > 0$ and $f(0.9) < 0$. By MVT, there exist $x^\star$ such that $f(x^\star) = 0$.

ii.) By using Sturm's theorem, it is straightforward to show that there exist only one real solution in the domain $(0.5, 0.9)$.

iii.) and iv.) Since f has a unique real solution, it is sufficient to show the sign of f is positive at 0.5 and negative at 0.9, that is already checked in the first step. □

Lemma B.2. *Consider the same function as in Lemma B.1, then f takes only negative values for $x > 0.95$.*

Proof. It is sufficient to show that $f(0.95) = -4.2309 < 0$ and $f(x) = 0$ has no real solution in the interval $(0.95, \infty)$. The latter follows directly from Sturm's theorem. □

Lemma B.3. *Unique $x^\star$ for the same function as in Lemma B.1 is between 0.8 and 0.81.*

Proof. $f(0.8) = 0.1398$ and $f(0.81) = -0.016$. The MVT ensures that $x^\star$ is in the desired interval. □

Now, we are ready to start our proof. First of all, we know that the real solution should satisfy $g > 1/2$. We also know that we have a real solution in the interval $(0.9, 0.95)$. Then, to prove that there is no other real solution for $g > 0.5$, we need to show the followings:

1.) Show that $F(\mu, g) < 0$ for any $g \in (0.5, 0.9)$ and for any $\mu \in [0, 1]$.

2.) Show that $F(\mu, g)$ has no real solution for any $g \geq 0.95$ and for any $\mu \in [0, 1]$.

3.) Show that for any $\mu \in [0, 1]$, $F(\mu, g)$ is increasing in g where $g \in [0.9, 0.95]$.

Before moving to step 1, we first decompose $F(\mu, g) = F_1(g) + \mu \tilde{F}_2(g)$ where

$F_1(g) = 2048g^{11} - 3072g^{10} + 512g^9 + 2048g^8 - 1408g^7 - 384g^6 + 480g^5 - 32g^4 - 72g^3 + 16g^2 + 4g - 1$,

$\tilde{F}_2(g) = -512g^8 + 256g^7 + 384g^6 - 192g^5 - 64g^4 + 48g^3 - 4g$.

Step 1

Notice that $F_2(g) \equiv \frac{\tilde{F}_2(g)}{16g(g-0.5)(g+0.5)} = -32g^5 + 16g^4 + 16g^3 - 8g^2 + 1$ as in Lemma B.1. Notice also that $16F_2(g) > \tilde{F}_2(g)$ for any $g \in (0.5, 0.9)$. This is since $(g - 0.5)(g + 0.5) < 1$ in that interval. Then, we conclude that for any $g \in (0.5, g^\star)$ and for any $\mu \in [0, 1]$, we have:

$F(\mu, g) = F_1(g) + \mu \tilde{F}_2(g) < F_1(g) + 16\mu F_2(g) < F_1(g)$, since $F_2(g) < 0$ in that interval. It is now sufficient to show that $F_1(g) < 0$ for any $g \in (0.5, g^\star)$.

Similarly, for any $g \in (g^\star, 0.9)$ and for any $\mu \in [0, 1]$, we have:

$F(\mu, g) = F_1(g) + \mu\tilde{F}_2(g) < F_1(g) + 16\mu F_2(g) < F_1(g) + 16F_2(g)$, since $F_2(g) > 0$ in that interval. It would be sufficient to show that $F_1(g) + F_2(g) < 0$ for any $g \in (g^\star, 0.9)$.

Step 1.1 Show that $F_1(g) < 0$ for any $g \in (0.5, g^\star)$.

It is sufficient to show that $F_1(0.5) < 0$ and $F_1(g) = 0$ has no real solution for any $g \in (0.5, 0.81)$. Here, we have looked for 0.81 rather than $g^\star$ since $0.81 > g^\star$ as in Lemma B.3. The direct calculation and applying Sturm's theorem show that $F_1(0.5) = -1$ and there is no real solution in that interval.

Step 1.2 Show that $F_1(g) + 16F_2(g) < 0$ for any $g \in (0.8, 0.9)$.

Similar to what we did in the previous case, we have $F_1(g) + 16F_2(g) = 2048g^{11} - 3072g^{10} + 512g^9 + 1536g^8 - 1152g^7 - 224g^5 + 160g^4 + 232g^3 - 112g^2 + 15$. Clearly, $F_1(0.8) + 16F_2(0.8) = -\frac{725895693}{48828125} < 0$, and Sturm's theorem guarantees that there is no real solution in that interval.

Step 2

Let's define the domain $D = \{(\mu, g) \mid \mu \in [0, 1] \quad \text{and} \quad g \in [0.95, \infty)\}$. It is then sufficient to show that $F(\mu, 0.95) > 0$ for any $\mu \in [0, 1]$, and $F(\mu, g) > 0$ for any $(\mu, g) \in D$. Clearly, $F(\mu, 0.95) = 42.2809 - 41.962\mu > 0$ for any $\mu \in [0, 1]$. Then we are left to show that $F(\mu, g) = F_1(g) + \mu 16g(g - 0.5)(g + 0.5)F_2(g) > 0$ for any $(\mu, g) \in D$. From Lemma B.2.3 we know that $F_2(g) > 0$ for any $g > 0.95$. Hence, it is now sufficient to show that $F_1(g) + F_2(g) > 0$ for any $g > 0.95$. Equivalently, it is sufficient to show that $F_1(0.95) + F_2(0.95) = 0.3189 > 0$ and $F_1(g) + F_2(g) = 0$ has no real root in the interval $[0.95, \infty)$. The latter again follows directly from Sturm's theorem.

Step 3

Let's define the domain $D = \{(\mu, g) \mid \mu \in [0, 1] \quad \text{and} \quad g \in [0.9, 0.95]\}$. It is then sufficient to show that $\frac{\partial F(\mu,g)}{\partial g} = F_g(\mu, g) > 0$ for any $(\mu, g) \in D$. Similar to what we did in step 1, we have $F_g(\mu, g) = G_1(g) + \mu G_2(g)$ where

$G_1(g) = 22528g^{10} - 30720g^9 + 4608g^8 + 16384g^7 - 9856g^6 - 2304g^5 + 2400g^4 - 128g^3 - 216g^2 + 32g + 4$,

$G_2(g) = -4096g^7 + 1792g^6 + 2304g^5 - 960g^4 - 256g^3 + 144g^2 - 4$.

It is clear that $G_2(0.9) < 0$ and $G_2(g) = 0$ has no real solution in the interval $[0.9, 0.95]$. Hence, $G_2(g) < 0$ for any $g \in [0.9, 0.95]$. Therefore, it is sufficient to show that $G_1(g) + G_2(g) = F_g(1, g) > 0$ for any $g \in [0.9, 0.95]$. It is then sufficient to show that $F_g(1, 0.9) = \frac{1603158822}{9765625} > 0$ and $F_g(1, g) = 0$ has no real solution in $g \in [0.9, 0.95]$. The latter is again directly from Sturm's theorem. This last point concludes both the proof of the step 3 and the uniqueness.

B.3. Proof of Proposition 3

B.3.1. Top-left comparison

For the top-left comparison, we need to show that $\frac{1}{2\lambda_1} > \frac{\sigma_u}{\sigma_v}$ for any $\mu \in (0, 1)$. Since g is an increasing function of μ and g takes values in the interval $[0.9, 0.95]$, it is sufficient to show that

$$\frac{8g^3}{\sqrt{4g^2 - 1}(8g^3 - 4g^2 + 1)} - 1 > 0 \qquad \text{for any} \qquad g \in [0.9, 0.95].$$

Since $g > 0.5$, the denominator is positive. It is then sufficient to show that

$$8g^3 > \sqrt{4g^2 - 1}(8g^3 - 4g^2 + 1) \qquad \text{for any} \qquad g \in [0.9, 0.95].$$

Since both sides are positive for any $g > 0.5$, after taking the square of both sides it is sufficient to show that

$$-256g^8 + 256g^7 + 64g^6 - 128g^5 + 48g^4 + 16g^3 - 12g^2 + 1 > 0 \qquad \text{for any} \qquad g \in [0.9, 0.95] \tag{B.10}$$

Our conjecture is that left side of the inequality (B.10) equals 0 when $\mu = 1$ and holds for any $\mu \in [0, 1)$. To prove our conjecture, we show the followings:

(i) There is only one real root for the left side of the above inequality in the interval

$[0.9, 0.95]$.

(ii) The inequality (B.10) is satisfied for $\mu = 0$, (indeed any $\mu \in [0, 1)$ works for this part).

(iii) The left side of the inequality (B.10) equals 0 when $\mu = 1$.

(i) This part directly follows from Sturm's theorem.

(ii) To find the unique value of g in the interval $[0.9, 0.95]$ when $\mu = 0$, we use the equation (B.9). The direct computation shows that $g^*(\mu = 0) \equiv 0.901$.Then inequality holds since the left side becomes approximately 5.049.

(iii) The difficulty part is that when $\mu = 1$, we cannot consider the approximate value of g^* which solves the $F(1, g(1)) = 0$. Instead, we use the polynomial itself, $F(1, g(1))$, to consider g^*. We can decompose the $F(1, g(1))$ as follows:

$$F(1, g(1)) = \left(g^3 - \frac{g^2}{2} + \frac{1}{8}\right)\left(g^8 - g^7 - \frac{g^6}{4} + \frac{g^5}{2} - \frac{3g^4}{16} - \frac{g^3}{16} + \frac{3g^2}{64} - \frac{1}{256}\right).$$

It is straightforward to show that the real roots of the first component is not in the interval $[0.9, 0.95]$. As a result, $\left(g^8 - g^7 - \frac{g^6}{4} + \frac{g^5}{2} - \frac{3g^4}{16} - \frac{g^3}{16} + \frac{3g^2}{64} - \frac{1}{256}\right) = 0$ gives the desired g^*. Notice that g^* also makes the left side of the inequality (B.10) 0, which completes the proof.

B.3.2. Top-right comparison

For the top-right comparison, we need to show that $\frac{\sigma_u}{\sigma_v} > \sqrt{\frac{2g-1}{2g(2h-1)}}\frac{\sigma_u}{\sigma_v}$ for any $\mu \in (0, 1)$. Equivalently, we need to show that $1 > \sqrt{\frac{1}{4g^2-1}}$ for any $\mu \in (0, 1)$, which directly follows from the fact that $g > 0.5$.

B.3.3. Bottom-left comparison

For the bottom-left comparison, we need to show that $\frac{\sigma_v}{2\sigma_u} > \frac{\sqrt{4g^2-1}(8g^3-4g^2+1)}{16g^3}\frac{\sigma_v}{\sigma_u}$ for any $\mu \in (0, 1)$. Equivalently, we need to show that $8g^3 > \sqrt{4g^2 - 1}(8g^3 - 4g^2 + 1)$ for any $\mu \in (0, 1)$, which directly follows from the proof of top-left comparison.

B.3.4. Bottom-right comparison

It is straightforward to show that the comparison holds for any $g > 0.5$.

B.4. Proof of Lemma 3.1

We know that the unique g is the solution to the equation $F(\mu, g) = 0$. By using the implicit function theorem, we have:

$$\frac{\partial F}{\partial \mu} + \frac{\partial F}{\partial g}\frac{\partial g}{\partial \mu} = 0,$$

which implies that

$$\frac{\partial g}{\partial \mu} = -\frac{\frac{\partial F}{\partial \mu}}{\frac{\partial F}{\partial g}},$$

where $\frac{\partial F}{\partial \mu} = -512g^8+256g^7+384g^6-192g^5-64g^4+48g^3-4g = 16g(g-0.5)(g+0.5)(-32g^5+16g^4+16g^3-8g^2+1)$ and $\frac{\partial F}{\partial g} = F_g(\mu, g)$ is given explicitly in step 3 of the proof of uniqueness. For the latter one, we have already shown that it is positive for any $(\mu, g) \in D$ where $D = \{(\mu, g) \mid \mu \in [0, 1] \quad \text{and} \quad g \in [0.9, 0.95]\}$. Now we are left to show that $F_\mu(\mu, g) \equiv \frac{\partial F}{\partial \mu} > 0$ for any $(\mu, g) \in D$. Since $F_\mu(\mu, g)$ is only a function of g and also we know that the unique value of g is between 0.9 and 0.95, it is sufficient to show that $F_\mu(\mu, g) < 0$ for any $g \in [0.9, 0.95]$. It is also sufficient to show that $f(g) = (-32g^5+16g^4+16g^3-8g^2+1) < 0$ for any $g \in [0.9, 0.95]$ since the unique value of g is higher than 0.5. What remains is to show that $f(g) < 0$. This follows from Lemma B.1, because the unique solution of this polynomial is between $[0.5, 0.9]$ and $f(0.95) < 0$. This completes the proof.

B.5. Coefficients of Price Discovery

Let's denote the auxiliary variables C_i's as follows:

$$C_1 = 256g^8 - 256g^7 + 64g^6 + 128g^5 - 48g^4 - 16g^3 + 12g^2 - 1,$$

$$C_2 = 4g^2 - 1, \qquad \text{and} \qquad C_3 = 8g^3 - 4g^2 + 1.$$

Then B_i's can be expressed as follows:

$$B_1 = \frac{C_2 C_3^2}{C_1}, \qquad B_2 = \frac{C_2}{4g^2}, \qquad B_3 = \frac{64g^6 C_2 C_3^2}{C_1^2}, \qquad B_4 = \frac{C_2}{16g^4},$$

$$B_5 = -\frac{64g^6 C_2 C_3^2}{C_1^2}, \qquad B_6 = -\frac{C_2}{16g^4}, \qquad B_7 = -\frac{2gC_2C_3}{C_1}.$$

Equivalently, we get the expression of

$$\tilde{B}_1(\mu) = \mu B_1 + (1-\mu)B_2, \qquad \tilde{B}_2(\mu) = \mu B_3 + (1-\mu)B_4 + \mu^2 B_5 + (1-\mu)^2 B_6 + 2\mu(1-\mu)B_7.$$

B.6. Proof of Proposition 5

For part (i), we need to show that $\tilde{B}_2(\mu) > 0$ for any $\mu \in (0,1)$, and 0 otherwise.

$$\tilde{B}_2(\mu) = \frac{\mu(1-\mu)(4g^2-1)^3(-32g^5+16g^4+16g^3-8g^2+1)^2}{16g^4(256g^8-256g^7+64g^6+128g^5-48g^4-16g^3+12g^2-1)^2} > 0.$$

We know that when $\mu \in (0,1)$, g takes values between 0.9 and 0.95. Notice that any components with g's are positive when $g \in [0.9, 0.95]$. It is straightforward to show that when μ is either 0 or 1, $\tilde{B}_2(\mu)$ is equal to 0. Notice that denominator in $\tilde{B}_2(\mu)$ never equals 0 when $\mu \in [0,1]$, which completes the proof.

The other parts follow from direct calculation.

References

Admati, A. R., Pfleiderer, P., 1988. A theory of intraday patterns: Volume and price variability. The Review of Financial Studies 1, 3–40.

Balcilar, M., Bouri, E., Gupta, R., Roubaud, D., 2017. Can volume predict bitcoin returns and volatility? a quantiles-based approach. Economic Modelling 64, 74–81.

Bhambhwani, S., Delikouras, S., Korniotis, G. M., 2019. Do fundamentals drive cryptocurrency prices? Unpublished Working Paper.

Biais, B., Bisiere, C., Bouvard, M., Casamatta, C., 2019. The blockchain folk theorem. The Review of Financial Studies 32, 1662–1715.

Biais, B., Bisiere, C., Bouvard, M., Casamatta, C., Menkveld, A. J., 2018. Equilibrium bitcoin pricing. Unpublished Working Paper.

Brandvold, M., Molnár, P., Vagstad, K., Valstad, O. C. A., 2015. Price discovery on bitcoin exchanges. Journal of International Financial Markets, Institutions and Money 36, 18–35.

Brunnermeier, M. K., 2001. Asset Pricing under Asymmetric Information: Bubbles, Crashes, Technical Analysis, and Herding. Oxford University Press, Oxford.

Brunnermeier, M. K., 2005. Information leakage and market efficiency. The Review of Financial Studies 18, 417–457.

Buterin, V., 2016. Ethereum: Platform review, opportunities and challenges for private and consortium blockchains. Discussion paper. Ethereum Foundation.

Catalini, C., Gans, J. S., 2016. Some simple economics of the blockchain. Unpublished Working Paper.

Chod, J., Trichakis, N., Tsoukalas, G., Aspegren, H., Weber, M., 2020. On the financing benefits of supply chain transparency and blockchain adoption. Management Science .

Collin-Dufresne, P., Fos, V., 2016. Insider trading, stochastic liquidity, and equilibrium prices. Econometrica 84, 1441–1475.

Cong, L. W., He, Z., 2019. Blockchain disruption and smart contracts. The Review of Financial Studies 32, 1754–1797.

Conrad, J. S., Hameed, A., Niden, C., 1994. Volume and autocovariances in short-horizon individual security returns. The Journal of Finance 49, 1305–1329.

Easley, D., O'Hara, M., Basu, S., 2019. From mining to markets: The evolution of bitcoin transaction fees. Journal of Financial Economics 134, 91–109.

Farboodi, M., Veldkamp, L., 2019. Long run growth of financial data technology. Unpublished Working Paper.

Gandal, N., Hamrick, J., Moore, T., Oberman, T., 2018. Price manipulation in the bitcoin ecosystem. Journal of Monetary Economics 95, 86–96.

Gervais, S., Kaniel, R., Mingelgrin, D. H., 2001. The high-volume return premium. The Journal of Finance 56, 877–919.

Holden, C. W., Subrahmanyam, A., 1992. Long-lived private information and imperfect competition. The Journal of Finance 47, 247–270.

Huddart, S., Hughes, J. S., Levine, C. B., 2001. Public disclosure and dissimulation of insider trades. Econometrica 69, 665–681.

Kaniel, R., Ozoguz, A., Starks, L., 2012. The high volume return premium: Cross-country evidence. Journal of Financial Economics 103, 255–279.

Kyle, A. S., 1985. Continuous auctions and insider trading. Econometrica pp. 1315–1335.

Li, T., Shin, D., Wang, B., 2019. Cryptocurrency pump-and-dump schemes. Unpublished Working Paper.

Llorente, G., Michaely, R., Saar, G., Wang, J., 2002. Dynamic volume-return relation of individual stocks. The Review of financial studies 15, 1005–1047.

Makarov, I., Schoar, A., 2019. Price discovery in cryptocurrency markets. In: *AEA Papers and Proceedings*, vol. 109, pp. 97–99.

Malinova, K., Park, A., 2017. Market design with blockchain technology. Unpublished Working Paper.

O'Hara, M., 1995. Market Microstructure Theory, vol. 108. Blackwell Publishers Cambridge, MA.

Pagnotta, E., Buraschi, A., 2018. An equilibrium valuation of bitcoin and decentralized network assets. Unpublished Working Paper.

Ron, D., Shamir, A., 2013. Quantitative analysis of the full bitcoin transaction graph. In: *International Conference on Financial Cryptography and Data Security*, Springer, pp. 6–24.

Saleh, F., 2019. Volatility and welfare in a crypto economy. Unpublished Working Paper.

Sockin, M., Xiong, W., 2018. A model of cryptocurrencies. Unpublished Working Paper.

Van Kervel, V., Menkveld, A. J., 2019. High-frequency trading around large institutional orders. The Journal of Finance 74, 1091–1137.

Yang, L., Zhu, H., 2020. Back-running: Seeking and hiding fundamental information in order flows. The Review of Financial Studies 33, 1484–1533.

Yermack, D., 2017. Corporate governance and blockchains. Review of Finance 21, 7–31.

Zhu, C., 2018. Big data as a governance mechanism. Unpublished Working Paper.

Zimmerman, P., 2020. Blockchain structure and cryptocurrency prices. Bank of England Working Paper.