Price Discovery on Decentralized Exchanges[1]

Agostino Capponi Ruizhe Jia Shihao Yu

March 15, 2023

[1]Agostino Capponi, Columbia University, Department of Industrial Engineering and Operations Research, ac3827@columbia.edu. Ruizhe Jia, Columbia University, Department of Industrial Engineering and Operations Research, rj2536@columbia.edu. Shihao Yu, Columbia University, Department of Industrial Engineering and Operations Research, sy3119@columbia.edu. We thank Darrel Duffie, Joel Hasbrouck, Sean Foley, Tālis Putniņš, Bart Zhou Yueshen, Marius Zoican, Konstantin Sokolov, and seminar participants at the Microstructure Exchange (Asia-pacific edition), Hong Kong Baptist University, Luohan Academy, Oxford University DeFi Working Group, Chaincode, and The University of Memphis for helpful comments. An earlier version of the paper circulated under the title "The Information Content of Blockchain Fees".

Price Discovery on Decentralized Exchanges

Abstract

In contrast to centralized exchanges, decentralized exchanges (DEXs) process orders in discrete time and require traders to bid a priority fee to determine the execution priority. We employ a structural vector-autoregressive model to provide evidence that the priority fee reveals the private information of a DEX trade, contributing to price discovery. A one standard deviation shock in the high-fee DEX trade flow leads to a permanent price impact between 4.27 and 8.16 basis points. We show that informed traders bid high fees not only to reduce execution risk but also to compete with each other. Using a unique dataset of Ethereum mempool orders, we lend support to the hypothesis that informed traders primarily compete on DEXs following a jump-bidding strategy.

1 Introduction

Price discovery, the process through which market participants reach a consensus about the fundamental value of an asset, is a key function of financial markets. How such a process realizes has been a central topic in market microstructure, and it largely depends on various aspects of asset trading including market structure (e.g., lit versus dark trading as in Zhu (2014); centralized versus OTC market as in Hagströmer and Menkveld (2019)), transparency rule (e.g., pre-trade transparency versus post-trade transparency, as in Bloomfield and O'Hara (1999); Boehmer, Saar, and Yu (2005)), new market participants (e.g., high-frequency traders as in Brogaard, Hendershott, and Riordan (2014) and trading constraints (e.g., short sell ban as in Boehmer and Wu (2013)).

Decentralized exchanges (DEXs) are trading venues built on public blockchains. They enable trading of digital assets without the need for centralized intermediaries, and have gained a sizable volume and market share since their inception.[1] Compared with centralized exchanges (CEXs), DEXs have two distinct features due to their reliance on the public blockchain infrastructure. First, unlike CEXs which typically operate a limit order book (LOB) and match buy and sell orders in real-time, most DEXs use an automated market maker (AMM) and execute orders in discrete time. The execution sequence of pending transactions is determined by the priority fee that traders on DEXs bid for their orders. The higher the bid fee and the earlier the order will be executed. Second, pending orders on DEXs are publicly visible, including information about the order size and the attached priority fee. In contrast, on CEXs with pre-trade transparency, only resting limit orders are public information. Incoming market orders can only be inferred from executions ex-post.

We study whether and how the unique trading mechanism and transparency rules of DEXs impact the price discovery process. Does priority fee bidding affect the trading strategy of informed traders? Does this fee, a crucial feature of DEX trades, convey any information? It remains to be seen ex-ante whether informed traders would bid a high fee to execute their orders. On the one

[1] The aggregated volume of spot trading in cryptocurrencies on DEXs hovers between 50 billion USD and 200 billion USD. These figures correspond to a market share between 10% and 20%. See `https://www.theblock.co/data/decentralized-finance/dex-non-custodial` for details.

hand, informed traders, especially the ones with short-lived and high-value information, would like to bid a higher fee to speed up their order execution. On the other hand, bidding a high fee might leak their information to predators (Brunnermeier and Pedersen, 2005) or back-runners (Yang and Zhu, 2020). Our goal is to shed light on these questions through an empirical market microstructure analysis.

We construct a data set consisting of executed trade data from Uniswap (the largest DEX) and Binance (the largest CEX), and with pending orders data from the public mempool of the Ethereum blockchain. We focus on six most traded token pairs during our sample period between November 18, 2020, and February 10, 2021. Using the executed trade data, we run a structural vector-autoregressive (structural VAR) model (Hasbrouck, 1991a; Hasbrouck, 1991b) to investigate the trade informativeness of DEX trade flows with different fee levels. The tick-by-tick mempool data tracks the complete history of all order submissions and modifications on the Ethereum network. Thus, it allows us to investigate in detail the fee-bidding strategy of competing traders on DEXs.

Our main findings are summarized below. We find that DEX trade flows with high fees reveal private information, and more so compared to DEX trade flow with low fees: for token pairs involving a non-stablecoin (e.g., Ethereum and Bitcoin), the permanent price impact of the high-fee DEX trade flow ranges between 4.27 and 8.16 basis points, while for the low-fee DEX trade flow it is between 0.41 and 0.94 basis points. These estimates remain robust if we control for CEX trade flow and size of DEX trades, meaning that the priority fee attached to DEX trades reveals private information beyond what is captured by these two other confounding factors.

Why do informed traders bid high fees to execute their orders on DEXs? Are they worried about not getting executed because of blockchain congestion? Such a motive naturally arises as informed traders have to compete with other users for limited block space. As the blockchain becomes congested and the marginal priority fee increases, informed traders increase their bids in order to execute their orders quickly. We show that this is not the only channel. Our analysis shows that DEX trades with an excessively high fee (relative to other transactions in the same block) are privately informed: excluding those trades significantly reduces the permanent price impact of the high-fee DEX trade flow to a range between 2.83 and 5.36 basis points. This result suggests that informed traders on DEXs may bid excessively high fees for their trades, which is less likely to be driven merely by their desire of avoiding execution risk, but more likely to result from them competing with each other.

How do informed traders compete on DEXs? Our analysis of tick-by-tick mempool order data shows that only a small proportion of DEX trades with excessively high fees, varying between 12.05% and 26.71% depending on the specific cryptocurrency pair, are likely to result from priority gas auctions (PGAs), a form of competition where traders competitively bid up fees to win the execution slot (Daian et al., 2020). Instead, informed traders start by placing a high initial fee bid to discourage other traders from bidding up. This bidding pattern is also referred to as *jump bidding*, and has been well documented in auction theory (Daniel and Hirshleifer, 1998; Avery, 1998). It can be rationalized by a high bidding cost on DEXs and the winner's curse problem. The finding shows that while bidding a high priority fee can leak information of informed traders, it can also serve as a signaling tool to deter competition from other traders.

The paper proceeds as follows. We review the literature in Section 2. In Section 3, we introduce institutional details of DEXs and their unique characteristics. In Section 4, we describe our dataset. We present the empirical methodology in Section 5, and discuss the results in Section 6. We conclude in Section 7.

2 Literature Review

Our paper relates to several streams of literature. Past studies in market microstructure have linked the private information contained in trades to their public characteristics[2], e.g., block trades versus non-block trades (Easley and O'Hara, 1987), odd-lot trades versus round-lot trades (O'Hara, Yao, and Ye, 2014), trades executed on ECNs versus the NASDAQ exchange (Barclay, Hendershott, and McCormick, 2003). We contribute to this literature by studying the private information content of

[2] Brogaard, Hendershott, and Riordan (2014) use proprietary data to investigate the information content of private trade characteristics e.g., HFT trades versus non-HFT trades.

the priority fee, a unique characteristic of DEX trades. We show that the priority fee of a trade conveys its private information and serves as a new public signal for trade informativeness.

Another related branch of literature analyzes trading strategies of informed traders and their implications on price discovery under different market and information structures. Facing no competition, a monopolistic informed trader reveals her information linearly in time through trades (Kyle, 1985). In contrast, price discovery is fast when there is competition among multiple privately informed traders (Holden and Subrahmanyam, 1992; Foster and Viswanathan, 1996; Back, Cao, and Willard, 2000) or impatience of informed traders due to uncertain timing of the public announcement of private information (Caldentey and Stacchetti, 2010) or short information horizon (Kaniel and Liu, 2006). In contrast, price discovery is slow when informed traders can time their trades (Collin-Dufresne and Fos, 2015). We contribute to this literature by analyzing informed trading patterns on DEXs. We show that price discovery realizes through informed traders bidding high fees to prioritize the execution of their orders and compete with other informed traders. Further, informed traders compete by following the jump bidding strategy, which is not possible on traditional CEXs because they do not have priority fee bidding.

Last, we contribute to the nascent yet rapidly growing literature on decentralized exchanges, and the role of priority fees in the provision of trading and liquidity incentives. Park (2021) focuses on the unintended consequence of public blockchain order processing, which exposes all pending DEX transactions to the risk of a "sandwich attack". He argues that, in theory, liquidity demanders are able to prevent frontrunning by choosing a very high priority fee. Capponi and Jia (2021) investigate the effect of DEX pricing rules on welfare and liquidity provision incentives. They show that arbitrageurs can always outbid liquidity providers in priority fee auctions, to exploit the price discrepancy between CEXs and DEXs, which in turn reduces incentives for liquidity provision. Barbon and Ranaldo (2022) compare the market quality of CEXs and DEXs. They find that the transaction costs of CEXs and DEXs are similar but DEX prices are less efficient partially because of priority fees. Lehar and Parlour (2021) contrast DEXs running an AMM with CEXs running a LOB and focus on the different trade-offs faced by liquidity providers. Aoyagi and Ito (2021) model the coexistence of an AMM-based DEX and a LOB-based CEX and study the resulting equilibrium in liquidity provision. Lehar, Parlour, and Zoican (2022) show that the priority fee drives the fragmentation of liquidity supply across DEX liquidity pools. Hasbrouck, Rivera, and Saleh (2022) demonstrate that increasing DEX trading fees can increase DEX trading volume. The main contribution of our work relative to this literature is to highlight the role of priority fees in the bidding strategies of informed traders and thus the consequent impact on the price discovery process on DEXs.

3 Institutional background of DEXs

DEXs are a type of exchange that operates on a decentralized blockchain network. Trades are executed through automated smart contracts which allow for peer-to-peer trading without the need for centralized intermediaries. To execute on a DEX, a trader must broadcast the order in the peer-to-peer network of the blockchain on which the DEX is deployed, and bid a priority fee. Once the order is received by validators, it becomes a pending order in their mempools. At discrete times, one validator is chosen to append the next block to the chain. As the block space is limited, the validator will execute orders in her mempool in descending order of priority fees. Note that DEX orders compete for blockspace with order flow unrelated to DeFi transactions (e.g. cryptocurrency payments and initial coin offerings) and pending in the mempool of the same validator.

It is worth highlighting several notable differences between DEXs and CEXs to best appreciate trading mechanisms on DEXs. First, unlike CEXs which typically process orders continuously, DEXs execute orders in discrete times depending on the underlying blockchain infrastructure.[3] Continuous trading on CEXs put fast traders at an advantage as the execution priority of orders depends on their arrival time. As a result, it leads to an arms race between high-frequency traders and can be detrimental to market liquidity (Budish, Cramton, and Shim, 2015; Aquilina, Budish,

[3]On the proof-of-work (PoW) Ethereum blockchain, the block time is random and a block is validated on average every 13 seconds. On September 6, 2022, Ethereum transitioned to proof-of-stake (PoS) and fixed the block time to 12 seconds.

and O'Neill, 2021). On DEXs, the arms race on trading speed is much less likely to occur because orders are executed at discrete times.

Second, traders on DEXs have to bid a priority fee to determine the execution priority of their orders. If traders are willing to pay a higher fee, it is more likely their orders will be executed quickly. Thus, the execution mechanism on DEXs is different from the frequent batch auction design proposed by Budish, Cramton, and Shim (2015) to mitigate the arms race between fast traders. Although orders are executed in batches for both mechanisms, there is no execution priority among the orders in a frequent batch auction as orders are crossed at a uniform price. In contrast, orders on DEXs are executed in descending order of their fees, and thus trades with higher fees are more likely to be executed earlier and receive better prices. Hence, instead of competing on speed on CEXs, traders on DEXs compete on priority fees.

Third, CEXs typically operate as a LOB market. In a LOB market, limit orders submitted by market makers are aggregated and incoming market orders are executed against the resting limit orders in the book. As a result, LOB prices can adjust as market makers revise their quotes responding to new information such as public news and the process involves no trade execution. For instance, Brogaard, Hendershott, and Riordan (2019) show that price discovery occurs predominantly through limit orders in the Canadian stock market. In comparison, most DEXs use an automated market maker (AMM) model. Unlike LOB markets, liquidity providers do not submit price quotes on AMMs. Instead, they provide liquidity by depositing tokens in the pool and incoming orders are executed against a pre-determined pricing curve such as a constant product function. As a result, AMMs prices can only adjust through trade executions. We refer to Capponi and Jia (2021) for additional details about liquidity provision and pricing curve characteristics on DEXs. [4]

[4] Another significant difference between CEXs and DEXs is that while the former typically requires users to transfer assets to their custody before trading, the latter allows traders to keep custody of their assets. Thus traders on DEXs are less subject to hacking risks.

4 Data

In this section, we provide details on the dataset used in our empirical analysis. DEXs and other decentralized finance (DeFi) services are nowadays hosted by many blockchains, including Ethereum, Binance Smart Chain (BSC), Tron, Arbitrum, Polygon, Avalanche, and Optimism. Among those, Ethereum has, by far, the dominant market share exceeding 50% in terms of total value locked (TVL).[5] Because of these estimates, we consider the Ethereum blockchain in our study. In addition, about 10% of cryptocurrency spot trading occurred on DEXs. Uniswap is currently the largest DEX by trading volume and accounts for more than half of the total DEX trading volume. The remaining 90% of crypto spot trading is executed on CEXs. The largest CEX by daily trading volume is Binance, which accounts for more than 60% of the CEX market share.[6] Given the dominant market shares of Uniswap and Binance, we focus on these two exchanges in our empirical market microstructure analysis. We describe executed trade data in Section 4.1, and mempool order data in Section 4.2.

4.1 Executed trade data

Our dataset covers trades executed on Binance, the largest CEX, and on Uniswap, the largest DEX, for six the most traded token pairs during the period from November 18, 2020, through February 10, 2021. Note that we begin our sample period after Uniswap's staking reward program was terminated to avoid including a structural break in token liquidity as the termination resulted in large token outflows and smaller pool sizes. In addition, our sample period ends before the first block to include Flashbots trades was mined.[7] Our six token pairs are USDC-USDT, DAI-USDT, ETH-USDT, WBTC-ETH, LINK-ETH, and AAVE-ETH, and they can be categorized into two types: "Stable" and "NonStable". "Stable" pairs include two stablecoins pegged to the US Dollar

[5]See https://defillama.com/chains for details.

[6]We refer to https://www.theblock.co/data/decentralized-finance/dex-non-custodial for trading volume estimates on DEXs, and to https://www.coingecko.com/en/exchanges for trading volume estimates of the top crypto exchanges.

[7]Flashbots provide off-chain private channels through which traders can send their orders directly to miners and thus avoid the risk of being front-run in the mempool.

(USDC-USDT, ETH-USDT). “NonStable” pairs include at least one non-stable token, i.e., which is not pegged to any fiat currency (ETH-USDT, WBTC-ETH, LINK-ETH, and AAVE-ETH). Binance trades are publicly available and collected from the Binance website[8], while Uniswap trades are collected through a proprietary node. Below, we provide a detailed description of each trade dataset.

Uniswap trades. Each Uniswap trade contains the following information types:

- Timestamp: the timestamp of the block in which the trade is included (to the precision of a second), the number of the block in which the trade is included, and the execution position of the trade in that block.
- Identifiers: hash, submission address, and nonce.
  - Hash: the hash is a unique identifier for each new order submitted to the network. Using the hash, we can match an executed trade with its original order. Note that when a trader modifies a pending order, the modified order will be assigned a new hash.
  - Submission address and nonce: nonce is used to track the orders sent from a given submission address. Specifically, the first order of a trader is assigned nonce “0”, her second order has a nonce “1”, and her N^{th} order has a nonce “N”. Importantly, a new order will not be executed if there are pending orders with smaller nonces from the same submission address. It means that if a trader wants to modify or cancel her pending order, she has to broadcast a new order with the same nonce and a higher priority fee. A validator will only execute the new order as she prioritizes orders with higher fees.[9] Through the submission address and the nonce number, we can link an executed or canceled order with its history of revisions.

[8] We refer to https://data.binance.vision/?prefix=data/spot/monthly/ for details.

[9] A trader can cancel a pending order by submitting a new order with the same nonce but a higher priority fee, where she transfers zero amount of the native token of the blockchain (e.g., ETH on the Ethereum blockchain) to her own wallet. The old order is effectively “canceled” as only the new order with a higher priority fee will be executed by validators.

- Trade characteristics: gas price, gas used, trade direction, and the amount of tokens that the trader deposits in and takes out from the liquidity pool.

  – Gas price and gas used: on the Ethereum blockchain, the priority fees are referred to as "gas fees".[10] The "gas used" of a transaction measures the fixed amount of computational resources needed for its execution. More complicated transactions require more computational work and thus consume a higher amount of gas. Upon bidding, Ethereum users choose the "gas price", that is, the unit price of gas they are willing to pay. Hence, the total gas fee paid by users is equal to the gas used multiplied by the gas price. Note that Ethereum validators sort and execute transactions in mempools in decreasing order of gas price.

  – Trade direction: it indicates whether it is a buy trade or sell trade in terms of the base token. We follow the convention used for currency pairs in the foreign exchange market and label the first token appearing in a pair as the base token and the second token as the quote token. For example, the token pair ETH-USDT has ETH as the base token and USDT as the quote token.

  – The amount of tokens that the trader deposits in and takes out from the liquidity pool: we use the amount of the base token swapped as the transaction size of the trade.

Binance trades. Each Binance trade record includes a unique identifier for the trade, the timestamp (to the precision of a millisecond), the transaction price, the transaction size in terms of the base token, and an indicator for whether the buyer uses a limit order or a market order, which tells us the direction of the trade: if the buyer uses a market order, then it is classified as a buy trade; otherwise, it is a sell trade.

[10]The London Upgrade of the Ethereum blockchain on August 5, 2021, implemented the EIP-1559, and decomposed the gas fee into two parts: base fee and priority fee (tips). The base fee is a reserve price every trader needs to pay, and adjusts to the congestion level of the network. The base fee gets burnt and thus is not earned by validators. The priority fee or tip instead is bid by traders to incentivize validators to include their transactions in the next block. Although our sample period precedes EIP-1559, we use the terminology priority fee to stress its role in determining the execution priority of pending orders.

In addition to executed trades, we obtain event updates of Binance's limit order book (to the precision of a second). With order book event updates, we are able to reconstruct the order book states and calculate the best bid, best ask, and the midquote on Binance, which we use to calculate token pair returns.

4.2 Mempool order data

We obtain tick-by-tick Ethereum mempool order data from Amberdata[11]. Our mempool data covers the same sample period of November 18, 2020, through February 10, 2021. The dataset includes every new order submission received in the mempool of nodes maintained by Amberdata, which either ends up being executed or left unexecuted. Each order comes with the following information: the hash, the timestamp when the order is received by the node (to the precision of a millisecond), the address of the trader, nonce, gas price, and gas limit (i.e. the maximum gas allowed to be used). With the mempool data, we can track the complete history of order revisions, if they occur, before the final order is executed and recorded as a trade. Hence, we are able to observe whether the trader increases the gas price attached to her order to get it executed.

4.3 Summary statistics of executed trades

In Table 1, we provide an overview of trading characteristics for our sample token pairs. We report summary statistics of their daily trading volume and their daily number of trades on Uniswap and Binance. Several observations are in order. First, trading in all six token pairs is fairly active. For instance, the average daily number of trades (daily trading volume) on Uniswap is 997 (≈ 3.4 million USDT), 8,560 (73,489 ETH ≈ 66 million USD) and 1,371 (31,644 ETH ≈ 28 million USD) for USDC-USDT, ETH-USDT and WBTC-ETH respectively.

Second, trading activity on Uniswap and Binance differs significantly across token pairs. For the two Stable token pairs, USDC-USDT and DAI-USDT, trading is much more active on Binance

[11] Amberdata is a US data company specializing in market data in decentralized finance.

Table 1. Summary statistics of daily trading statistics on Uniswap and Binance. This table reports, for each token pair, summary statistics of daily trading volume (TradingVolume) and number of trades (TradeCount) on Uniswap and Binance respectively. N refers to the number of days in our sample period.

(a) Stable token pairs. Trading volume is denominated in thousand USDT.

Pair		N	Mean	SD	Min	Med	Max
USDC-USDT	TradingVolume-Uniswap	85	3426	1747	646	3456	7577
	TradeCount-Uniswap	85	997	397	504	884	3085
	TradingVolume-Binance	85	96681	59806	22936	82593	276362
	TradeCount-Binance	85	51403	21583	15724	47647	116379
DAI-USDT	TradingVolume-Uniswap	85	1494	1361	56	1155	5830
	TradeCount-Uniswap	85	658	403	174	570	2068
	TradingVolume-Binance	85	11575	10451	2224	9210	77831
	TradeCount-Binance	85	9174	7925	1525	7341	58558

(b) NonStable token pairs. Trading volume is denominated in ETH.

Pair		N	Mean	SD	Min	Med	Max
ETH-USDT	TradingVolume-Uniswap	85	73489	37752	36923	62131	263356
	TradeCount-Uniswap	85	8560	1700	6311	8155	16419
	TradingVolume-Binance	85	1444426	709203	493012	1281734	4245010
	TradeCount-Binance	85	994231	524099	272746	915584	2577496
WBTC-ETH	TradingVolume-Uniswap	85	31644	17748	9014	27141	87965
	TradeCount-Uniswap	85	1371	592	646	1127	3338
	TradingVolume-Binance	85	2023	1993	135	1258	9984
	TradeCount-Binance	85	7886	7529	289	5332	35191
LINK-ETH	TradingVolume-Uniswap	85	10779	6295	3437	9406	42520
	TradeCount-Uniswap	85	1054	380	574	961	2682
	TradingVolume-Binance	85	4387	2687	1071	3856	13598
	TradeCount-Binance	85	10459	6793	2223	9391	29514
AAVE-ETH	TradingVolume-Uniswap	85	7368	4177	1766	6366	29936
	TradeCount-Uniswap	85	609	253	261	551	1514
	TradingVolume-Binance	85	2135	1510	408	1627	10143
	TradeCount-Binance	85	6829	5410	1131	5511	36964

than on Uniswap. For example, the average daily trading volume on Binance is about 96 million USDT for USDC-USDT, more than an order of magnitude larger than that on Uniswap. This is because trading is cheaper on Binance than on Uniswap, as the latter imposes a larger price impact due to the convexity of the bonding curve, and requires to pay an additional priority fee. The transaction cost is an important factor when trading Stable token pairs, as Stable coin transactions are not information but liquidity driven. In contrast, for NonStable token pairs, trading is in general more active on Uniswap than Binance. Take, for example, WBTC-ETH. Its average daily trading volume is about 31644 ETH on Uniswap, much larger than 2023 ETH on Binance.

In Table 2, we further report summary statistics of the execution price, gas price and trade size of Uniswap trades for our six sample token pairs. The average trade size of a Uniswap trade is fairly large and amounts to about 4,360 USDT ($\approx$ 4,360 USD), 8.59 ETH ($\approx$ 7,661 USD), and 23.07 ETH ($\approx$ 20,577 USD) for USDC-USDT, ETH-USDT, and WBTC-ETH respectively. Second, the gas price attached to Uniswap trades varies considerably across trades. Take WBTC-ETH as an example. While a Uniswap trade in WBTC-ETH has an average gas price of 126.13 Gwei (1Gwei = 10^{-9}ETH), its standard deviation is 220.12, which is about twice the size of the mean. Such a large variation can result from either change in the overall congestion of the Ethereum network, or from traders' bidding high fees to trade on the information.

5 Empirical Methodology

The de facto standard approach to estimating the private information contained in trades has been proposed by Hasbrouck (1991a) and Hasbrouck (1991b). The security return and trades are modeled as a structural vector-autoregressive (VAR) system that characterizes the dynamic interactions between them. With the structural VAR model, one can estimate the persistent impact of trades on security price, a proxy for private information, by computing the cumulative impulse responses of security return to trade innovations over a substantially long period. In addition, one is able to decompose the total efficient price variance, a measure of the total amount of information, into a component correlated with trades and a component uncorrelated with trades. The former component reflects the amount of private information conveyed through trades while the latter captures the amount of public information such as news.

Hasbrouck (1991a) and Hasbrouck (1991b) consider security return and a single aggregated

Table 2. Summary statistics of Uniswap trades. This table reports, for each token pair, summary statistics of the transaction price (TxPrice), transaction size (TxSize), and gas price (GasPrice). Gas price is denominated in Gwei, which equals to 10^{-9} ETH. N refers to the number of trades for each token pair during our sample period.

(a) Stable token pairs. Transaction size is denominated in thousand USDT.

TokenPair	Variable	N	Mean	SD	1%	10%	Median	90%	99%
USDC-USDT	TxPrice	84779	1.00	0.00	0.99	1.00	1.00	1.00	1.01
	GasPrice	84779	90.91	81.52	16.00	33.00	71.00	164.00	400.00
	TxSize	84779	3.43	8.51	0.01	0.11	1.04	7.66	40.05
DAI-USDT	TxPrice	55919	1.00	0.00	0.99	1.00	1.00	1.01	1.01
	GasPrice	55919	91.08	121.44	18.00	35.00	73.00	155.00	397.82
	TxSize	55919	2.27	5.15	0.01	0.08	0.76	5.05	26.02

(b) NonStable token pairs. Transaction size is denominated in ETH.

TokenPair	Variable	N	Mean	SD	1%	10%	Median	90%	99%
ETH-USDT	TxPrice	727600	891.94	379.26	474.95	546.93	653.14	1397.87	1751.67
	GasPrice	727600	99.77	189.92	15.30	30.00	70.00	181.00	530.00
	TxSize	727600	8.59	34.08	0.01	0.13	1.37	15.34	124.89
WBTC-ETH	TxPrice	116520	30.48	5.44	22.56	23.88	31.44	38.28	42.33
	GasPrice	116520	126.13	220.12	17.00	37.00	88.00	240.00	652.00
	TxSize	116520	23.07	59.40	0.02	0.23	3.99	64.76	235.89
LINK-ETH	TxPrice	89630	0.02	0.00	0.01	0.01	0.02	0.02	0.03
	GasPrice	89630	114.48	242.18	16.00	34.00	78.89	205.70	669.82
	TxSize	89630	10.22	24.36	0.02	0.19	2.82	27.20	86.39
AAVE-ETH	TxPrice	51811	0.16	0.05	0.09	0.11	0.15	0.27	0.31
	GasPrice	51811	110.91	177.61	15.56	30.72	80.00	203.00	565.62
	TxSize	51811	12.08	19.93	0.03	0.20	4.85	29.95	88.13

signed trade flow. While different functional forms of the aggregated signed trade flow, such as the trade direction indicator and the squared trade flow, can be included in the structural VAR, all trade variables essentially originate from the same aggregated signed trade flow. Recent studies extend the approach and include more than one trade flow so that one can compare the informativeness of different trade flows, for example, ECN trades versus NASDAQ trades (Barclay, Hendershott, and McCormick, 2003), odd-lot trades versus round-lot trades (O'Hara, Yao, and Ye, 2014), and orders versus trades (Fleming, Mizrach, and Nguyen, 2018; Brogaard, Hendershott, and Riordan, 2019).

To examine whether DEX trades with a higher priority fee convey more private information, we follow the literature and estimate a structural VAR model with CEX return and DEX trade flows of different fee levels. Below, we first introduce our fee-level classification method. Then we detail our structural VAR specification.

5.1 Construction of DEX trade flows with different fee levels

We adopt a quantile-based, rolling-window approach to construct DEX trade flows with three different priority fee levels. Specifically, to classify trades in the current block t, we first sort them together with all trades within the last 20 non-empty blocks, i.e., block $t-20$ to block $t-1$ based on their priority fee in descending order. Then we label trades in block t located in the top quartile (i.e., above 75% quantile) as high-fee trades; trades located in the bottom quartile (i.e., below 25% quantile) as low-fee trades; all other trades as mid-fee trades (i.e., between 25% and 75% quantile). Last, we compute the (signed) trade flow at a given fee level i in a block t as:

$$x_t^i = \sum_k d_{t,k}^i s_{t,k}^i. \tag{1}$$

where k indexes trades with fee level i in block t, d_k is the trade direction indicator (+1 for buyer-initiated trades and -1 for seller-initiated trades), and s_k is the trade size. So, for each block, we construct three DEX trade flows: $x_t^{\text{LowFee-DEX}}$, $x_t^{\text{MidFee-DEX}}$, and $x_t^{\text{HighFee-DEX}}$. If we have no observation for one type of trade in a block, then the corresponding trade flow simply takes a zero value.

It is worth highlighting two choices made in the above classification method. First, we choose three different fee levels, high-fee (above 75% quantile), mid-fee (between 25% and 75% quantile), and low fee (below 25% quantile), instead of two, high-fee (above 50%) and low-fee (above 50%), to guarantee a clear distinction between the high-fee and low-fee group.

Second, we set the length of the rolling window to 20 non-empty blocks to balance the following two competing forces. On the one hand, a too-short window makes our quantile estimates noisy due to a small number of trades. For example, if we only use the current block and if that contains only a few trades, then two trades with very similar gas prices will fall into different categories. On the other hand, a too-long window might include trades with priority fees bid too long ago to reflect the current congestion level of the blockchain. Admittedly, the choice of 20 non-empty blocks remains arbitrary. As a robustness check, we repeat the fee level classification with window lengths of 10 and 40 blocks, and the structural VAR results stay qualitatively the same. We report the detailed results in Appendix A.2.

5.2 Structural VAR specification

A general structural VAR model can be specified as follows:

$$Ay_t = \alpha + \Phi_1 y_{t-1} + \cdots + \Phi_p y_{t-p} + \varepsilon_t \tag{2}$$

where $\Phi_1 \ldots \Phi_p$ are system matrices of the lagged terms of the structural VAR model. ε_t is the vector of structural innovations and satisfies the following conditions: $\mathrm{E}(\varepsilon_t) = 0$; $\mathrm{E}(\varepsilon_t \varepsilon_t') = \Sigma_\varepsilon$; $\mathrm{E}(\varepsilon_t \varepsilon_s') = 0$ for $s \neq t$. Note that as the contemporaneous relations between the endogenous variables are directly modeled in A, the covariance matrix Σ_ε is diagonal. y_t is the endogenous variable vector, and A is the structural matrix capturing the contemporaneous correlations between the endogenous variables.

Choice of the endogenous variable vector For our baseline specification, we include the midquote return on Binance and three trade flows on Uniswap with different priority fee levels in the endogenous variable vector y_t, that is,

$$y_t = \begin{pmatrix} r_t^{\text{CEX}} & x_t^{\text{LowFee-DEX}} & x_t^{\text{MidFee-DEX}} & x_t^{\text{HighFee-DEX}} \end{pmatrix}'. \tag{3}$$

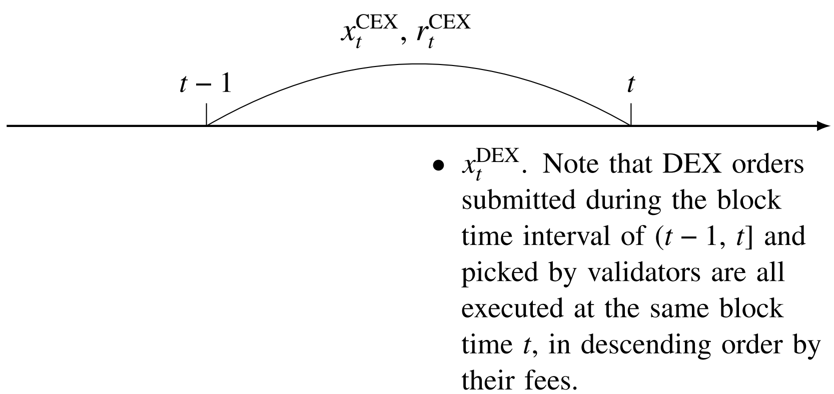
where t indexes block time, and r_t^{CEX} is the Binance midquote return from block time $t-1$ to t. $x_t^{\text{LowFee-DEX}}$, $x_t^{\text{MidFee-DEX}}$ and $x_t^{\text{HighFee-DEX}}$ denote Uniswap trade flows in block t, respectively with low, mid and high priority fee levels, as specified in the above priority fee level classification.

We use the midquote return on Binance, instead of Uniswap, in the endogenous variable vector to make possible a dichotomy that identifies public information with the return innovation and private information with the trade flow innovations. As discussed in Hasbrouck (1991a) (p. 190), the dichotomy is most likely to fail when "there are market features which impair the quote revision process and thereby constrain the quote revisions from fully reflecting public information". On Uniswap, which runs AMMs, market makers provide liquidity by depositing tokens in the pools instead of submitting limit orders. Thus, it is impossible for them to revise their quotes in response to public information. Instead, the spot price on AMMs can only be adjusted by the execution of incoming trades following a pre-determined pricing curve. The consequence is that both private information and public information will be wrongly attributed to trades and thus overestimate their information. In contrast, Binance runs a traditional LOB where market makers can adjust their quotes without the need for trades, fitting the dichotomy.

As we include the return on Binance, which executes orders based on a continuous clock, and trade flows on Uniswap, which executes orders based on a discrete clock, in the same structural VAR model, we need a timestamp convention that encompasses both clocks. To do so, we define r_t^{CEX} as the log difference between the Binance midquote at block time $t-1$ and t respectively. All three Uniswap trade flows, $x_t^{\text{LowFee-DEX}}$, $x_t^{\text{MidFee-DEX}}$ and $x_t^{\text{HighFee-DEX}}$, are computed based on trades executed at block time t as in Equation 1. In contrast, trade flow on Binance, x_t^{CEX}, which we use in the later robustness check section, is computed based on trades executed on Binance between the block time $t-1$ and t. In Figure 1, we provide a visual illustration of the chosen convention.

Because r_t^{CEX} is defined over block time $t-1$ and t and Uniswap trade flows, $x_t^{\text{LowFee-DEX}}$, $x_t^{\text{MidFee-DEX}}$ and $x_t^{\text{HighFee-DEX}}$, are defined at block time t, it might appear at first glance that the Binance return happens before the Uniswap trade flows. In actuality, these two can affect each other contemporaneously. Although Uniswap trades are all executed at the same block time t, market makers on Binance can monitor the pending orders in mempool and have a good guess of which orders will be executed according to their priority fees. Then they can adjust their quotes accordingly on Binance based on the mempool information. For example, if there are many pending buy orders in the mempool, market makers might revise up their quotes on Binance. We detail the strategy for dealing with this contemporaneous causality issue below.

Figure 1. Timestamp convention. This figure illustrates our time convention. t is block time. r_t^{CEX} is the log return from Binance defined over the time interval between $t-1$ and t. Note that we do not have quote updates from Binance. The return is calculated based on trade prices, not midquotes. x_t^{CEX} is the trade flow on Binance obtained by summing the trades executed between block time $t-1$ and t (See Equation 1). x_t^{DEX} is the signed trade flow on Uniswap at block time t.



Permanent price impact and information shares of trade flows After estimating the structural VAR model, we can easily obtain the vector moving average (VMA) representation with structural innovations as below:

$$y_t = \Theta(L)\varepsilon_t = \Theta_0\varepsilon_t + \Theta_1\varepsilon_{t-1} + \Theta_2\varepsilon_{t-2} + \cdots \tag{4}$$

where $\Theta(L)$ is the polynomial of the lag operator $\Theta(L) = \Theta_0 + \Theta_1 L + \Theta_2 L^2 + \cdots$ and $\Theta_0, \Theta_1, \ldots$ are the VMA system matrices. Then the permanent price impact (PPI) of a trade flow variable k is defined as the cumulative impulse response of the midquote return to a unit shock in the trade flow, that is,

$$\text{PPI}_k = \frac{\sum_{j=0}^{\infty} \partial r_{t+j}^{\text{CEX}}}{\partial \varepsilon_{k,t}} = [\Theta(1)]_{1,k}, \quad k > 1 \tag{5}$$

where $[\Theta(1)]_{1,k}$ denotes the $(1, k)$-*th* entry of $\Theta(1) = \Theta_0 + \Theta_1 + \Theta_2 + \ldots$.

One challenge of measuring the private information content of trades through price impact measures is that trades can move prices through two confounding effects: inventory control effects and asymmetric information effects. For example, upon executing buy trades, market makers can revise their quotes upwards either because they would like to induce future sell trades to revert their inventory back to the original level or because they learn positive private information from the buy trades. However, as pointed out in Hasbrouck (1991a), we can partially resolve these two effects if we measure the impact of trades on the security prices over a substantially long period. The idea is simple. Inventory control effects are inherently transitory: market makers revise their quotes back to the pre-trade level as their inventories revert back). In contrast, private information conveyed through trades due to asymmetric information is permanently incorporated into the security price.

In addition to permanent price impacts, we compute another trade informativeness measure, the information shares of the trade flow variables (Hasbrouck, 1991b). The basic idea is that we can decompose the logarithm of the midquote q_t into an efficient price component m_t and a microstructure noise term s_t:

$$q_t = m_t + s_t. \tag{6}$$

The efficient price component m_t is a random walk with innovation w_t: $m_t = m_{t-1} + w_t$ where $Ew = 0$, $Ew_t^2 = \sigma_w$, and $Ew_t w_\tau = 0$ for $\tau \neq t$. The microstructure noise term s_t is a zero-mean process that is jointly covariance stationary with w_t. Thus, the variance of the innovation σ_w has a natural interpretation as the total new information incorporated into the efficient price, which can be attributed to both public news and private information conveyed through trades.

As the above structural random walk decomposition is unobservable, to estimate the efficient price variance and decompose it to a trade-correlated, private information component and a trade-uncorrelated, public information, component, we need to resort to its reduced-form representation. First, we can rewrite the VMA representation above (see Equation 4) as follows:

$$\begin{pmatrix} r_t^{\text{CEX}} \\ \boldsymbol{x_t} \end{pmatrix} = \begin{pmatrix} \Theta^a(L) & \boldsymbol{\Theta^b(L)} \\ \boldsymbol{\Theta^c(L)} & \boldsymbol{\Theta^d(L)} \end{pmatrix} \begin{pmatrix} \varepsilon_{r,t} \\ \boldsymbol{\varepsilon_{x,t}} \end{pmatrix} \tag{7}$$

where $\boldsymbol{x_t}$[12] is the vector containing the Uniswap DEX trade flows. $\Theta^a(L)$ and $\boldsymbol{\Theta^b(L)}$, $\Theta^c(L)$ and

[12]We use normal font for scalar variables and bold font for vectors/matrices.

$\boldsymbol{\Theta^d}(L)$ are the polynomial of the lag operators with Θ^a, $\boldsymbol{\Theta^b}$, $\boldsymbol{\Theta^c}$ and $\boldsymbol{\Theta^d}$ being the conformable VMA system matrices in the return and trade equations. $\text{Var}(\varepsilon_{r,t}) = \sigma^2_{\varepsilon_r}$ and $\text{Var}(\boldsymbol{\varepsilon_{x,t}}) = \Sigma_{\varepsilon_x}$.

Skipping detailed proofs (See Hasbrouck, 1991b), we have:

$$\sigma^2_w = \boldsymbol{\Theta^b}(\mathbf{1})\Sigma_{\varepsilon_x}\boldsymbol{\Theta^b}(\mathbf{1})' + [\Theta^a(1)]^2\sigma^2_{\varepsilon_r}. \tag{8}$$

where the first component represents the private information conveyed through trades and the second component represents the public information through the news.

For a given A matrix, the structural innovations $\boldsymbol{\varepsilon_t}$ have zero contemporaneous correlations by construction, which implies that the covariance matrix Σ_{ε_x} is diagonal. Thus, we can further attribute the private information component to the contribution of each trade flow uniquely: $\boldsymbol{\Theta^b}(\mathbf{1})\Sigma_{\varepsilon_x}\boldsymbol{\Theta^b}(\mathbf{1})' = \sum_k[\Theta^b_k(1)]^2\sigma^2_{\varepsilon_k}$ where k indexes trade flow. $[\Theta^b_k(1)]^2$ is the k-th element of the vector of $\boldsymbol{\Theta^b}(\mathbf{1})$ and $\sigma^2_{\varepsilon_k}$ is the k-th element of the diagonal of Σ_{ε_x}. Finally, to normalize each variable's information contribution, an absolute measure, to its information share, a relative measure bounded between 0 and 1, we divide it by the total efficient price variance. Formally, the information share (IS) of the trade flow variable k is computed as:

$$\text{IS}_k = \frac{[\Theta^b_k(1)]^2\sigma^2_{\varepsilon_k}}{\sigma^2_w} \tag{9}$$

Compared with permanent price impact, the information share measure is a more comprehensive measure of trade informativeness as it weighs the permanent price impact of the trade flow variable by its own structural innovation variance. So, if two trade flow variables have the same permanent price impact, the one with a larger innovation (unexpected) variance will have a larger information share.

Resolution of the contemporaneous correlations To pin down the permanent price impacts (PPI) and information shares (IS) of the Uniswap trade flows, one needs to specify the A matrix in the structural VAR model as it enters into the reduced-form system matrices (see Equation 2), which dictates the contemporaneous causality between the endogenous variables. For example, at the same block time t, is it the Binance return that causes changes in the Uniswap trade flows or the other way around? In addition, is it the high-fee Uniswap trade flow that causes changes in the low-fee trade flow or the other way around? The specific contemporaneous assumption affects the computation of the two trade informativeness measures.

In some earlier applications of the structural VAR approach (see, e.g., Hasbrouck, 1991a; Hasbrouck, 1991b; Fleming, Mizrach, and Nguyen, 2018; Brogaard, Hendershott, and Riordan, 2019), trades are assumed to contemporaneously cause quote revisions, not vice versa. This timing convention is based on the following sequence of quote and trade interactions often assumed in theoretical models of dealers' market (see, e.g., Glosten and Milgrom, 1985): market makers leave their quotes at $t-1$; trades and possibly public news arrive at t; after observing the trades and public news, market makers revise their quotes, incorporating information from both sources.

Adopting the same timing convention is problematic in our application. First, our estimation frequency is block-by-block, which on average is about 12 seconds. Given such a relatively long interval, the contemporaneous correlations between the Binance return and Uniswap trade flows can be quite large. Second, even though we could impose a specific contemporaneous causality between the Binance return and Uniswap trade flows, it is not straightforward to impose one among the Uniswap trade flows themselves. Although Uniswap trades are executed in descending order of priority fees, meaning the high-fee trade flow is executed before the mid-fee and low-fee trade flows, all are executed at the same block time t. For example, low-fee trades might be submitted before the high-fee trades and thus enter the mempool earlier, they can be executed at the same block time t.

Given the above two concerns, we do not impose a specific contemporaneous causality among the endogenous variables ex-ante. Instead, we try all possible contemporaneous causalities based on a generic recursive structure to obtain the lower and upper bound of the permanent price impacts and information shares of the trade flows (Hasbrouck, 1995; Barclay, Hendershott, and McCormick, 2003; O'Hara, Yao, and Ye, 2014). In econometric terms, we use Cholesky factorization to decompose the reduced-form covariance matrix, which is equivalent to imposing a recursive structure on the A matrix. Without loss of generality, we let the A matrix be a lower-triangular matrix. Thus we assume that the first variable in the endogenous variable vector, y_t, contemporaneously causes the second variable onwards and the second variable contemporaneously causes the third variable onwards, and so on. For example, if we place the high-fee Uniswap trade flow as the first variable, we assume that it contemporaneously causes both the Binance return and the other two Uniswap trade flows. Thus, we are likely to obtain an upper bound for the permanent price impact and information share of the high-fee Uniswap trade flow. In contrast, if we place the high-fee Uniswap trade flow as the last variable, we assume that both the Binance return and the other two Uniswap trade flows contemporaneously cause the high-fee Uniswap trade flow. Thus, we are likely to obtain a lower bound for the permanent price impact and information share of the high-fee Uniswap trade flow.

So, we consider all 24 (4!) possible sequences of the endogenous variable vector to compute the upper and lower band for each endogenous variable. In addition, We implement the structural VAR estimation in the following ways: (1) the model is estimated at block-by-block frequency, although the priority fee level classification is based on a 20-block rolling window; (2) we set the number of lags in the structural VAR model to 5. In Appendix A.3, we change the number of lags included in the structural VAR model to 10 and 20, and show that estimation results remain qualitatively the same; (3) As the base currency varies across token pairs, to ease comparison and aggregation across token pairs, we standardize all trade flow variables such that they have zero mean and unit variance. Hence, the impulse responses reported below should be interpreted as permanent price impacts in basis points per standard deviation increase in the trade flow. In appendix A.1, we use the unstandardized DEX trade flows and the results stay qualitatively the same.

6 Empirical results

In what follows, we present the main results from our empirical analysis. First, in Section 6.1 we report results from our structural VAR analysis and show our key finding: high-fee DEX trade flow contains more private information than low-fee DEX trade flow. Then, in Section 6.2, we conduct several robustness checks for our key finding. Last, in Section 6.3, we provide plausible economic channels behind our key finding.

6.1 Priority fees and price discovery

Below we examine whether priority fees play an important role in the price discovery process through DEX trade flows. To do so, we estimate a structural VAR model as in Equation 3 where we include CEX return and DEX trade flows with different fee levels. In Section 6.1.1, we provide summary statistics of the CEX return, DEX trade flows, and CEX trade flow, variables used in the structural VAR model. We then analyze the impulse response analysis and report the permanent price impact and information shares of DEX trade flows in Section 6.1.2 and Section 6.1.3 respectively. In Section 6.1.4, we examine the speed of price discovery through DEX trade flows.

6.1.1 Summary statistics of CEX return and DEX trade flows

Before discussing the estimation results from the structural VAR model, for each token pair, we report summary statistics of the return and trade flow variables in Table 3. Several observations are in order. First, as expected, returns of NonStable token pairs are much more volatile. For instance, per-block-time (≈12 seconds) standard deviation of Binance return, r_t^{CEX}, is about 0.79, 10.27, and 9.12 basis points for USDC-USDT, ETH-USDT, and WBTC-ETH respectively. These results are expected because NonStable pairs consist of risky tokens such as Bitcoin and Ethereum and thus their prices respond to both short-term liquidity shocks and long-term information shocks. In contrast, Stable token pairs are only affected by short-term liquidity shocks as both of their tokens

Table 3. Summary statistics of CEX return, CEX trade flow, and DEX trade flow variables. This table reports, for each token pair, summary statistics of the return and trade flow variables used in the structural VAR estimation. r_t^{CEX} is Binance return from block time $t-1$ to t. x_t^{CEX} is Binance trade flow. x_t^{DEX} is Uniswap trade flows. $x_t^{\text{LowFee-DEX}}$, $x_t^{\text{MidFee-DEX}}$ and $x_t^{\text{HighFee-DEX}}$ are Uniswap trade flows consisting of trades from the low-, mid- and high-fee category in block t. Both r_t^{CEX} and r_t^{DEX} are in basis points. N refers to the number of blocks for each token pair during our sample period.

(a) Stable token pairs. All trade flow variables are denominated in thousand USD.

		N	Mean	SD	Min	50%	Max
USDC-USDT	r_t^{CEX}	66949	-0.00	0.79	-52.68	0.00	37.12
	x_t^{CEX}	66949	1.58	112.29	-3892.72	0.00	3277.79
	x_t^{DEX}	66949	0.01	7.95	-268.57	-0.02	227.86
	$x_t^{\text{LowFee-DEX}}$	66949	0.02	2.62	-118.09	0.00	85.00
	$x_t^{\text{MidFee-DEX}}$	66949	0.03	4.82	-145.88	0.00	150.00
	$x_t^{\text{HighFee-DEX}}$	66949	-0.04	5.75	-268.57	0.00	227.86
DAI-USDT	r_t^{CEX}	45868	-0.00	1.45	-46.36	0.00	40.56
	x_t^{CEX}	45868	-1.20	33.79	-1162.58	0.00	778.95
	x_t^{DEX}	45868	0.01	5.14	-142.20	-0.00	141.16
	$x_t^{\text{LowFee-DEX}}$	45868	0.02	1.90	-60.01	0.00	50.65
	$x_t^{\text{MidFee-DEX}}$	45868	-0.00	3.12	-81.14	0.00	64.00
	$x_t^{\text{HighFee-DEX}}$	45868	-0.01	3.55	-142.20	0.00	94.93

are pegged to the US Dollar.

Second, consistent with the liquidity summary statistics in Table 1, the magnitude of trade flows on Uniswap versus Binance differs significantly across token pairs. For the two Stable token pairs, USDC-USDT and DAI-USDT, the magnitude of the trade flow is much larger on Binance than on Uniswap. For example, the standard deviation of per-block-time trade flow of USDC-USDT on Binance is about 112 thousand USD, more than an order of magnitude larger than that of about eight thousand USD on Uniswap. In contrast, for the rest of the token pairs except for ETH-USDT, absolute trade flow is larger on Uniswap than on Binance. For example, the standard deviation of per-block-time trade flow of WBTC-ETH is about 56 ETH on Uniswap compared with 10 ETH on Binance.

Third, for all token pairs on Uniswap, trade flows with high fees are larger in magnitude than flows with middle and low gas fees. For example, the standard deviation of ETH-USDT high-fee trade flow is 33.18 ETH, which is more than three times larger than that of low-fee trade flow.

(b) NonStable token pairs. All trade flow variables are denominated in ETH.

		N	Mean	SD	Min	50%	Max
ETH-USDT	r_t^{CEX}	370291	0.03	10.27	-476.61	0.00	368.22
	x_t^{CEX}	370291	-0.32	221.19	-7370.94	0.11	10152.33
	x_t^{DEX}	370291	0.15	40.76	-3111.34	0.04	2154.22
	$x_t^{\text{LowFee-DEX}}$	370291	-0.03	10.29	-2345.49	0.00	1241.70
	$x_t^{\text{MidFee-DEX}}$	370291	-0.06	21.37	-1897.53	0.00	2147.57
	$x_t^{\text{HighFee-DEX}}$	370291	0.23	33.18	-3498.28	0.00	2217.48
WBTC-ETH	r_t^{CEX}	81892	-0.05	9.12	-269.32	0.00	245.93
	x_t^{CEX}	81892	-0.02	9.93	-395.21	0.00	1991.97
	x_t^{DEX}	81892	-0.25	56.17	-2750.21	0.22	2331.24
	$x_t^{\text{LowFee-DEX}}$	81892	0.07	15.87	-475.92	0.00	698.13
	$x_t^{\text{MidFee-DEX}}$	81892	0.07	36.64	-2750.21	0.00	726.66
	$x_t^{\text{HighFee-DEX}}$	81892	-0.40	39.13	-771.15	0.00	2331.24
LINK-ETH	r_t^{CEX}	72951	-0.07	16.10	-494.76	0.00	467.55
	x_t^{CEX}	72951	-0.47	16.73	-2047.56	0.00	432.04
	x_t^{DEX}	72951	-0.08	22.57	-1187.08	0.00	652.36
	$x_t^{\text{LowFee-DEX}}$	72951	-0.04	5.32	-202.07	0.00	161.16
	$x_t^{\text{MidFee-DEX}}$	72951	-0.10	14.47	-1187.08	0.00	652.36
	$x_t^{\text{HighFee-DEX}}$	72951	0.06	16.11	-432.35	0.00	541.94
AAVE-ETH	r_t^{CEX}	42975	0.14	29.89	-509.77	0.00	582.37
	x_t^{CEX}	42975	-0.31	10.83	-676.27	0.00	239.78
	x_t^{DEX}	42975	0.14	19.59	-417.79	0.10	374.95
	$x_t^{\text{LowFee-DEX}}$	42975	0.07	5.51	-150.28	0.00	225.81
	$x_t^{\text{MidFee-DEX}}$	42975	0.02	12.78	-417.79	0.00	192.39
	$x_t^{\text{HighFee-DEX}}$	42975	0.05	13.75	-221.06	0.00	374.95

6.1.2 Permanent price impacts of DEX trade flows

If DEX trade flows with high fees contain more private information than those with low fees, we expect the former to have a larger permanent price impact. In a structural VAR framework, the permanent price impact of a particular trade flow is estimated by the cumulative impulse responses of return to its unexpected component, as specified in Equation 5. In Table 4, we report the cumulative impulse responses of CEX return to DEX trade flows with different fee levels. In addition, to account for contemporaneous relations between the endogenous variables, we report the upper and lower bound of each variable's permanent price impact based on estimates across all possible orderings of the Cholesky decomposition. The results show that high-fee DEX trade flow is more

Table 4. Permanent price impact of DEX trade flows with different priority fee levels. This table reports the permanent price impacts of the DEX trade flows with high, medium, and low priority fee levels. Permanent price impacts are defined as the cumulative impulse responses of the CEX return to DEX trade flow in the structural VAR model (see Equation 5). Upper bounds (UB) and lower bounds (LB) are obtained by considering all possible sequences of the recursive contemporaneous causality among the endogenous variables. The last column reports the difference between the lower bound of the permanent price impact of the high-fee DEX trade flow and the lower bound of the permanent price impact of the low-fee DEX trade flow. The estimation of the structural VAR is done for each pair-day and statistical inference is based on variations in the pair-day estimates. Row variables are response variables and column variables are shock variables. CEX return is in basis points. DEX trade flows are standardized and thus in their standard deviations. *, ** and *** indicate significance levels at 1%, 5% and 10% respectively.

Variable	$x^{\text{LowFee-DEX}}$		$x^{\text{MidFee-DEX}}$		$x^{\text{HighFee-DEX}}$		$\Delta^{\text{HighFee - LowFee}}$
	LB	UB	LB	UB	LB	UB	LB - UB
Stable	0.0	0.01	0.0	0.0	0.0	0.01	−0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
NonStable	0.41***	0.94***	1.93***	3.86***	4.27***	8.16***	3.33***
	(0.1)	(0.12)	(0.17)	(0.27)	(0.22)	(0.37)	(0.22)

informed than low-fee DEX trade flow.

We begin by discussing the results for NonStable token pairs. These pairs consist of at least one non-stable coin (such as Bitcoin and Ethereum), and thus have frequent arrival of private information. Overall, the results show that DEX trade flows with higher fees have larger permanent price impacts. Specifically, the permanent price impact of the high-fee DEX trade flow, $x_t^{\text{HighFee-DEX}}$, has a lower and upper bound of 4.27 and 8.16 basis points respectively, meaning that a one standard deviation positive shock to the high-fee DEX trade flow leads to a permanent increase of CEX prices between 4.27 and 8.16 basis points. In contrast, the permanent price impact of low-fee DEX trade flow, $x_t^{\text{LowFee-DEX}}$, has a lower and upper bound of 0.41 and 0.94 basis points. Thus, a one standard deviation positive shock to the low-fee DEX trade flow only leads to a permanent increase of CEX prices between 0.41 and 0.94 basis points. To make a conservative comparison between the permanent price impacts of the high-fee DEX trade flow and the low-fee DEX trade flow, we calculate the difference between the lower bound of the former and the upper bound of the latter, resulting in a value of 3.33 basis points (see the last column $\Delta^{\text{HighFee}-\text{LowFee}}$). A sample t-test shows the positive difference is statistically significant. Note that the difference is economically significant as well. It is about three times larger than the permanent price impact of the low-fee DEX trade flow.

Table 5. Information shares of DEX trade flows with different priority fee levels. This table reports the information shares of the CEX return and DEX trade flows with different priority fee levels. Information shares are computed using the formula in Equation 9. Upper bounds (UB) and lower bounds (LB) are obtained by considering all possible sequences of the recursive contemporaneous causality among the endogenous variables. The last column reports the difference between the lower bound of the information share of the high-fee DEX trade flow and the lower bound of the information share of the low-fee DEX trade flow. The estimation of the structural VAR model is done for each pair-day and statistical inference is based on pair-day estimates. Numbers in brackets are standard errors.

Variable	r^{CEX}		$x^{\text{LowFee-DEX}}$		$x^{\text{MidFee-DEX}}$		$x^{\text{HighFee-DEX}}$		$\Delta^{\text{HighFee - LowFee}}$
	LB	UB	LB	UB	LB	UB	LB	UB	LB - UB
Stable	97.28	97.99	0.78	1.05	0.63	1.01	0.59	0.67	-0.46^{***}
	(0.24)	(0.18)	(0.1)	(0.13)	(0.08)	(0.14)	(0.07)	(0.07)	(0.15)
NonStable	77.09	91.08	0.55	0.81	1.52	4.17	6.73	18.06	5.92^{***}
	(0.79)	(0.44)	(0.05)	(0.07)	(0.14)	(0.3)	(0.4)	(0.67)	(0.41)

Next, we discuss the estimation results for pairs of Stable tokens. Stable token pairs carry little private or public information because both tokens of the pair are stable coins pegged to the US Dollar. Hence, without short-term liquidity shocks, token pairs should always be priced at one. As a result, traders of Stable pairs are either liquidity traders who would like to exchange one stablecoin for the other or arbitrageurs who respond to public information such as transitory price discrepancy of the token pairs between CEXs and DEXs. Both types of trades can only impose a transitory impact on the prices, but not a permanent one. These results are consistent with intuition: the cumulative impulse responses of DEX trade flows are statistically insignificant, regardless of fee levels.

6.1.3 Information shares of DEX trade flows

In addition to permanent price impacts, we further compute the information shares of the DEX trade flows of different priority fee levels. The information share measure considers both the permanent price impact of a trade flow variable and its own innovation (unexpected) variance. So, if a trade flow is harder to predict based on return and trade flow history or it has larger variability, it will have a larger information share. Table 5 reports the results.

We first turn to results for the NonStable token pairs. There are two main observations. First, CEX return innovation contributes the largest share, ranging between 77.09% and 91.08%, to price discovery. As the contribution of the CEX return innovation reflects the relative amount of public information incorporated into the efficient price, it indicates that the CEX midquote changes are predominantly driven by the arrival of public news. More precisely, upon observing public news, market makers on CEX quickly revise their stable quotes before the arrival of trades on DEX. Thus, the information content of the public news is reflected in the quote changes directly, not predicted by trades.[13]

Second, in line with the permanent price impact results, high-fee DEX trade flow contributes a much larger share to price discovery than low-fee DEX trade flow: the information share of the high-fee DEX trade flow has a lower and upper bound of 6.73% and 18.06% while that of the low-fee DEX trade flow has a lower and upper bound of 0.55% and 0.81%. Again, we compute the difference between the lower bound of the information share of the high-fee DEX trade flow and the upper bound of the information share of the low-fee DEX trade flow. The difference is 5.92% and statistically and economically significant. In contrast, for Stable token pairs, CEX return innovation itself contributes to almost all (between 97.28% and 97.99%) price discovery while DEX trade flows contribute barely any.

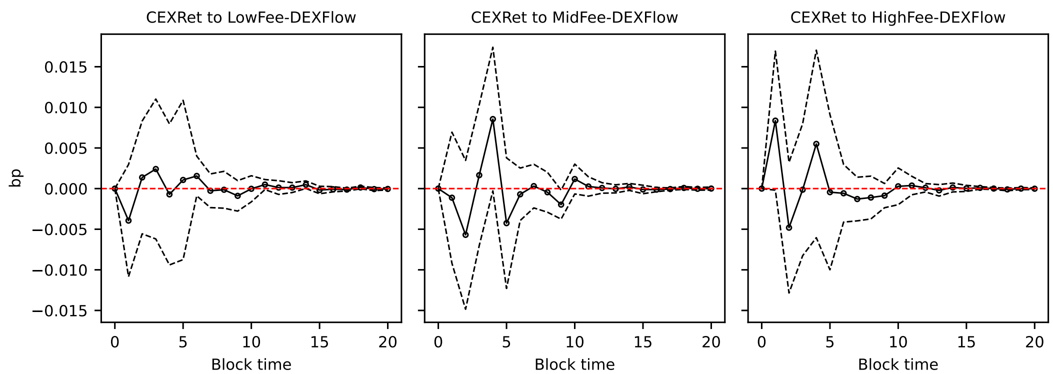
6.1.4 DEX trade flows and speed of price discovery

In previous sections, we have shown that the high-fee DEX trade flow has a much larger permanent price impact and information share than the low-fee DEX trade flow, contributing more to price discovery. However, as the permanent price impact is defined as the cumulative impulse responses of CEX return, it can not speak to the speed of price discovery. How quickly does the CEX price adjust to the private information revealed through DEX trade flows? To examine it, we turn to the dynamics of impulse responses of CEX return to DEX trade flows.

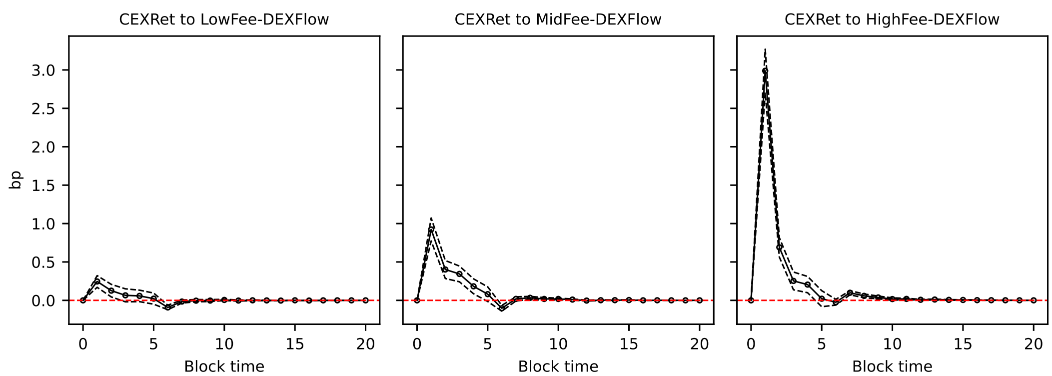
[13] Previous applications of the approach report similar numbers: 65% for a sample of 177 NYSE stocks in the first quarter of 1989 (Hasbrouck, 1991b), between 72% and 88% for a sample of NASDAQ stocks in June 2000 (Barclay, Hendershott, and McCormick, 2003)

Figure 2. Impulse response functions of CEX return to DEX trade flows with different priority fee levels. This figure plots the full impulse responses of the CEX return to a one-standard-deviation shock to trade flows of high-fee, mid-fee, and low-fee levels over the horizon of 20 blocks. Impulse responses are obtained by converting the estimated structural VAR model to its VMA form. To have conservative estimates of the price impacts of the high-fee DEX trade flow, we use the following sequence of variables: CEX return, low-fee DEX trade flow, mid-fee DEX trade flow, and high-fee DEX trade flow (see Equation 3) for the Cholesky decomposition. CEX return is measured in basis points and DEX trade flows are standardized and thus measured in standard deviation units. We estimate the structural VAR model for each pair-day, and the statistical inference is based on pair-day estimates. Dashed black lines represent 95% confidence bands.

(a) Stable token pairs.



(b) NonStable token pairs.



To obtain conservative estimates for the return impulse responses to the DEX trade flows, we use the following sequence for the endogenous variables: CEX return, low-fee DEX trade flow, mid-fee DEX trade flow, and high-fee DEX trade flow. In other words, we impose the restriction that the CEX return contemporaneously causes the DEX trade flows, but not vice versa. Under such a restriction, we estimate the structural VAR model and plot the impulse responses of the CEX return to a one-standard-deviation shock to the DEX trade flows for the Stable and NonStable pairs in Panel (a) and Panel (b) of Figure 2 the respectively.

There are several important observations. First, consistent with the permanent price impact results reported in Table 4, the impulse responses of CEX return to DEX trade flow for Stable pairs are statistically insignificant over all periods, regardless of its priority fee levels. Second, for NonStable pairs, the return impulse responses are significant and peak at the first subsequent period (block time $t = 1$). Note that given the specific recursive structure that we impose, the return impulse responses to all three DEX trade flows at the contemporaneous period ($t = 0$) are zero. In addition, the return impulse responses drop significantly from the second period (block time $t = 2$), especially to a shock in the high-fee DEX trade flow. It indicates that CEX return responds significantly and quickly to DEX trade flows and most price discovery through DEX trade flow is realized within the subsequent block time. Hence, traders are able to learn the private information contained in the DEX trade flow quickly and update their beliefs on the new price.

6.2 Robustness checks

6.2.1 Robustness: Accounting for CEX trade flow

Informed traders might execute their trades both on centralized and decentralized exchanges. Thus, trade flow on the centralized exchange can contain private information. To control for it, we include Binance trade flow, x_t^{CEX}, in the endogenous variable vector:

$$y_t = \begin{pmatrix} r_t^{\text{CEX}} & x_t^{\text{CEX}} & x_t^{\text{LowFee-DEX}} & x_t^{\text{MidFee-DEX}} & x_t^{\text{HighFee-DEX}} \end{pmatrix}' \tag{10}$$

where x_t^{CEX} is the signed trade flow on Binance aggregated between block time $t-1$ and t. r_t^{CEX}, $x_t^{\text{LowFee-DEX}}$, $x_t^{\text{MidFee-DEX}}$ and $x_t^{\text{HighFee-DEX}}$ have all been introduced above, and we recall that they represent Binance return and Uniswap trade flows with low, mid and high gas fee levels respectively.

As in the baseline case, we do not impose specific contemporaneous relations among the endogenous variables as there is no certain economic reasoning for them. For example, it is not clear whether it is Binance trades that contemporaneously cause Uniswap trades or the other way around. For one thing, informed traders might split their orders and trade simultaneously on both exchanges. For another, even if there is a specific sequence between the two trade flows, it might be only observable at frequencies higher than the block time we currently use. As a result, we consider all possible recursive contemporaneous relations among the endogenous variables to obtain the upper and lower bounds of their permanent price impacts and information shares. Specifically, we impose the A matrix as a lower triangular matrix and permute the sequence of the endogenous variables. As we now have five variables in total, we consider 60 (5!) different sequences.

Table 6 and 7 report the results for permanent price impacts and information shares respectively. The key takeaway is that the results remain qualitatively the same as the baseline model: high-fee DEX trade flow has a much larger permanent price impact and information share than low-fee DEX trade flows. For NonStable pairs, the permanent price impact of the high-fee DEX trade flow has an upper and lower bound of 2.53 and 4.12 basis points respectively. In contrast, the lower bound of the permanent price impact of the low-fee DEX trade flow is negative, albeit marginally significant, and the upper bound is not significant.

In addition, there are two points worth discussing. First, we note that controlling for CEX trade flow results in a decrease in the permanent price impact of the high-DEX trade flows. For NonStable token pairs, the upper and lower bound of the permanent price impact of the high-fee DEX trade flow are 4.27 and 8.76 basis points without the CEX trade flow. They drop to 2.53 and 4.12 basis points respectively after CEX trade flow is controlled for. In addition, the upper and lower bound of the permanent price impact of the CEX trade flow are 0.93 and 3.89 basis points, which are comparable with those of the permanent price impact of the high-fee DEX trade flow. It indicates that privately informed traders indeed trade on both the CEX and DEX, possibly to reduce the price impact of the trades.

6.2.2 Robustness: Controlling for the confounding effect of trade size

Priority fee is a fixed cost regardless of the trade size. Thus, traders are willing to pay a higher priority fee for large trades as it is relatively cheaper. So, trade size and fee are positively correlated. In addition, it is a well-known fact that large trades tend to have a larger price impact than small

Table 6. Robustness: Permanent price impact of DEX trade flows with different priority fee levels. This table reports the permanent price impacts of the CEX trade flow and DEX trade flows with high, medium, and low priority fee levels. Permanent price impacts are defined as the cumulative impulse responses of the CEX return to DEX trade flow in the structural VAR model (see Equation 5). Upper bounds (UB) and lower bounds (LB) are obtained by considering all possible sequences of the recursive contemporaneous causality among the endogenous variables. The last column reports the difference between the lower bound of the permanent price impact of the high-fee DEX trade flow and the lower bound of the permanent price impact of the low-fee DEX trade flow. The estimation of the structural VAR is done for each pair-day and statistical inference is based on variations in the pair-day estimates. Row variables are response variables and column variables are shock variables. CEX return is in basis points. DEX trade flows are standardized and thus in their standard deviations. *, ** and *** indicate significance levels at 1%, 5% and 10% respectively.

Variable	x^{CEX}		$x^{\text{LowFee-DEX}}$		$x^{\text{MidFee-DEX}}$		$x^{\text{HighFee-DEX}}$		$\Delta^{\text{HighFee - LowFee}}$
	LB	UB	LB	UB	LB	UB	LB	UB	LB - UB
Stable	0.13^{***}	0.37^{***}	0.0	0.01	0.0	0.0	0.0	0.0	-0.01
	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
NonStable	0.93^{***}	3.89^{***}	-0.1^{*}	-0.03	0.24^{***}	0.56^{***}	2.53^{***}	4.12^{***}	2.56^{***}
	(0.13)	(0.24)	(0.05)	(0.05)	(0.07)	(0.1)	(0.3)	(0.42)	(0.31)

Table 7. Robustness: Information shares of DEX trade flows with different priority fee levels. This table reports the information shares of the CEX return, CEX trade flow, and DEX trade flows with different priority fee levels. Information shares are computed using the formula in Equation 9. Upper bounds (UB) and lower bounds (LB) are obtained by considering all possible sequences of the recursive contemporaneous causality among the endogenous variables. The last column reports the difference between the lower bound of the information share of the high-fee DEX trade flow and the lower bound of the information share of the low-fee DEX trade flow. The estimation of the structural VAR model is done for each pair-day and statistical inference is based on pair-day estimates. Numbers in brackets are standard errors.

Variable	r^{CEX}		x^{CEX}		$x^{\text{LowFee-DEX}}$		$x^{\text{MidFee-DEX}}$		$x^{\text{HighFee-DEX}}$		$\Delta^{\text{HighFee - LowFee}}$
	LB	UB	LB	UB	LB	UB	LB	UB	LB	UB	LB - UB
Stable	70.5	91.03	6.89	26.96	0.79	1.06	0.62	0.96	0.61	0.7	-0.45^{***}
	(1.33)	(0.64)	(0.61)	(1.35)	(0.1)	(0.14)	(0.08)	(0.14)	(0.08)	(0.08)	(0.16)
NonStable	73.08	92.44	1.3	15.17	0.15	0.17	0.27	0.57	5.41	13.91	5.25^{***}
	(1.29)	(0.86)	(0.2)	(0.68)	(0.02)	(0.02)	(0.04)	(0.08)	(0.69)	(1.23)	(0.69)

trades (Easley and O'Hara, 1987). Thus, trade size has a potential confounding effect on the price impact of fees.

To alleviate the concern, we further partition DEX trades based on their size in addition to fees and examine whether, within the same size group, trades with higher fees have a larger price impact. Specifically, we classify DEX trades into two size groups: a large-size group consisting of trades with a size above its 90% quantile and a small-size group consisting of trades with a size below its 90% quantile. Thus, our large-size group captures very large trades on the right tail of the size distribution. We choose 90% quantile as the cutoff point so that for the large-size group, the average trade size is similar across our three fee levels as shown in Table 8. For some token pairs such as WBTC-ETH, LINK-ETH, and AAVE-ETH, within the large-trade group, the average trade size of the mid-fee group is even larger than the high-fee group.

Table 8. Average trade size by trade size group and priority fee level. N refers to the number of trades in our sample. This table reports, for each token pair, the average size of trades categorized into different size groups and fee levels. The size group "Below Q90(TxSize)" includes trades with a size below its 90% quantile. The size group "Above Q90(TxSize)" includes trades with a size above its 90% quantile. The trade size for Stable pairs (USDC-USDT and DAI-USDT is in thousand USDT and the trade size for NonStable pairs (ETH-USDT, WBTC-ETH, LINK-ETH, and AAVE-ETH) is in ETH.

	GasPriceLevel	LowFee	MidFee	HighFee
TokenPair	TxSizeLevel			
USDC-USDT	Below Q90(TxSize)	1.23	1.48	2.01
	Above Q90(TxSize)	17.55	19.93	21.72
DAI-USDT	Below Q90(TxSize)	0.87	1.02	1.31
	Above Q90(TxSize)	12.19	12.87	13.90
ETH-USDT	Below Q90(TxSize)	1.65	2.12	3.16
	Above Q90(TxSize)	55.44	63.57	68.60
WBTC-ETH	Below Q90(TxSize)	5.79	9.27	17.04
	Above Q90(TxSize)	148.94	151.65	126.60
LINK-ETH	Below Q90(TxSize)	2.86	4.63	9.31
	Above Q90(TxSize)	50.90	62.22	49.73
AAVE-ETH	Below Q90(TxSize)	3.74	6.86	12.67
	Above Q90(TxSize)	55.44	58.84	48.02

Based on our size and fee grouping above, we construct six DEX trade flows: small-size and low-fee DEX trade flow ($x^{\text{S-L-DEX}}$), small-size and medium-fee DEX trade flow ($x^{\text{L-M-DEX}}$), small-size and high-fee DEX trade flow ($x^{\text{L-H-DEX}}$), large-size and low-fee DEX trade flow ($x^{\text{L-L-DEX}}$), large-size and medium-fee DEX trade flow ($x^{\text{L-M-DEX}}$), and large-size and high-fee DEX trade flow ($x^{\text{L-H-DEX}}$). Then we estimate a structural VAR model based on the six DEX trade flows.

Table 9 and Table 10 report the permanent price impacts and information shares of the six DEX trade flow by size group and fee level. Focusing on the NonStable token pairs, the results show that, consistent with the literature, large trades in general contain more private information. We see that overall, DEX trade flows in the large-trade group ($x^{\text{L-L-DEX}}$, $x^{\text{L-M-DEX}}$, and $x^{\text{L-H-DEX}}$) have larger price impacts and information shares than flows in the small-trade group ($x^{\text{S-L-DEX}}$, $x^{\text{S-M-DEX}}$, and $x^{\text{S-H-DEX}}$). More importantly, the results further show that, within the same trade size group, high-fee DEX trade flow has a larger price impact than medium-fee and low-fee flows. Focus on the large trade size group where the average trade size is very similar across different priority fee levels. The permanent price impact of the high-fee DEX trade flow has a lower and upper bound of 4.54 and 9.22 basis points, while that of the low-fee DEX trade flow has a lower and upper bound of 0.49 and 1.14 basis points. Last, it is worth noting that the difference between the price impact of high-fee and low-fee DEX trade flows is more pronounced for the large-trade group, which reflects the positive interaction effect between priority fees and trade size. In summary, the above results show that the priority fee, a unique feature of DEX trades, contains additional private information content not captured by the trade size.

Table 9. Robustness: Permanent price impact of DEX trade flows with different priority fee levels and trade size groups. This table reports the permanent price impacts of the CEX trade flow and DEX trade flows by priority fee levels and trade size levels. Permanent price impacts are defined as the cumulative impulse responses of the CEX return to DEX trade flow in the structural VAR model (see Equation 5). Upper bounds (UB) and lower bounds (LB) are obtained by considering all possible sequences of the recursive contemporaneous causality among the endogenous variables. The last column reports the difference between the lower bound of the permanent price impact of the large-size, high-fee DEX trade flow and the lower bound of the permanent price impact of the large-size, low-fee DEX trade flow. The estimation of the structural VAR is done for each pair-day and statistical inference is based on variations in the pair-day estimates. Row variables are response variables and column variables are shock variables. CEX return is in basis points. DEX trade flows are standardized and thus in their standard deviations. *, ** and *** indicate significance levels at 1%, 5% and 10% respectively.

Variable	$x^{\text{S-L-DEX}}$		$x^{\text{S-M-DEX}}$		$x^{\text{S-H-DEX}}$		$x^{\text{L-L-DEX}}$		$x^{\text{L-M-DEX}}$		$x^{\text{L-H-DEX}}$		$\Delta^{\text{L-H - L-L}}$
	LB	UB	LB	UB	LB	UB	LB	UB	LB	UB	LB	UB	LB - UB
Stable	−0.01	0.0	−0.01	0.0	0.0	0.0	0.0	0.01*	0.0	0.0	0.0	0.0	−0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
NonStable	−0.07	0.11	0.2	0.58***	0.44***	0.88***	0.49***	1.14***	2.25***	4.6***	4.54***	9.22***	3.4***
	(0.13)	(0.14)	(0.14)	(0.14)	(0.17)	(0.18)	(0.13)	(0.16)	(0.22)	(0.33)	(0.26)	(0.45)	(0.27)

Table 10. Robustness: Information shares of DEX trade flows with different priority fee levels and trade size groups. This table reports the information shares of the CEX return, CEX trade flow, and DEX trade flows with different priority fee levels. Information shares are computed using the formula in Equation 9. Upper bounds (UB) and lower bounds (LB) are obtained by considering all possible sequences of the recursive contemporaneous causality among the endogenous variables. The last column reports the difference between the lower bound of the information share of the large-size, high-fee DEX trade flow and the lower bound of the information share of the large-size, low-fee DEX trade flow. The estimation of the structural VAR model is done for each pair-day and statistical inference is based on pair-day estimates. Numbers in brackets are standard errors.

Variable	r^{CEX}		$x^{\text{S-L-DEX}}$		$x^{\text{S-M-DEX}}$		$x^{\text{S-H-DEX}}$		$x^{\text{L-L-DEX}}$		$x^{\text{L-M-DEX}}$		$x^{\text{L-H-DEX}}$		$\Delta^{\text{LH - LL}}$
	LB	UB	LB	UB	LB	UB	LB	UB	LB	UB	LB	UB	LB	UB	LB - UB
Stable	94.38	95.35	1.07	1.26	0.81	0.99	0.9	1.04	0.63	0.81	0.66	0.95	0.52	0.61	−0.3***
	(0.52)	(0.51)	(0.21)	(0.21)	(0.13)	(0.13)	(0.2)	(0.21)	(0.09)	(0.1)	(0.12)	(0.12)	(0.08)	(0.09)	(0.11)
NonStable	74.86	90.1	0.55	0.68	0.59	0.67	0.7	0.93	0.53	0.79	1.61	4.56	5.73	17.78	4.95***
	(0.95)	(0.49)	(0.07)	(0.08)	(0.07)	(0.07)	(0.09)	(0.11)	(0.06)	(0.08)	(0.18)	(0.37)	(0.4)	(0.76)	(0.41)

6.3 Priority fees and information: Economic channels

In the above section, we have shown that high-fee DEX trade flow contains more private information than low-fee DEX trade flow, suggesting that privately informed traders bid high fees to execute their orders on DEXs. Next, we provide plausible economic channels to explain the results and use mempool order data to test them.

6.3.1 Two potential economic channels

Channel #1: Execution risk due to blockchain congestion Trading on DEXs is not the only activity on a blockchain. Other non-DEX activities such as payment transfer, borrowing and lending, non-fungible token (NFTs) auctions, and initial coin offerings (ICOs) take up limited block space as well. In particular, if there is a surge of non-DEX activities which make blocks congested, the marginal priority fee needed to execute a transaction increases, driving up the transaction cost for traders on DEXs.

During such times, in contrast to a patient and uninformed trader, a trader who possesses short-lived private information, e.g, over the next several blocks, might bid a high fee to avoid execution risk if the gain from her trade is large.[14] Ideally, she would like to set her bid to the marginal fee to guarantee execution in the next block. However, the marginal fee of the next block is not perfectly predictable. For example, even if the informed trader actively monitors all pending orders received by its mempool, due to network latency, pending orders seen by her can be different from the ones seen by the validators. As a result, she will bid a fee higher than the expected marginal fee to reduce her execution risk.

What this implies is that, if an informed trader only faces execution risk, she will choose a high, but not too high blockchain priority fee for her trades, compared with other transactions in the same block. In terms of the block position, her trades will likely be located around the middle

[14]We note that impatient and uninformed traders (e.g., liquidity traders who receive marginal calls and have to liquidate their positions) can bid high priority fees to avoid execution risk as well. However, their trades contain no private information and thus can not drive our findings in the above section that high-fee trades are more informative.

of the block, but not at the very top.

Channel #2: Competition among informed traders An informed trader will bid a high fee if the blockchain network is congested. However, this may not be the only channel; she might bid a high fee if she faces competition from other traders.

It is unclear ex-ante whether such a channel exists as theoretical literature has mixed predictions about informed trading and its implications on price discovery. Competition arises when private information is not only possessed by one informed trader; instead, there are multiple traders who receive either the same or highly correlated private signals (See, e.g., Holden and Subrahmanyam, 1992; Foster and Viswanathan, 1996; Back, Cao, and Willard, 2000).

Another possibility is that there are "back-runners" (Yang and Zhu, 2020) or "predators" (Brunnermeier, 2005) who are not endowed with private signals but infer them from public signals such as order imbalance or priority fees in the context of DEXs. However, informed traders can select the timing of their trades. For example, they might trade when the liquidity of the target token pairs is high such that their trades lead to a low price impact and can not be easily detected (Collin-Dufresne and Fos, 2015).

When facing competition from other traders with the same or similar information, an informed trader might have to bid a priority fee much higher than the rest of non-DEX transactions in the same block, especially when the potential profit from the information is high. In such cases, we might observe DEX trades with excessively high fees located at the very top of the block.

6.3.2 Do privately informed traders compete on DEXs?

Identify "excessively-high-fee trades" As explained above, competition among informed traders can lead to excessively high priority fees for DEX trades compared with other non-DEX transactions executed in the same block. How high a fee needs to be in order to be regarded as "excessive"? To choose the right threshold for the priority fee, we use the inter-quartile range (IQR) method, a commonly used outlier detection approach in statistics.[15] Specifically, for each block, we first calculate the 25% quantile (Q25) and 75% quantile (Q75) of the priority fees of all executed transactions in the block[16], including both DEX trades and non-DEX transactions. Then we calculate the IQR, defined as the difference between the 75% quantile and 25% quantile, that is, IQR = Q75 − Q25. Finally, we obtain the threshold Q75 + 1.5 × IQR and label DEX trades with a priority fee higher than the threshold as "excessively-high-fee trades".[17]

Information content of "excessively-high-fee trades" Note that DEX trades with excessively high fees or located at the very top of the block can include three different types of trades: (1) trades driven by competition among privately informed traders; (2) trades driven by competition among arbitrageurs on public information (e.g., price discrepancies between CEXs and DEXs); (3) trades by impatient and uninformed traders (e.g., liquidation trades triggered by marginal calls). However, only the first type of trades, which are driven by competition among privately informed traders, contain private information and thus can have permanent price impacts on the CEX returns.

To examine whether our identified trades include the first type of trades with private information, we reconstruct DEX trade flows with different priority fee levels excluding all "excessively-high-fee" trades and then re-implement the structural VAR analysis. The idea is that if a significant share of high-fee trades results from competition among privately informed traders, we should see their permanent price impact become significantly smaller in magnitude after we exclude the "excessively-high-fee trades".

Table 11 reports the permanent price impacts of the DEX trade flows when "excessively-high-fee trades" are excluded. Focus on the NonStable pairs. It shows that, compared with the baseline

[15]We prefer the IQR method, a quantile-based approach, over other outlier detection methods based on standard deviations as the priority fee distribution is not normal but right-skewed.

[16]We obtain the executed transactions data on the Ethereum blockchain from Blockchair (https://gz.blockchair.com/ethereum/transactions/).

[17]Alternatively, one can identify such trades based on their block position. As transactions executed in the same block are ranked based on their priority fees in descending order. Thus, transactions with higher priority fees will be placed more at the front of the block. Specifically, one can choose a threshold for the block position, say top 10%, and then label DEX trades located more before the threshold. We tested the alternative approach and the results are qualitatively the same.

Table 11. Permanent price impact of DEX trade flows with different priority fee levels: Excluding "excessively-high-fee trades". This table reports the permanent price impacts of the DEX trade flows with high, medium, and low priority fee levels. Permanent price impacts are defined as the cumulative impulse responses of the CEX return to DEX trade flow in the structural VAR model (see Equation 5). Upper bounds (UB) and lower bounds (LB) are obtained by considering all possible sequences of the recursive contemporaneous causality among the endogenous variables. The last column reports the difference between the lower bound of the permanent price impact of the high-fee DEX trade flow and the lower bound of the permanent price impact of the low-fee DEX trade flow. The estimation of the structural VAR is done for each pair-day and statistical inference is based on variations in the pair-day estimates. Row variables are response variables and column variables are shock variables. CEX return is in basis points. DEX trade flows are standardized and thus in their standard deviations. *, ** and *** indicate significance levels at 1%, 5% and 10% respectively.

Variable	$x^{\text{LowFee-DEX}}$		$x^{\text{MidFee-DEX}}$		$x^{\text{HighFee-DEX}}$		$\Delta^{\text{HighFee - LowFee}}$
	LB	UB	LB	UB	LB	UB	LB - UB
Stable	0.0	0.01	0.0	0.0	0.0	0.0	−0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
NonStable	0.52***	1.01***	2.04***	3.97***	2.83***	5.36***	1.82***
	(0.13)	(0.14)	(0.19)	(0.28)	(0.21)	(0.31)	(0.21)

results where all trades are included in Table 4, the lower and upper bound of the permanent price impact of the high-fee DEX trade flow drop significantly in magnitude from 4.27 and 8.16 basis points to 2.83 and 5.36 basis points. The results illustrate that our key results—high-fee DEX trade flow is more privately informed —are in part driven by competition among privately informed traders in addition to them avoiding execution risk.

6.3.3 How do informed traders compete on DEXs?

In the above section, we have shown that competition among privately informed traders on DEXs is a significant driving force of our key finding that high-fee DEX trade flow contains more private information. Next, we investigate what fee bidding strategy informed traders use to compete with each other on DEXs.

Identify trades from priority gas auctions (PGAs) As pending orders in the mempools are publicly visible to all traders who actively monitor them, one natural bidding strategy is that informed traders competitively bid up their priority fees, a process known as the priority gas auction (PGA) in the literature (Daian et al., 2020). But is it the dominant bidding strategy? For an executed trade to quality as a PGA trade, we require the following criteria:

1. **The executed trade has at least one matched mempool order with the same submission address and nonce.** Recall that a trader on DEX needs to attach a number called "nonce" to each of her orders. The most important property of a nonce is that each number can only be used once and it must be used in a consecutively increasing order. For example, a new order broadcast by a trader needs to have a new nonce increased by 1 compared with the previous order. More importantly, a trader's order with a larger nonce cannot be executed before one with a smaller nonce. This implies that if a trader wants to modify her pending order, e.g., increase the fee, she needs to broadcast a new order with the same nonce as the pending one. Hence, the first criterion on submission address and nonce guarantees that the matched mempool orders are previous revisions of the final executed order.

2. **The gas price of the executed trade must be higher than that of its matched order(s).** We observe the gas price attached to both mempool orders and the executed trade. The second criterion requires that the executed trade must have a higher gas fee than its matched order(s) (i.e., those with the same submission address and nonce) so that we capture trades associated with fee competition.

3. **All matched orders of the executed trade must arrive at the mempool within five blocks.** Specifically, to be matched with a trade executed at block time t, orders must arrive in the mempool during the block time interval of $(t-5, t]$. We believe gas bidding due to competition should happen within a fairly short time window. If the window is too long, the bid update is more likely to result from patient liquidity traders revising their fees to reduce the waiting time.

Fraction of PGA trades We implement the foregoing identification strategy above and Table 12 reports, for each token pair, the fraction of PGA trades for both trades with excessively high fees ("excessively-high-fee trades") and other trades ("other trades"). There are two notable observations. First, the overall fraction of executed trades identified as PGA trades is very small. For example, for the group of "Other trades", less than 5% of them are identified as PGA trades across the six token pairs.

Table 12. Percentages of priority gas auction (PGA) trades. This table shows the fraction of trades identified as priority gas auction (PGA) trades, for "excessively-high-fee trades" and other trades.

TokenPair	ExplicitCompetition ExcessiveGas	Non-PGA trades	PGA trades
USDC-USDT	Other trades	98.06	1.94
	Excessively-high-fee trades	96.24	3.76
DAI-USDT	Other trades	97.90	2.10
	Excessively-high-fee trades	95.60	4.40
ETH-USDT	Other trades	97.68	2.32
	Excessively-high-fee trades	87.95	12.05
LINK-ETH	Other trades	95.06	4.94
	Excessively-high-fee trades	73.29	26.71
WBTC-ETH	Other trades	96.61	3.39
	Excessively-high-fee trades	84.79	15.21
AAVE-ETH	Other trades	95.39	4.61
	Excessively-high-fee trades	81.72	18.28

Surprisingly, even if we zoom in on the "excessively-high-fee trades" which include trades likely driven by competition, only a minority of them are identified as PGA trades. Across the six token pairs, the fraction of PGA trades out of "excessively-high-fee trades" varies between 1.94% for USDC-USDT and 26.71% for LINK-ETH. The result suggests that the PGA type of bidding strategy is not the dominant one used by informed traders. Instead of competitively bidding up the fee, they start with bidding a very high fee, which resembles the jump bidding strategy in auction theory (Daniel and Hirshleifer, 1998; Avery, 1998).

The motivation for adopting such a bidding strategy is that, by bidding a high fee in the first place, an informed trader can discourage competition from other traders. First, by bidding a high fee, she can signal that her valuation of the information is high, and if bidding is costly, it is optimal for potential competitors to drop out. Second, even if all traders value the information the same and there is no bidding cost, it remains optimal for others to drop out as winning over an aggressive bid from the jump bidder subjects one to a greater Winner's Curse.

7 Conclusion

In this paper, we study the price discovery process on decentralized exchanges (DEXs). Unlike CEXs, DEXs execute orders in batches and determine their execution priority based on the priority fee bid by traders. Using a structural VAR model, we have shown that high-fee DEX trades reveal more private information than low-fee trades. In addition, we tested the possible economic channels driving this high-fee bidding behavior using a unique data set of Ethereum mempool orders. Our findings indicate that informed traders bid high fees not only to reduce execution risk due to blockchain congestion, but also to compete with each other via a jump bidding strategy.

We have shown that the priority fee bidding mechanism of DEXs plays a crucial role in price discovery. However, we can not claim that such a trading mechanism is conducive to overall market quality. While the priority fee bidding mechanism can potentially lead to faster information revelation, it might not help liquidity providers mitigate the adverse selection risk as losses are shared among the liquidity providers and thus each of them does not have the incentive to outbid informed traders. As a result, liquidity providers might provide less liquidity on DEXs (Capponi and Jia, 2021). Future studies should take both price discovery and market liquidity into consideration to fully assess the effect of the priority fee biding mechanism on the market quality of DEXs.

Table A1. Robustness: Permanent price impact of DEX trade flows with different priority fee levels: Unstandardized trade flows. This table reports the permanent price impacts of the CEX trade flow and DEX trade flows with high, medium, and low priority fee levels. Permanent price impacts are defined as the cumulative impulse responses of the CEX return to DEX trade flow in the structural VAR model (see Equation 5). Upper bounds (UB) and lower bounds (LB) are obtained by considering all possible sequences of the recursive contemporaneous causality among the endogenous variables. The last column reports the difference between the lower bound of the permanent price impact of the high-fee DEX trade flow and the lower bound of the permanent price impact of the low-fee DEX trade flow. The estimation of the structural VAR is done for each pair-day and statistical inference is based on variations in the pair-day estimates. Row variables are response variables and column variables are shock variables. CEX return is in basis points. DEX trade flows standardized. For Stable pairs, the DEX trade flows are in thousand USDT. For NonStable pairs, the DEX trade flows are in ETH. *, ** and *** indicate significance levels at 1%, 5% and 10% respectively.

Variable	$x^{\text{LowFee-DEX}}$		$x^{\text{MidFee-DEX}}$		$x^{\text{HighFee-DEX}}$		$\Delta^{\text{HighFee - LowFee}}$
	LB	UB	LB	UB	LB	UB	LB - UB
Stable	0.01	0.03	0.0	0.0	0.0	0.0	−0.02
	(0.02)	(0.02)	(0.0)	(0.01)	(0.01)	(0.01)	(0.02)
NonStable	0.08^{***}	0.18^{***}	0.14^{***}	0.3^{***}	0.25^{***}	0.52^{***}	0.07^{***}
	(0.02)	(0.03)	(0.01)	(0.02)	(0.02)	(0.03)	(0.03)

A Other robustness checks

In the appendix, we conduct three other robustness checks.

A.1 Unstandardized DEX trade flows

In our baseline estimation, we standardize the DEX trade flows to have zero mean and unit variance so that we have a fair comparison across token pairs. As a result, permanent price impacts of the DEX trade flows mean cumulative return impulse responses to a one standard deviation shock in the trade flows. As a robustness check, we use the unstandardized, original levels of the DEX trade flows instead. Specifically, for Stable pairs, the DEX trade flows are in thousand USDT and, for NonStable pairs, they are in ETH. Table A1 reports the estimation results. They show that in terms of per-ETH permanent price impact, the high-DEX trade flow remains more informative than the low-fee DEX trade flow.

Table A2. Permanent price impact of DEX trade flows with different priority fee levels: Fee level classification based on a rolling window of alternative lengths. This table reports the permanent price impacts of the DEX trade flows with high, medium, and low priority fee levels. Permanent price impacts are defined as the cumulative impulse responses of the CEX return to DEX trade flow in the structural VAR model (see Equation 5). Upper bounds (UB) and lower bounds (LB) are obtained by considering all possible sequences of the recursive contemporaneous causality among the endogenous variables. The last column reports the difference between the lower bound of the permanent price impact of the high-fee DEX trade flow and the lower bound of the permanent price impact of the low-fee DEX trade flow. The estimation of the structural VAR is done for each pair-day and statistical inference is based on variations in the pair-day estimates. Row variables are response variables and column variables are shock variables. CEX return is in basis points. DEX trade flows are standardized and thus in their standard deviations. *, ** and *** indicate significance levels at 1%, 5% and 10% respectively.

(a) Fee level classification based on a rolling window of 10 blocks.

Variable	$x^{\text{LowFee-DEX}}$		$x^{\text{MidFee-DEX}}$		$x^{\text{HighFee-DEX}}$		$\Delta^{\text{HighFee - LowFee}}$
	LB	UB	LB	UB	LB	UB	LB - UB
Stable	0.0	0.01	0.0	0.0	0.0	0.01	0.0
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
NonStable	0.48***	1.08***	2.08***	4.01***	4.18***	8.04***	3.11***
	(0.12)	(0.15)	(0.17)	(0.26)	(0.22)	(0.37)	(0.23)

(b) Fee level classification based on a rolling window of 40 blocks.

Variable	$x^{\text{LowFee-DEX}}$		$x^{\text{MidFee-DEX}}$		$x^{\text{HighFee-DEX}}$		$\Delta^{\text{HighFee - LowFee}}$
	LB	UB	LB	UB	LB	UB	LB - UB
Stable	0.01	0.01	−0.01***	−0.01*	0.01**	0.01**	0.0
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
NonStable	0.51***	1.04***	1.71***	3.6***	4.32***	8.19***	3.28***
	(0.1)	(0.14)	(0.16)	(0.26)	(0.23)	(0.38)	(0.24)

A.2 Priority fee level classification

In the baseline estimation, we use 20 blocks as the length of the rolling window in fee-level classification. As a robustness check, we try two different window lengths, 5 blocks, and 10 blocks, to classify DEX trades and then redo the structural VAR estimation. Table A2 reports the estimation results of the cumulative return impulse responses based on DEX trade flows from the two alternative gas level classifications. It shows that the results are largely unchanged compared with the baseline results in Table 4.

A.3 Lag order choice

In our baseline specification for the structural VAR model, we include lagged return and trade flow variables of the last five blocks. As a robustness check, we vary the number of lags in the structural VAR specification. Table A3 report the permanent price impacts when the number of lags is set to 10 and 20 respectively. It shows that the results are qualitatively the same as the baseline results.

Table A3. Permanent price impact of DEX trade flows with different priority fee levels: Alternative number of lags in the structural VAR specification. This table reports the permanent price impacts of the DEX trade flows with high, medium, and low priority fee levels. Permanent price impacts are defined as the cumulative impulse responses of the CEX return to DEX trade flow in the structural VAR model (see Equation 5). Upper bounds (UB) and lower bounds (LB) are obtained by considering all possible sequences of the recursive contemporaneous causality among the endogenous variables. The last column reports the difference between the lower bound of the permanent price impact of the high-fee DEX trade flow and the lower bound of the permanent price impact of the low-fee DEX trade flow. The estimation of the structural VAR is done for each pair-day and statistical inference is based on variations in the pair-day estimates. Row variables are response variables and column variables are shock variables. CEX return is in basis points. DEX trade flows are standardized and thus in their standard deviations. *, ** and *** indicate significance levels at 1%, 5% and 10% respectively.

(a) 10 lags of CEX return and DEX trade flows included in the structural VAR.

Variable	$x^{\text{LowFee-DEX}}$		$x^{\text{MidFee-DEX}}$		$x^{\text{HighFee-DEX}}$		$\Delta^{\text{HighFee - LowFee}}$
	LB	UB	LB	UB	LB	UB	LB - UB
Stable	0.01	0.02^{***}	0.0	0.01	−0.01	0.0	-0.03^{**}
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
NonStable	0.57^{***}	1.18^{***}	2.27^{***}	4.23^{***}	4.82^{***}	8.69^{***}	3.64^{***}
	(0.16)	(0.18)	(0.21)	(0.3)	(0.27)	(0.41)	(0.29)

(b) 20 lags of CEX return and DEX trade flows included in the structural VAR.

Variable	$x^{\text{LowFee-DEX}}$		$x^{\text{MidFee-DEX}}$		$x^{\text{HighFee-DEX}}$		$\Delta^{\text{HighFee - LowFee}}$
	LB	UB	LB	UB	LB	UB	LB - UB
Stable	0.03^{***}	0.04^{***}	0.01	0.01	−0.01	0.0	-0.05^{***}
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
NonStable	0.83^{***}	1.4^{***}	2.34^{***}	4.25^{***}	4.9^{***}	8.7^{***}	3.49^{***}
	(0.22)	(0.23)	(0.26)	(0.33)	(0.34)	(0.45)	(0.4)

References

Aoyagi, Jun and Yuki Ito (Mar. 2021). *Coexisting Exchange Platforms: Limit Order Books and Automated Market Makers*. SSRN Scholarly Paper. Rochester, NY.

Aquilina, Matteo, Eric Budish, and Peter O'Neill (Dec. 2021). "Quantifying the High-Frequency Trading "Arms Race"". In: *The Quarterly Journal of Economics* 137.1, pp. 493–564.

Avery, Christopher (Apr. 1998). "Strategic Jump Bidding in English Auctions". In: *Rev Econ Studies* 65.2, pp. 185–210.

Back, Kerry, C. Henry Cao, and Gregory A. Willard (Oct. 2000). "Imperfect Competition among Informed Traders". In: *The Journal of Finance* 55.5, pp. 2117–2155.

Barbon, Andrea and Angelo Ranaldo (Oct. 2022). *On The Quality Of Cryptocurrency Markets: Centralized Versus Decentralized Exchanges*. SSRN Scholarly Paper. Rochester, NY.

Barclay, Michael J., Terrence Hendershott, and D. Timothy McCormick (Dec. 2003). "Competition among Trading Venues: Information and Trading on Electronic Communications Networks". In: *The Journal of Finance* 58.6, pp. 2637–2665.

Bloomfield, Robert and Maureen O'Hara (Jan. 1999). "Market Transparency: Who Wins and Who Loses?" In: *Rev. Financ. Stud.* 12.1, pp. 5–35.

Boehmer, Ekkehart, Gideon Saar, and Lei Yu (2005). "Lifting the Veil: An Analysis of Pre-trade Transparency at the NYSE". In: *The Journal of Finance* 60.2, pp. 783–815.

Boehmer, Ekkehart and Juan (Julie) Wu (Feb. 2013). "Short Selling and the Price Discovery Process". In: *Rev. Financ. Stud.* 26.2, pp. 287–322.

Brogaard, Jonathan, Terrence Hendershott, and Ryan Riordan (Aug. 2014). "High-Frequency Trading and Price Discovery". In: *Rev. Financ. Stud.* 27.8, pp. 2267–2306.

— (Aug. 2019). "Price Discovery without Trading: Evidence from Limit Orders". In: *The Journal of Finance* 74.4, pp. 1621–1658.

Brunnermeier, Markus K. (2005). "Information Leakage and Market Efficiency". In: *Rev. Financ. Stud.* 18.2, pp. 417–457.

Brunnermeier, Markus K. and Lasse Heje Pedersen (Aug. 2005). “Predatory Trading”. In: *The Journal of Finance* 60.4, pp. 1825–1863.

Budish, Eric, Peter Cramton, and John Shim (Nov. 2015). “The High-Frequency Trading Arms Race: Frequent Batch Auctions as a Market Design Response*”. In: *The Quarterly Journal of Economics* 130.4, pp. 1547–1621.

Caldentey, René and Ennio Stacchetti (2010). “Insider Trading With a Random Deadline”. In: *Econometrica* 78.1, pp. 245–283.

Capponi, Agostino and Ruizhe Jia (Mar. 2021). *The Adoption of Blockchain-based Decentralized Exchanges*. SSRN Scholarly Paper. Rochester, NY.

Collin-Dufresne, Pierre and Vyacheslav Fos (Aug. 2015). “Do Prices Reveal the Presence of Informed Trading?: Do Prices Reveal the Presence of Informed Trading?” In: *The Journal of Finance* 70.4, pp. 1555–1582.

Daian, Philip et al. (May 2020). “Flash Boys 2.0: Frontrunning in Decentralized Exchanges, Miner Extractable Value, and Consensus Instability”. In: *2020 IEEE Symposium on Security and Privacy (SP)*, pp. 910–927.

Daniel, Kent D. and David A. Hirshleifer (July 1998). *A Theory of Costly Sequential Bidding*. SSRN Scholarly Paper. Rochester, NY.

Easley, David and Maureen O’Hara (Sept. 1987). “Price, Trade Size, and Information in Securities Markets”. In: *Journal of Financial Economics* 19.1, pp. 69–90.

Fleming, Michael J., Bruce Mizrach, and Giang Nguyen (Sept. 2018). “The Microstructure of a U.S. Treasury ECN: The BrokerTec Platform”. In: *Journal of Financial Markets* 40, pp. 2–22.

Foster, F. Douglas and S. Viswanathan (1996). “Strategic Trading When Agents Forecast the Forecasts of Others”. In: *The Journal of Finance* 51.4, pp. 1437–1478.

Glosten, Lawrence R. and Paul R. Milgrom (Mar. 1985). “Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders”. In: *Journal of Financial Economics* 14.1, pp. 71–100.

Hagströmer, Björn and Albert J. Menkveld (Dec. 2019). “Information Revelation in Decentralized Markets”. In: *The Journal of Finance* 74.6, pp. 2751–2787.

Hasbrouck, Joel (Mar. 1991a). “Measuring the Information Content of Stock Trades”. In: *The Journal of Finance* 46.1, pp. 179–207.

— (July 1991b). “The Summary Informativeness of Stock Trades: An Econometric Analysis”. In: *Rev. Financ. Stud.* 4.3, pp. 571–595.

— (Sept. 1995). “One Security, Many Markets: Determining the Contributions to Price Discovery”. In: *The Journal of Finance* 50.4, pp. 1175–1199.

Hasbrouck, Joel, Thomas J. Rivera, and Fahad Saleh (Aug. 2022). *The Need for Fees at a DEX: How Increases in Fees Can Increase DEX Trading Volume*. SSRN Scholarly Paper. Rochester, NY.

Holden, Craig W. and Avanidhar Subrahmanyam (1992). “Long-Lived Private Information and Imperfect Competition”. In: *The Journal of Finance* 47.1, pp. 247–270.

Kaniel, Ron and Hong Liu (2006). “So What Orders Do Informed Traders Use?” In: *The Journal of Business* 79.4, pp. 1867–1913.

Kyle, Albert S. (Nov. 1985). “Continuous Auctions and Insider Trading”. In: *Econometrica* 53.6, p. 1315.

Lehar, Alfred and Christine A. Parlour (Aug. 2021). *Decentralized Exchanges*. SSRN Scholarly Paper. Rochester, NY.

Lehar, Alfred, Christine A. Parlour, and Marius Zoican (Nov. 2022). *Liquidity Fragmentation on Decentralized Exchanges*. SSRN Scholarly Paper. Rochester, NY.

O’Hara, Maureen, Chen Yao, and Mao Ye (Oct. 2014). “What’s Not There: Odd Lots and Market Data: What’s Not There: Odd Lots and Market Data”. In: *The Journal of Finance* 69.5, pp. 2199–2236.

Park, Andreas (2021). “The Conceptual Flaws of Constant Product Automated Market Making”. In: *SSRN Journal*.

Yang, Liyan and Haoxiang Zhu (Apr. 2020). "Back-Running: Seeking and Hiding Fundamental Information in Order Flows*". In: *The Review of Financial Studies* 33.4. Ed. by Andrew Karolyi, pp. 1484–1533.

Zhu, Haoxiang (Mar. 2014). "Do Dark Pools Harm Price Discovery?" In: *Rev. Financ. Stud.* 27.3, pp. 747–789.