The Information Content of Blockchain Fees[1]

Agostino Capponi Ruizhe Jia Shihao Yu

Current version: September 26, 2022

[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 Bart Zhou Yueshen, Sean Foley, Tālis Putniņš and other participants at the Microstructure Exchange online seminar series (Asia-pacific edition) for helpful comments. We thank Prof. Arthur Gervais from Imperial College London for providing the Ethereum blockchain mempool data.

Abstract

Trading at decentralized exchanges (DEXs) requires traders to bid blockchain fees to determine the execution priority of their orders. We employ a structural vector-autoregressive (structural VAR) model to provide evidence that DEX trades with high blockchain fees not only reveal more private information, but also respond more to public price innovations on centralized exchanges (CEXs), contributing to price discovery. Using a unique dataset of Ethereum mempool orders, we further demonstrate that high blockchain fees do not result from traders competing with each other on private or public information. Rather, our analysis lends support to the hypothesis that they bid high blockchain fees to reduce the execution risk of their orders due to blockchain crowding.

1 Introduction

Decentralized exchanges (DEXs) are trading venues built on public blockchains, and they enable the trading of digital assets without the need for centralized intermediaries. Since their inception, DEXs have attracted a sizable trading volume and market share in crypto spot trading. As of May 2022, the market share of DEXs is roughly 15%, and the total trading volume on DEXs reaches at about $200 billion in May 2022.

DEXs collect and process orders through public blockchains. As a result, order execution on DEXs has two features distinct from that of centralized exchanges (CEXs) running a central limit order book (CLOB): first, orders are processed in discrete batches by blockchain validators; second, traders have to bid a blockchain fee (e.g., gas fee for Ethereum blockchain[1]) when submitting their orders. The blockchain fees determine the execution priority of orders because validators execute received orders based on blockchain fees, from high to low.

Compared with traditional CLOB markets where each order has two main components: price and trading size, an order on DEXs features an additional third component: blockchain fee. It is a natural question to ask whether the blockchain fee, which reflects traders' willingness to execute their trades timely, conveys any information? If so, does it reveal private information or simply respond to public information available? Are there any plausible economic channels that can explain the information content of blockchain fees?

To answer the above questions, we construct a data set that consists of executed trade data of Uniswap (the largest DEX) and Binance (the largest CEX) and tick-by-tick mempool data on the Ethereum blockchain. We focus on the eight most traded token pairs during our sample period between November 1, 2020, and March 24, 2021. With the executed trade data, we use a structural vector-autoregressive (structural VAR) model to investigate the information content of DEX trade flows with different levels of blockchain fee. The proprietary tick-by-tick mempool data tracks all orders submitted to the Ethereum network, as well as the gas fee bid by submitters. Thus, it allows

[1] In the rest of the paper, we use blockchain fee and gas fee interchangeably as we focus on the Ethereum blockchain.

us to identify potential competition among traders by examining whether they raise the blockchain fees to outbid their competitors and get their pending orders executed first.

Our main findings are summarized below. First of all, we find that DEX trades with high blockchain fees reveal private information, and more so compared with DEX trades with low blockchain fees: for token pairs involving a non-stable coin (e.g., Ethereum and Bitcoin), a shock to high-blockchain-fee DEX trade flow leads to a much larger permanent impact on market price than that to low-blockchain-fee trade flow. For example, for NonStable-NonStable token pairs (such as Bitcoin-Ethereum) where prices are most affected by fundamental value shocks, a positive shock of one standard deviation to the high-blockchain-fee DEX trade flow results in a permanent increase of about 6.5 basis points in market price. In contrast, a shock of the same size to low-blockchain-fee DEX trade flow permanently moves market price by only about 1.74 basis points. The only exception is DEX trades between two stablecoins (such as USDC-USDT). In such a case, DEX trades, regardless of the blockchain fee level, do not convey any private information, because the efficient exchange rate of two stablecoins is common knowledge and fixed at one.

In addition to being more privately informed, DEX trades with higher blockchain fees are found to be more responsive to market price innovations on CEXs. Our analysis shows that a shock to CEX price results in larger impulse responses of high-blockchain-fee DEX trade flow compared with low-blockchain-fee DEX trade flow. Take NonStable-NonStable token pairs as an example. A positive shock of one percentage to CEX price leads to an increase of about two standard deviations in the high-blockchain-fee DEX trade flow. For DEXs such as Uniswap which rely on an exogenous bonding curve to determine pricing schedule, the price cannot adjust through quote revision but has to be changed through trades. DEX trades with high blockchain fees help update the stale DEX price and make it more efficient, although they respond to public information already reflected in the quotes of CEX.

Why do privately informed traders and public information arbitrageurs bid high blockchain fees? A natural answer is that they compete to prioritize their order executions: they consecutively raise blockchain fees attached to their orders in order to outbid their competitors. Such competition arises either due to private information known to multiple informed traders or because of released public information seen to all. Surprisingly, our analysis of tick-by-tick mempool data shows that such a "trader competition" channel is unlikely to hold. First, we find that traders rarely increase their blockchain fee bids: only around 1% of executed orders see their gas fee revised up in the mempool during the same block. It suggests that the blockchain fee of virtually all orders is determined ex-ante when they are submitted, and they are not raised subsequently as a result of competition. Second, we further demonstrate that high-blockchain-fee DEX trades which are likely to be involved with trader competition are less privately informed and less responsive to price innovations on CEXs.

Instead, we propose an alternative explanation that is more consistent with our findings: for both traders trading on their short-lived private information and arbitrageurs trading on newly released public information, they bid high blockchain fees to avoid execution risks that they can not perfectly anticipate, such as a surge of other non-trading related transactions on the Ethereum blockchain. Such a "blockchain crowding" channel is further supported by our finding that high-blockchain-fee DEX trades are more likely to be profitable as, under competition, trading profits should be competed away.

Our paper relates to several strands of literature. First, past studies 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).[3] We contribute to the literature by studying the information content of blockchain fees, a featuring characteristic of trades executed on DEXs besides price and trade size. As informed traders on DEXs have to bid blockchain fees to get their orders executed, blockchain fees can

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

[3]A related literature is on the trading strategy of the informed trader(s) in various settings, e.g., a monopolistic informed trader (Kyle, 1985) or 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 the private information (Caldentey and Stacchetti, 2010) or short information horizon (Kaniel and Liu, 2006).

potentially serve as a new public signal revealing the private information contained in DEX trades.

Second, our paper further contributes to the nascent yet rapidly growing literature on decentralized exchanges. Few papers have discussed blockchain fees in various contexts. Park (2021) focuses on the unintended consequence of public blockchain order processing, i.e., all pending DEX transactions are subject to the risk of "sandwich attack". He mentions that in theory, liquidity demanders are able to prevent frontrunning by choosing a very high blockchain fee. Capponi and Jia (2021) investigates how the choice of DEX pricing rules affects welfare and liquidity provision incentives. They show that arbitrageurs can always outbid liquidity providers, in blockchain fees auctions, to exploit the price discrepancy between CEX and DEX, which reduces incentives for liquidity provision. Barbon and Ranaldo (2021) compares the price efficiency of CEX and DEX. They argue that the low price efficiency of Uniswap can be partially attributed to high blockchain fees which are fixed costs for traders. Parlour and Lehar (2021) and Aoyagi and Ito (2021) abstract away from blockchain fees and focus more on the novel liquidity provision structure of AMMs. A significant contribution of our work relative to the existing literature is highlighting how blockchain fees convey both private and public information. In DEX, market price can only be revised through trading, and it is trades with high blockchains fees that help DEX keep track of efficient prices.

More broadly, our paper relates to the literature on the "arm race" among traders on public information. Previous papers analyze the issue in the context of traditional central limit order book (CLOB) markets and they show that its impact on market liquidity ultimately depends on which trader type is faster: market liquidity worsens if arbitrageurs become faster (Biais, Foucault, and Moinas, 2015; Budish, Cramton, and Shim, 2015; Foucault, Hombert, and Roşu, 2016), while it improves if market makers become faster (Hoffmann, 2014; Jovanovic and Menkveld, 2016). The emphasis on speed lies in the continuous trading nature of the traditional CLOB market so that the fastest trader will win the race and seize the trading opportunity. In contrast, trading on DEXs runs in discrete time and thus speed is no longer a determinant factor in winning the races. Instead, traders compete by bidding blockchain fees, which are public signals to all, in order to get their orders executed first. We contribute to the literature by focusing on such a unique form of public information race on DEXs and analyzing its impact on price efficiency.

The remainder of the paper proceeds as follows: Section 2 introduces the inner workings of the DEXs and their unique features. Section 3 provides an overview of the dataset. Section 4 details the empirical methodology. Section 5 shows the results. Section 6 concludes.

2 Institutional background

In this section, we briefly introduce DEXs and characteristics of trade execution on DEXs, with a focus on blockchain fees.

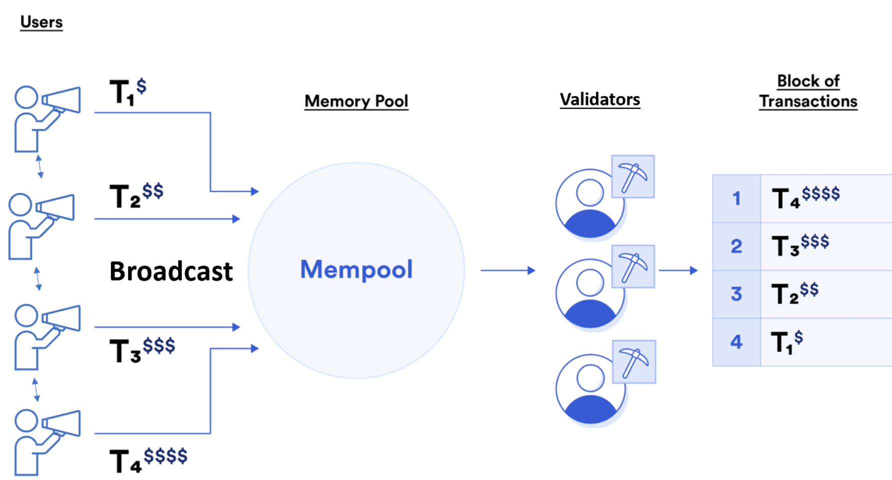
2.1 DEXs

DEXs are blockchain-based smart contracts. As of July 2022, around 15% crypto spot tradings occur on DEXs, and Uniswap is the largest DEX which accounts for more than half of the DEX trading volume (see, Block, 2022). The remaining 85% of crypto spot tradings are executed on centralized exchanges, and the largest CEX is Binance which takes up more than 75% of the CEX market share. Different from CEX which utilizes limit order books, most DEXs are in the form of Automated Market Makers (AMM). In AMMs, liquidity providers are depositors, and the pricing schedules are determined by some function pre-coded in the smart contract. A more detailed description of AMMs can be found at Capponi and Jia, 2021.

2.2 Trade execution on DEXs and blockchain fees

DEXs rely on blockchain networks (typically Ethereum) to receive, process, and execute orders. To execute a trade at a DEX, a trader typically has to first broadcast the transaction details in the blockchain network and bid a blockchain fee for her order. The transaction details reveal trade information even before the trade is executed, such as the address of the DEX and the intended trade size and price. If the transaction is received by a validator, then it will be pending in the mempool of that validator. Blockchain fees determine execution priority. If a validator is chosen to validate the next block, then she will execute the transaction in her mempool in the decreasing order of blockchain fee bid. Since blocks in a blockchain are produced in discrete time, DEX orders are also processed discretely in batches. Because each block has a maximum capacity, transactions with too low blockchain fees will not be included in the block, or they need to wait for a long time before being executed. Figure 1 depicts the order execution mechanism of DEX.

Figure 1. Trading mechanism on DEXs.



Every transaction broadcasted by a trader is associated with a number called "nonce". Each nonce can only be used once and must be used in increasing order. In other words, the first order of a trader has a nonce "0", her second order has a nonce "1", and her N^{th} order has a nonce "N". When a trader wants to modify her pending order or increase the blockchain fees bid, she has to broadcast a new transaction with the same nonce as the pending one and increase the blockchain fees bid. A validator who receives this new transaction will not execute the old one because a validator will prioritize transactions with higher blockchain fees, and each nonce can only be used once by one trader. For the same reason, sending a new transaction with decreased blockchain fee will never work, because the old transaction with higher blockchain fees will still be executed first.

In the Ethereum blockchain, the blockchain fees are referred to as "gas fees". The execution of each transaction costs a fixed amount of computation resource, measured by "gas used". When bidding blockchain fees, Ethereum users specify "gas price", i.e., how much they are willing to pay per unit of gas. The total gas fees paid by users as blockchain fees are equal to the gas used multiplied by the gas price bid. Ethereum validators sort and execute transactions in mempools in the decreasing order of gas price.

3 Data

In this section, we describe our dataset. In Section 3.1, we introduce our executed trade data, and our tick-by-tick mempool order data in Section 3.2.

3.1 Executed trade data

Our dataset covers trades executed on Binance and Uniswap for eight most actively traded token pairs during our sample period, USDC-USDT, DAI-USDT, ETH-USDT, ETH-USDC, ETH-DAI, WBTC-ETH, LINK-ETH, AAVE-ETH, between November 1, 2020 and March 24, 2021. The eight token pairs fall into three types respectively: "Stable-Stable" pairs, "Stable-NonStable" pairs, and "NonStable-NonStable" pairs. "Stable-Stable" pairs include two stablecoins pegged to one US Dollar (USDC-USDT, ETH-USDT). "Stable-NonStable" pairs include one non-stable token, which is not pegged to any fiat currency, and one stablecoin (ETH-USDT, ETH-USDC, ETH-DAI). "NonStable-NonStable" includes two non-stable tokens (WBTC-ETH, LINK-ETH, AAVE-ETH). Binance trades are publicly available and collected from its website[4], while Uniswap trades are collected through a proprietary node. Below we provide a detailed description of each of the two datasets.

Uniswap trades Each Uniswap trade contains the hash, the address of the trader, the timestamp of the block in which the trade is included (to the precision of second), the number of the block in which the trade is included, the execution position of the trade in that block, gas price, gas used,

[4] See https://data.binance.vision/?prefix=data/spot/monthly/.

trade direction indicating whether it is a buy trade or sell trade in terms of the base token[5], the amount of tokens in the liquidity pool before and after the trades, and the amount of tokens that the trader deposits in and takes out from the liquidity pool.

Based on the amount of tokens in the liquidity pool before a trade, we can compute the prevailing "midquote" just before the trade. For example, if there are y amount of token Y and x amount of X before the trade, the prevailing midquote is simply the ratio of the amount of two tokens in the pool, x/y. Note that on AMMs like Uniswap, there are no quotes. Thus, "midquote" is defined in a broader sense as the hypothetical price for an infinitesimal trade. In addition, based on the amount of tokens that the trader deposits in and takes out from the liquidity pool, we can compute the transaction price of the trade. For example, if Δy amount of token Y is swapped for x amount of token X, then the transaction price is simply the ratio of the amount of two tokens swapped, i.e., $\Delta x/\Delta y$. Last, we use the amount of the base token swapped as the transaction size of the trade, that is, $|\Delta y|$.

Binance trades Each Binance trade record includes a unique identifier for the trade, the timestamp (to the precision of 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 a trade: if the buyer uses a market order, then it is classified as a buy trade; otherwise, it is a sell trade.

3.2 Mempool order data

In addition to executed trade data, we obtain proprietary tick-by-tick mempool order data through a proprietary node.[6] The proprietary node records every new order submission it receives in its mempool, which either ends up with being executed or left unexecuted. Each order has the follow-

[5] Note that 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, for the token pair ETH-USDT, ETH is the base token and USDT is the quote token.

[6] The proprietary node and the data are both provided by Prof. Arthur Gervais from Imperial College.

ing information: the hash, the timestamp when the order is received by the node (to the precision of millisecond), address of the trader, nonce, gas price, and the gas limit (i.e. the maximum gas allowed to be used). Our mempool data covers orders received in the mempool for the eight token pairs in the period of February 7, 2021 through March 24, 2021. With the mempool data, we can track the complete history of order revisions, if they occur, before the final order is executed and turns into a trade. Thus, we are able to see whether the trader bid up the gas price attached to her order in order to get it executed.

3.3 Summary statistics of executed trades

To have a high-level picture of the liquidity of our sample token pairs, in Table 1, we report summary statistics of their daily trading volume and number of trades on Uniswap and Binance. There are several notable observations. First, trading in all eight token pairs is fairly active. For instance, the average daily number of trades (daily trading volume) on Uniswap is 869 ($\approx$ 3.8 million USD), 8601 ($\approx$ 80 million USD) and 1304 ($\approx$ 34 million USD) for USDC-USDT, ETH-USDT and WBTC-ETH respectively. Second,

Second, trading activity on Uniswap versus Binance differs significantly across token pairs. For USDC-USDT, DAI-USDT, and ETH-USDT, trading is much more active on Binance than Uniswap. For example, the average daily trading volume on Binance is about 107 million USDT and about 1.24 million ETH for USDC-USDT and ETH-USDT respectively, more than an order of magnitude larger than that on Uniswap. In contrast, for the rest of token pairs, ETH-USDC, ETH-DAI, WBTC-ETH, LINK-ETH and AAVE-ETH, trading is more active on Uniswap than Binance. Take WBTC-ETH as an example. Its average daily trading volume is about 31604 ETH on Uniswap, much larger than 1684 ETH on Binance.

Then, in Table 2, we report summary statistics of the execution price, gas price and trade size of Uniswap trades for our eight sample token pairs. First, the average trade size of a Uniswap trade is fairly large, and is about 4360 USD, 8.42 ETH ($\approx$ 9268 USD), and 24.15 ETH ($\approx$ 26,177

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.

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

Pair		N	Mean	SD	Min	Med	Max
USDC-USDT	TradingVolume-Uniswap	145	3804	2257	364	3885	11641
	TradeCount-Uniswap	145	869	351	415	773	3085
	TradingVolume-Binance	145	107896	69910	16396	102837	398522
	TradeCount-Binance	145	50266	21949	14182	47198	135409
DAI-USDT	TradingVolume-Uniswap	145	1258	1152	54	960	5830
	TradeCount-Uniswap	145	494	368	161	352	2068
	TradingVolume-Binance	145	12086	10135	694	9946	77831
	TradeCount-Binance	145	8279	6782	1211	7071	58558

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

Pair		N	Mean	SD	Min	Med	Max
ETH-USDT	TradingVolume-Uniswap	145	72574	38805	36596	61494	263356
	TradeCount-Uniswap	145	8601	1452	6311	8308	16419
	TradingVolume-Binance	145	1243663	635988	438952	1052452	4245010
	TradeCount-Binance	145	948849	479112	181369	915584	2577496
ETH-USDC	TradingVolume-Uniswap	145	79559	39893	32878	70621	302051
	TradeCount-Uniswap	145	7595	1405	4681	7549	13851
	TradingVolume-Binance	145	27927	20587	5808	21350	142110
	TradeCount-Binance	145	19292	13222	3158	16554	80061
ETH-DAI	TradingVolume-Uniswap	145	54327	69493	9992	39382	746637
	TradeCount-Uniswap	145	3834	1353	1591	3547	7786
	TradingVolume-Binance	145	8737	8823	245	6529	51489
	TradeCount-Binance	145	17390	17435	285	14180	103522

USD) for USDC-USDT, ETH-USDT and WBTC-ETH respectively. Second, 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 131.37 Gwei (1Gwei = 10^{-9}ETH), its standard deviation is 474.40, which is more than triple the size of the mean. Such a large variation can result from either the changing overall crowding level of the Ethereum network or traders' gas fee bidding behaviors.

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

Pair		N	Mean	SD	Min	Med	Max
WBTC-ETH	TradingVolume-Uniswap	145	31604	21973	6278	26293	139189
	TradeCount-Uniswap	145	1304	531	646	1141	3338
	TradingVolume-Binance	145	1684	1688	18	1261	9984
	TradeCount-Binance	145	6260	6397	92	4503	35191
LINK-ETH	TradingVolume-Uniswap	145	9348	5870	1949	7958	42520
	TradeCount-Uniswap	145	920	366	367	873	2682
	TradingVolume-Binance	145	4463	2489	1071	3975	13598
	TradeCount-Binance	145	11630	7337	2200	11226	43491
AAVE-ETH	TradingVolume-Uniswap	145	6442	4418	819	5595	29936
	TradeCount-Uniswap	145	536	277	136	502	1514
	TradingVolume-Binance	145	1876	1381	348	1546	10143
	TradeCount-Binance	145	6272	4708	1075	5322	36964

4 Methodology

To examine whether DEX trades with a higher blockchain fee are more informative, we follow Hasbrouck (1991a) and estimate a structural vector-autoregressive (structural VAR) model. In the structural VAR model, we include CEX return and DEX trade flows with different blockchain fee levels as endogenous variables. Thus we can compute the cumulative impulse response of return to a trade flow variable, i.e., its permanent price impact, which is regarded as a measure of its private information content. Barclay, Hendershott, and McCormick (2003) and O'Hara, Yao, and Ye (2014) apply the same approach to examine the informativeness of odd-lot trades versus round-lot trades and ECN trades versus market-maker trades respectively.

4.1 Baseline structural VAR specification

A general structural VAR model can be specified as follows:

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

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

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

Pair	Variable	N	Mean	SD	1%	10%	Median	90%	99%
USDC-USDT	TxPrice	123086	1.00	0.00	0.99	1.00	1.00	1.00	1.01
	GasPrice	123086	105.20	93.92	16.00	33.00	86.00	192.00	420.00
	TxSize	123086	4.36	11.57	0.01	0.12	1.30	10.00	50.00
DAI-USDT	TxPrice	68626	1.00	0.00	0.99	1.00	1.00	1.01	1.01
	GasPrice	68626	97.13	116.51	16.94	33.10	79.00	171.00	400.00
	TxSize	68626	2.54	5.50	0.00	0.07	0.85	6.01	27.65

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

Pair	Variable	N	Mean	SD	1%	10%	Median	90%	99%
ETH-USDT	TxPrice	1242492	1102.82	534.86	384.03	468.39	1156.36	1808.08	1949.73
	GasPrice	1242492	115.09	186.48	15.51	30.00	93.00	202.00	527.50
	TxSize	1242492	8.42	106.26	0.01	0.15	1.45	13.70	122.23
ETH-USDC	TxPrice	1096573	1143.46	513.77	385.26	484.57	1228.99	1803.08	1949.10
	GasPrice	1096573	121.27	215.92	16.00	34.21	99.00	210.00	551.00
	TxSize	1096573	10.45	38.41	0.02	0.18	1.75	21.90	140.44
ETH-DAI	TxPrice	551912	1012.30	482.78	380.62	467.07	902.06	1760.85	1931.49
	GasPrice	551912	111.90	308.95	15.70	31.00	82.50	195.80	554.00
	TxSize	551912	12.91	109.34	0.01	0.10	1.43	33.53	159.55

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

Pair	Variable	N	Mean	SD	1%	10%	Median	90%	99%
WBTC-ETH	TxPrice	185927	30.85	4.70	22.67	24.33	31.61	36.96	41.89
	GasPrice	185927	131.37	474.40	16.10	34.43	102.00	230.00	631.83
	TxSize	185927	24.15	73.05	0.01	0.21	3.60	63.23	288.72
LINK-ETH	TxPrice	130257	0.02	0.00	0.01	0.01	0.02	0.03	0.03
	GasPrice	130257	121.54	257.07	15.10	31.00	88.80	213.00	686.41
	TxSize	130257	10.13	23.94	0.02	0.19	2.62	27.53	86.58
AAVE-ETH	TxPrice	74587	0.17	0.06	0.07	0.11	0.15	0.26	0.30
	GasPrice	74587	115.72	180.75	15.10	29.00	88.80	205.81	600.00
	TxSize	74587	11.98	19.83	0.03	0.20	4.86	29.50	86.60

where $\Phi_1 \dots \Phi_p$ are standard system matrices of the VAR model. ε_t is the vector of structural innovations and satisfies the following conditions: $E(\varepsilon_t) = 0$; $E(\varepsilon_t\varepsilon_t') = \Sigma_\varepsilon$; $E(\varepsilon_t\varepsilon_s') = 0$ for $s \neq t$. What's more important is y_t, the vector endogenous variable vector, and A, the structural matrix capturing the contemporaneous correlations between the endogenous variables. Including which endogenous variables in y_t and assuming what kind of contemporaneous correlations between the endogenous variables is up to the choice of the researchers, which normally is based on some economic reasoning. We will detail our specifications below and provide our rationale for the choice.

Our baseline specification for the structural VAR model is as follows:

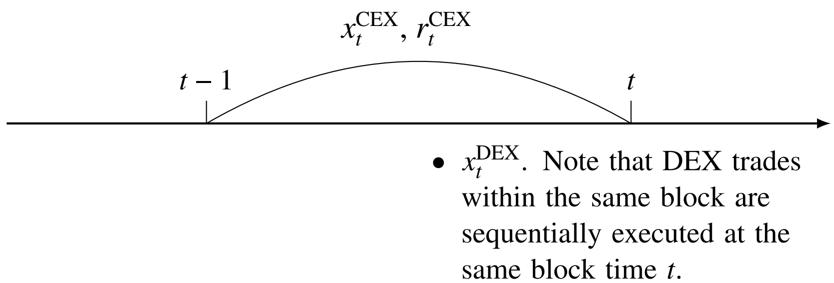
$$y_t = \begin{pmatrix} r_t^{\text{CEX}} & x_t^{\text{LowGas-DEX}} & x_t^{\text{MidGas-DEX}} & x_t^{\text{HighGas-DEX}} \end{pmatrix}', \qquad A = \begin{pmatrix} 1 & a_{12} & a_{13} & a_{14} \\ 0 & 1 & 0 & 0 \\ 0 & a_{32} & 1 & 0 \\ 0 & a_{42} & a_{43} & 1 \end{pmatrix} \tag{2}$$

where t indexes block time. r_t^{CEX} is Binance return from block time $t-1$ to t. Note that we do not have quote updates, but only trades, from Binance. So the return is calculated based on trade prices, not midquotes. $x_t^{\text{LowGas-DEX}}$, $x_t^{\text{MidGas-DEX}}$ and $x_t^{\text{HighGas-DEX}}$ are Uniswap trade flows in block t with low, mid and high gas price levels respectively.

Timestamp convention As Binance runs a continuous central limit order book while Uniswap executes trades in batches based on block time. It is important to specify the timestamp convention we use for Binance return and Uniswap trade flows. In Figure 2, we provide a visual illustration. Specifically, r_t^{CEX} is the log difference between the price of the last Binance trade before block time $t-1$ and that of the last Binance trade before block time t. Then all Uniswap trade flows, $x_t^{\text{LowGas-DEX}}$, $x_t^{\text{MidGas-DEX}}$ and $x_t^{\text{HighGas-DEX}}$, are computed based on trades executed in batch at block time t. Next, we will detail our timestamp convention and our gas fee level classification scheme.

Gas fee level classification In the structural VAR specification above, we include DEX trade flows with three different levels of gas price. Now, we introduce our classification scheme. For any

Figure 2. 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 signed trade flow on Binance summed over the time interval between block time $t-1$ and t. x_t^{DEX} is the signed trade flow on Uniswap at block time t.



classification, finding the right benchmark gas fee is crucial. We adopt a rolling-window approach to calculate the benchmark gas price. Specifically, to classify trades in the current block t, we first sort them together with all trades within the last 20 non-empty blocks[7], i.e., block $t-20$ to block $t-1$ based on their gas fee in descending order. Then trades in block t which fall in the top quartile (i.e., above 75% quantile) is labeled as high-gas-price trades, $x_t^{\text{HighGas-DEX}}$; Trades which fall in the bottom quartile (i.e., below 25% quantile) are labeled as low-gas-price trades, $x_t^{\text{LowGas-DEX}}$; All other trades are labeled as mid-gas-price trades (i.e., between 25% and 75% quantile), $x_t^{\text{MidGas-DEX}}$. Thus for each block, we construct three DEX trade flows. For some blocks, we might only have observations for only one or two types of trades, then trade flows for the remaining type(s) are simply zero.

In Table 3 we report summary statistics of gas price and trade size of trades within each of the three gas price category classified above. By construction, the average gas price of trades within the high-gas group should be higher than that of trades within the mid- and low-gas groups. Take ETH-USDT as an example. The average gas price of a trade classified as high-gas trade is 170.98 Gwei (1 Gwei = 10^{-9} ETH), which is almost double the size of that of a low-gas trade. One notable point is that the average trade size of a trade in the high-gas category is much higher than that of a trade in the mid- and low-gas categories. For example, the average trade size of a ETH-USDT

[7] As a robustness check, we redo the classification based on trades within the last five, ten blocks instead and all results are qualitatively the same. In Appendix A.1, we report the structural VAR results based on alternative window lengths.

trade within the high-gas category is 17.72 ETH, which is more than five times that within the low-gas category.

Table 3. Summary statistics of trade characteristics by gas fee level. This table reports, for each token pair, summary statistics of the gas price (GasPrice) and transaction size (TxSize) by gas fee level. Gas price is denominated in Gwei, which equals 10^{-9} ETH.

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

Pair	Variable	GasLevel	N	Mean	SD	Median
USDC-USDT	GasPrice	LowGas	29932	82.00	57.73	68.00
		MidGas	60522	96.14	68.66	80.00
		HighGas	32632	143.29	139.21	117.00
	TxSize	LowGas	29932	3.00	8.12	0.98
		MidGas	60522	3.95	11.05	1.14
		HighGas	32632	6.37	14.57	2.20
DAI-USDT	GasPrice	LowGas	17056	73.57	50.70	61.00
		MidGas	32797	90.15	63.40	76.00
		HighGas	18773	130.72	196.34	104.00
	TxSize	LowGas	17056	1.90	4.28	0.55
		MidGas	32797	2.31	5.06	0.78
		HighGas	18773	3.52	6.90	1.19

Resolution of the contemporaneous correlations Last, the structural matrix A is specified such that we only allow the following contemporaneous relations: (1) all three DEX trade flow variables cause CEX return but not vice versa, which is normally assumed in market microstructure; (2) low-gas DEX trade flow causes mid-gas and high-gas DEX trade flows; (3) mid-gas DEX trade flow causes high-gas DEX trade flow. We impose such a recursive structure so that we seek to obtain a lower bound of the price impact of high-gas DEX trade flow. Similarly, O'Hara, Yao, and Ye (2014) assume odd-lot trades cause round-lot and mixed-lot trades in one of their specification to obtain a lower bound of the price impact of odd-lot trades.

Permanent price impact of trade flows After we have estimated the structural VAR model, we can easily obtain the vector moving average (VMA) representation to compute the impulse

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

Pair	Variable	GasLevel	N	Mean	SD	Median
ETH-USDT	GasPrice	LowGas	266642	88.21	67.25	74.00
		MidGas	676442	100.95	76.10	85.00
		HighGas	299408	170.99	350.67	128.00
	TxSize	LowGas	266642	3.54	19.04	0.83
		MidGas	676442	6.22	135.13	1.24
		HighGas	299408	17.72	71.87	3.22
ETH-USDC	GasPrice	LowGas	235734	92.43	70.59	80.00
		MidGas	594722	106.48	80.76	91.00
		HighGas	266117	179.85	410.45	135.00
	TxSize	LowGas	235734	4.30	18.83	0.94
		MidGas	594722	7.89	32.35	1.54
		HighGas	266117	21.59	57.04	4.19
ETH-DAI	GasPrice	LowGas	123305	82.09	68.38	65.00
		MidGas	292993	96.50	83.46	76.00
		HighGas	135614	172.27	603.48	115.23
	TxSize	LowGas	123305	5.36	38.55	0.75
		MidGas	292993	10.42	45.81	1.20
		HighGas	135614	25.14	206.29	4.48

responses of return and trade variables to shocks in the structural innovations:

$$y_t = \Theta(L)\varepsilon_t = \Theta_0\varepsilon_t + \Theta_1\varepsilon_{t-1} + \Theta_2\varepsilon_{t-2} + \cdots \tag{3}$$

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

$$\mathrm{PPI}_k = \frac{\sum_{j=0}^{\infty} \partial r_{t+j}}{\partial \varepsilon_{k,t}} = [\Theta(1)]_{1,k} \tag{4}$$

where $[\Theta(1)]_{1,k}$ denotes the $(1,k)$-*th* element of $\Theta(1)$, the impulse response of return to trade flow variable k.

(c) NonStable-NonStable token pairs. Transaction size is in ETH.

Pair	Variable	GasLevel	N	Mean	SD	Median
WBTC-ETH	GasPrice	LowGas	43947	93.62	73.42	81.00
		MidGas	94110	113.76	89.82	97.00
		HighGas	47870	200.66	920.11	144.00
	TxSize	LowGas	43947	10.86	42.08	1.58
		MidGas	94110	21.21	71.58	3.15
		HighGas	47870	42.14	92.33	14.72
LINK-ETH	GasPrice	LowGas	31427	81.82	68.14	64.64
		MidGas	65276	100.83	86.63	80.00
		HighGas	33554	199.03	478.83	130.00
	TxSize	LowGas	31427	4.44	10.98	1.18
		MidGas	65276	9.38	26.37	2.22
		HighGas	33554	16.91	26.12	10.39
AAVE-ETH	GasPrice	LowGas	18186	75.26	59.33	59.00
		MidGas	37196	97.12	78.39	78.00
		HighGas	19205	190.04	322.35	127.86
	TxSize	LowGas	18186	5.58	13.14	1.40
		MidGas	37196	11.74	21.37	4.27
		HighGas	19205	18.49	19.95	16.91

Information share of trade flows In addition to permanent price impact, we can compute the so-called information shares of the trade flow variables via the approach of random walk decomposition (See Hasbrouck, 1991b, for detailed proofs). In words, the information share measure weighs the permanent price impact of a trade flow variable $[\Theta(1)]_{1,k}$ by its own structural innovation variance, $\sigma^2_{\varepsilon_k}$. So if two trade flow variables have the same permanent price impact, the information share of the one which arrives at the market more frequently will have a larger information share. Mathematically, the information share (IS) of trade flow variable k to price discovery is computed as:

$$\mathrm{IS}_k = \frac{[\Theta(1)]^2_{1,k}\sigma^2_{\varepsilon_k}}{\sum_k [\Theta(1)]^2_{1,k}\sigma^2_{\varepsilon_k}} \tag{5}$$

Implementation details We implement the structural VAR estimation in the following ways: (1) the model is estimated at block-by-block frequency, although the gas price level classification is based on a 20-block rolling window; (2) we set the number of lags in the structural VAR model to

5[8]; (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. So the impulse responses reported below should be interpreted as permanent price impacts in basis points per standard deviation increase in the trade flow.

4.2 An alternative structural VAR specification for robustness

In addition to Uniswap, traders can execute their trades on CEXs such as Binance as well. To control for cross-venue arbitrage trades, we include Binance trade flow in the endogenous variable vector and thus have the following alternative specification:

$$y_t = \begin{pmatrix} r_t^{\text{CEX}} & x_t^{\text{CEX}} & x_t^{\text{LowGas-DEX}} & x_t^{\text{MidGas-DEX}} & x_t^{\text{HighGas-DEX}} \end{pmatrix}', \quad A = \begin{pmatrix} 1 & a_{12} & a_{13} & a_{14} & a_{15} \\ 0 & 1 & 0 & 0 & 0 \\ 0 & a_{32} & 1 & 0 & 0 \\ 0 & a_{42} & a_{43} & 1 & 0 \\ 0 & a_{52} & a_{53} & a_{54} & 1 \end{pmatrix} \tag{6}$$

x_t^{CEX} is the signed trade flow on Binance aggregated between block time $t-1$ and t. r_t^{CEX}, $x_t^{\text{LowGas-DEX}}$, $x_t^{\text{MidGas-DEX}}$ and $x_t^{\text{HighGas-DEX}}$ are all defined above and represent CEX return and Uniswap trade flows with low, mid and high gas fee levels respectively.

Similarly, we specify the structural matrix A in such a way that we only allow CEX trade flow to contemporaneously affect DEX trade flows, but not vice versa. In addition, as in the first specification, we further assume the following contemporaneous relations: (1) low-gas DEX trade flow causes mid-gas and high-gas DEX trade flows; (3) mid-gas DEX trade flow causes high-gas DEX trade flow. Again, we impose such a recursive structure so that we seek to obtain a lower bound of the price impact of high-gas DEX trade flow. In terms of economics, we impose such

[8]In Appendix A.2, we change the number of lags included in the structural VAR model to 10 and 20, and show that estimation results are qualitatively the same.

structural restrictions in order to control for informed traders splitting their trades on both CEXs and DEXs. The idea is intuitive: assuming traders trade their private information on both CEXs and DEXs, we should expect CEX and DEX trade flows to be highly correlated. So after the CEX trade flow is controlled, if CEX prices still respond to DEX trade flows, it must be the case that the DEX trade flows contain other private information.

5 Results

5.1 Summary statistics of return and trade flows

Table 4. Summary statistics of return and 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 of (See Equation 2 and Equation 6). 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. $x_t^{\text{LowGas-DEX}}$, $x_t^{\text{MidGas-DEX}}$ and $x_t^{\text{HighGas-DEX}}$ are Uniswap trade flows consisting of trades from the low-, mid- and high-gas category in block t. Both r_t^{CEX} and r_t^{DEX} are in basis point.

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

		N	Mean	SD	Min	50%	Max
USDC-USDT	r_t^{CEX}	109436	0.00	0.97	-93.50	0.00	78.48
	x_t^{CEX}	109436	-1.61	146.78	-4994.91	0.00	7305.06
	x_t^{DEX}	109436	0.01	10.93	-403.21	-0.01	500.00
	$x_t^{\text{LowGas-DEX}}$	109436	0.03	3.97	-145.13	0.00	430.50
	$x_t^{\text{MidGas-DEX}}$	109436	0.06	6.88	-206.70	0.00	500.00
	$x_t^{\text{HighGas-DEX}}$	109436	-0.08	7.51	-403.21	0.00	348.85
DAI-USDT	r_t^{CEX}	63442	-0.00	2.53	-89.16	0.00	93.20
	x_t^{CEX}	63442	0.21	46.93	-1162.58	0.00	1009.53
	x_t^{DEX}	63442	0.01	5.78	-142.20	-0.00	141.16
	$x_t^{\text{LowGas-DEX}}$	63442	0.01	2.23	-61.21	0.00	58.82
	$x_t^{\text{MidGas-DEX}}$	63442	0.01	3.60	-81.14	0.00	68.87
	$x_t^{\text{HighGas-DEX}}$	63442	-0.02	3.87	-142.20	0.00	94.93

Before discussing the estimation results from the structural VAR model, we first report, for each token pair, summary statistics of the return and trade flow variables in Table 4. There are several notable observations. First, as expected, returns of NonStable-NonStable token pairs are most

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

		N	Mean	SD	Min	50%	Max
ETH-USDT	r_t^{CEX}	636959	0.02	9.52	-551.79	0.00	362.82
	x_t^{CEX}	636959	-0.57	201.40	-7800.81	0.04	13897.08
	x_t^{DEX}	636959	0.17	45.18	-4304.51	0.02	5398.51
	$x_t^{\text{LowGas-DEX}}$	636959	-0.03	10.23	-2935.73	0.00	1241.70
	$x_t^{\text{MidGas-DEX}}$	636959	-0.08	23.93	-2579.05	0.00	2147.57
	$x_t^{\text{HighGas-DEX}}$	636959	0.28	36.87	-4304.51	0.00	5398.51
ETH-USDC	r_t^{CEX}	593951	0.02	10.42	-456.54	0.00	353.11
	x_t^{CEX}	593951	-0.08	23.98	-3693.72	0.00	2345.59
	x_t^{DEX}	593951	0.20	49.79	-7426.57	-0.08	8305.92
	$x_t^{\text{LowGas-DEX}}$	593951	-0.12	10.45	-1558.00	0.00	1086.42
	$x_t^{\text{MidGas-DEX}}$	593951	-0.14	27.15	-2612.99	0.00	3005.27
	$x_t^{\text{HighGas-DEX}}$	593951	0.46	39.81	-7426.57	0.00	8305.92
ETH-DAI	r_t^{CEX}	381166	0.04	11.98	-529.28	0.00	949.91
	x_t^{CEX}	381166	-0.03	7.10	-850.26	0.00	679.90
	x_t^{DEX}	381166	0.25	51.99	-9281.89	-0.00	3300.12
	$x_t^{\text{LowGas-DEX}}$	381166	-0.02	15.78	-6140.64	0.00	2165.82
	$x_t^{\text{MidGas-DEX}}$	381166	0.21	28.50	-1718.83	0.00	2841.28
	$x_t^{\text{HighGas-DEX}}$	381166	0.06	42.20	-11038.04	0.00	3196.28

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

		N	Mean	SD	Min	50%	Max
WBTC-ETH	r_t^{CEX}	156096	0.00	12.89	-360.42	0.00	300.74
	x_t^{CEX}	156096	0.07	8.90	-502.95	0.00	1991.97
	x_t^{DEX}	156096	0.02	77.20	-3465.19	0.18	4369.22
	$x_t^{\text{LowGas-DEX}}$	156096	0.05	18.64	-1690.72	0.00	1250.70
	$x_t^{\text{MidGas-DEX}}$	156096	0.23	48.26	-2954.34	0.00	4369.22
	$x_t^{\text{HighGas-DEX}}$	156096	-0.27	56.60	-3465.19	0.00	2344.65
LINK-ETH	r_t^{CEX}	113044	-0.04	22.72	-639.83	0.00	828.08
	x_t^{CEX}	113044	-0.46	17.64	-2047.56	0.00	603.75
	x_t^{DEX}	113044	-0.04	22.19	-1187.08	-0.06	661.38
	$x_t^{\text{LowGas-DEX}}$	113044	-0.04	5.44	-202.07	0.00	255.32
	$x_t^{\text{MidGas-DEX}}$	113044	-0.10	14.54	-1187.08	0.00	652.36
	$x_t^{\text{HighGas-DEX}}$	113044	0.09	15.63	-432.35	0.00	661.38
AAVE-ETH	r_t^{CEX}	66875	0.17	38.81	-429.98	0.00	420.53
	x_t^{CEX}	66875	-0.28	10.68	-676.27	0.00	493.72
	x_t^{DEX}	66875	0.12	18.88	-417.79	0.04	509.38
	$x_t^{\text{LowGas-DEX}}$	66875	0.05	5.43	-150.28	0.00	329.71
	$x_t^{\text{MidGas-DEX}}$	66875	0.03	12.84	-417.79	0.00	509.38
	$x_t^{\text{HighGas-DEX}}$	66875	0.05	12.84	-221.06	0.00	374.95

volatile, followed by NonStable-Stable token pairs, and Stable-Stable token pairs. For instance, per-block-time (≈15 seconds) standard deviation of Binance return, r_t^{CEX}, is about 0.97, 9.49, and 12.88 basis points for USDC-USDT, ETH-USDT and WBTC-ETH respectively. The results are not surprising as both NonStable-Stable and NonStable-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-Stable token pairs are only affected by short-term liquidity shocks as both of their tokens 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 USDC-USDT, DAI-USDT and ETH-USDT, the magnitude of trade flow is much more larger on Binance than Uniswap. For example, standard deviation of per-block-time trade flow for USDC-USDT on Binance is about 150 thousand USD, more than 10 times larger than that of only about 10.93 thousand USD on Uniswap. In contrast, for the rest of token pairs, ETH-USDC, ETH-DAI, WBTC-ETH, LINK-ETH and AAVE-ETH, absolute trade flow is larger on Uniswap than Binance. Take WBTC-ETH as an example. The standard deviation of per-block-time trade flow is about 77 Ethereum on Uniswap compared with 8.9 Ethereum on Binance.

Third, across all token pairs, Uniswap trade flows with higher gas price are larger in magnitude. For ETH-USDT, the standard deviation of Uniswap high-gas trade flow is 36.87 Ethereum, which is more than three times larger than that of low-gas trade flow.

5.2 Blockchain fees and private information

Now, we investigate whether DEX trades with high blockchain fees contain more private information. To do so, we estimate a structural VAR model (See Equation 2) where we include CEX return and DEX trade flows with different levels of blockchain fees.

5.2.1 Permanent price impacts of DEX trade flows

If DEX trade flow with high blockchain fees contains more private information than that with low blockchain fees, we should see the former has a larger permanent price impact. In a structural VAR framework, permanent price impact of a particular trade flow can be estimated by the cumulative impulse responses of return to its unexpected component (See Equation 4). In Table 5, we report the cumulative impulse responses of CEX return to DEX trade flows with different blockchain fee levels. The results show that high-blockchain-fee DEX trade flow has a larger permanent price impact and thus contains more private information. We discuss the detailed results below.

We first look at the estimation results for token pairs of the NonStable-Stable and NonStable-NonStable types. Token pairs in both types have at least one non-stable coin and thus should experience frequent private information shocks. The results show that the cumulative impulse response of CEX return, r_t^{CEX}, to high-blockchain-fee DEX trade flow, $x_t^{\text{HighGas-DEX}}$, is statistically significant. In addition, for both pair types, the cumulative impulse responses of CEX return to high-blockchain-fee DEX trade flow are much larger in magnitude compared with mid- and low-blockchain-fee DEX trade flows. For instance, for NonStable-Stable token pairs, one standard deviation of positive shock to high-blockchain-fee DEX trade flow leads to a 3.41 basis points increase in CEX return while that to low-blockchain-fee trade flow leads to no significant increase. For NonStable-NonStable token pairs, although one standard deviation of positive shock to low-blockchain-fee DEX trade flow leads to a statistically significant increase in CEX return, its economic magnitude is much smaller than high-blockchain-fee DEX trade flow.

Then we turn to the estimation results for token pairs in the Stable-Stable type, which can serve as a placebo test and show that our structural VAR model works. For Stable-Stable token pairs, there is little private or public information as both tokens of the pairs are stable coins pegged to the US Dollar. So without short-term liquidity shocks, token pairs should always be priced at one. As a result, traders who trade Stable-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. For both types of trades, they can only result in a transitory impact on the prices, not permanent ones. The estimation results are consistent with our expectations: the cumulative return impulse responses to DEX trade flows are statistically insignificant, regardless of their blockchain fee levels.

Table 5. Cumulative impulse responses between CEX return and DEX trade flows with different gas price levels. This table reports the impulse responses between the CEX return and DEX trade flows with different gas price levels, cumulative over 20 blocks. Impulse responses are obtained by estimating the structural VAR model specified in Equation 2. The estimation is done for each pair-day and statistical inference is based on pair-day estimates. Row variables are response variables and column variables are shock variables. CEX return is in basis point. DEX trade flows are standardized and thus in their standard deviations. *, ** and *** indicate significance levels at 1%, 5% and 10% respectively.

PairType	Variable	r_t^{CEX}	$x_t^{\text{LowGas-DEX}}$	$x_t^{\text{MidGas-DEX}}$	$x_t^{\text{HighGas-DEX}}$
Stable-Stable	r_t^{CEX}	0.63***	0.0	0.0	−0.01
		(0.02)	(0.01)	(0.01)	(0.01)
	$x_t^{\text{LowGas-DEX}}$	0.01	1.01***	−0.02*	−0.02*
		(0.01)	(0.03)	(0.01)	(0.01)
	$x_t^{\text{MidGas-DEX}}$	−0.01	−0.06***	0.92***	−0.04***
		(0.02)	(0.01)	(0.02)	(0.02)
	$x_t^{\text{HighGas-DEX}}$	0.01	−0.13***	−0.28***	0.85***
		(0.01)	(0.03)	(0.03)	(0.03)
NonStable-Stable	r_t^{CEX}	0.97***	0.0	0.47***	3.41***
		(0.01)	(0.04)	(0.08)	(0.19)
	$x_t^{\text{LowGas-DEX}}$	0.0***	1.0***	−0.02***	−0.02***
		(0.0)	(0.01)	(0.01)	(0.01)
	$x_t^{\text{MidGas-DEX}}$	0.01***	−0.08***	0.97***	0.03***
		(0.0)	(0.01)	(0.01)	(0.01)
	$x_t^{\text{HighGas-DEX}}$	0.04***	−0.13***	−0.22***	1.06***
		(0.0)	(0.01)	(0.01)	(0.01)
NonStable-NonStable	r_t^{CEX}	0.81***	1.74***	3.47***	6.5***
		(0.01)	(0.31)	(0.37)	(0.54)
	$x_t^{\text{LowGas-DEX}}$	0.0***	0.97***	−0.01	−0.04***
		(0.0)	(0.01)	(0.01)	(0.01)
	$x_t^{\text{MidGas-DEX}}$	0.01***	−0.06***	0.99***	−0.01
		(0.0)	(0.01)	(0.01)	(0.01)
	$x_t^{\text{HighGas-DEX}}$	0.02***	−0.08***	−0.16***	1.07***
		(0.0)	(0.01)	(0.02)	(0.02)

Table 6. Information shares of DEX trade flows with different gas price levels. This table reports the information shares of the CEX return and DEX trade flows with different gas price levels. Information shares are computed based on Equation 5. The estimation is done for each pair-day and statistical inference is based on pair-day estimates. Numbers in brackets are standard errors.

PairType Variable	Stable-Stable	NonStable-Stable	NonStable-NonStable
r_t^{CEX}	97.98	89.07	84.17
	(0.27)	(0.48)	(0.96)
$x_t^{\text{LowGas-DEX}}$	0.6	0.21	1.84
	(0.14)	(0.03)	(0.35)
$x_t^{\text{MidGas-DEX}}$	0.76	0.49	3.66
	(0.16)	(0.07)	(0.49)
$x_t^{\text{HighGas-DEX}}$	0.66	10.23	10.32
	(0.14)	(0.46)	(0.73)

5.2.2 Information shares of DEX trade flows

In Table 6 we further compute the information shares of DEX trade flows with different blockchain fee levels. The information share approach not only considers the permanent price impact of a trade flow variable, but also its own (unexpected) variances (See Equation 5). For example, if two trade flow variables have the same permanent price impacts, the one with a larger (unexpected) variance will have a larger information share.

The results show that, for NonStable-Stable and NonStable-NonStable token pairs, while CEX return itself contributes the largest share to price discovery, which reflects public information, that of the high-blockchain-fee DEX trade flow has a much larger information share than that of the low-blockchain-fee DEX trade flow. For NonStable-Stable (NonStable-NonStable) pair type, the high-blockchain-fee DEX trade flow contributes about 10.23% (10.32%) to price discovery, which is much larger than 0.49% (3.66%) of the mid-blockchain-fee trade flow, and 0.21% (1.84%) of the low-blockchain-fee DEX trade flow. In contrast, for Stable-Stable token pairs, CEX return itself contributes virtually all (97.88%) price discovery. It shows that, for Stable-Stable token pairs, DEX trade flows contain barely no private information. In addition, for Stable-Stable token pairs, DEX trade flows of different blockchain fee levels contribute more or less the same.

5.3 Blockchain fees and public information

Upon the release of new public information, the exchange price at CEX can be quickly adjusted to the efficient price through quote revisions. In contrast, in most DEXs, prices are determined by a pre-coded pricing function, so DEX prices are unable to be immediately adjusted and can only be updated through the trades executed in the subsequent blocks. The response of DEX trade flows to CEX price allows us to investigate whether DEX trades with higher blockchain fees are more responsive to unexpected public information and help move DEX price towards the efficient price.

In Table 5, the first column reports the cumulative impulse responses of DEX trade flows with different blockchain fee levels to one basis point of positive shock to CEX return. It shows that, for Stable-Stable token pairs, the response of DEX trade flows to CEX return is not statistically significant, for all three blockchain fee levels. As argued above, Stable-Stable pairs such as USDC-USDT rarely have any public information, and the efficient exchange rate, 1, is common knowledge for all traders. As a result, DEX trade flows for stablecoins are not responsive to CEX price changes as they are transitory in nature.

In contrast, for NonStable-Stable and NonStable-NonStable token pairs, the cumulative impulse responses of DEX trade flows to shocks to CEX return are statistically significant and their magnitudes increase in blockchain fee levels. For instance, one basis point of positive shock to CEX return of NonStable-Stable token pairs leads to an increase of high-blockchain-fee DEX trade flow in the same direction for about 0.04 standard deviations, which is about four times larger than the response of mid-blockchain DEX trade flow. As CEX return innovations represent the arrival of public information shocks, the results show that it is mainly high-blockchain-fee DEX trades that help update DEX prices to new equilibrium levels.

In summary, the above results suggest that DEX trade flow with high blockchain fees not only contains more private information, but also responds more to public price innovations on CEXs. Thus, high-blockchain-fee trade flow plays important roles in the price discovery process of crypto trading. On the one hand, it reveals private information and incorporates it into the market prices.

On the other hand, it helps update stale prices on DEXs so that prices across DEXs and CEXs are more aligned.

5.4 Full dynamics of the impulse responses between CEX return and DEX trade flows

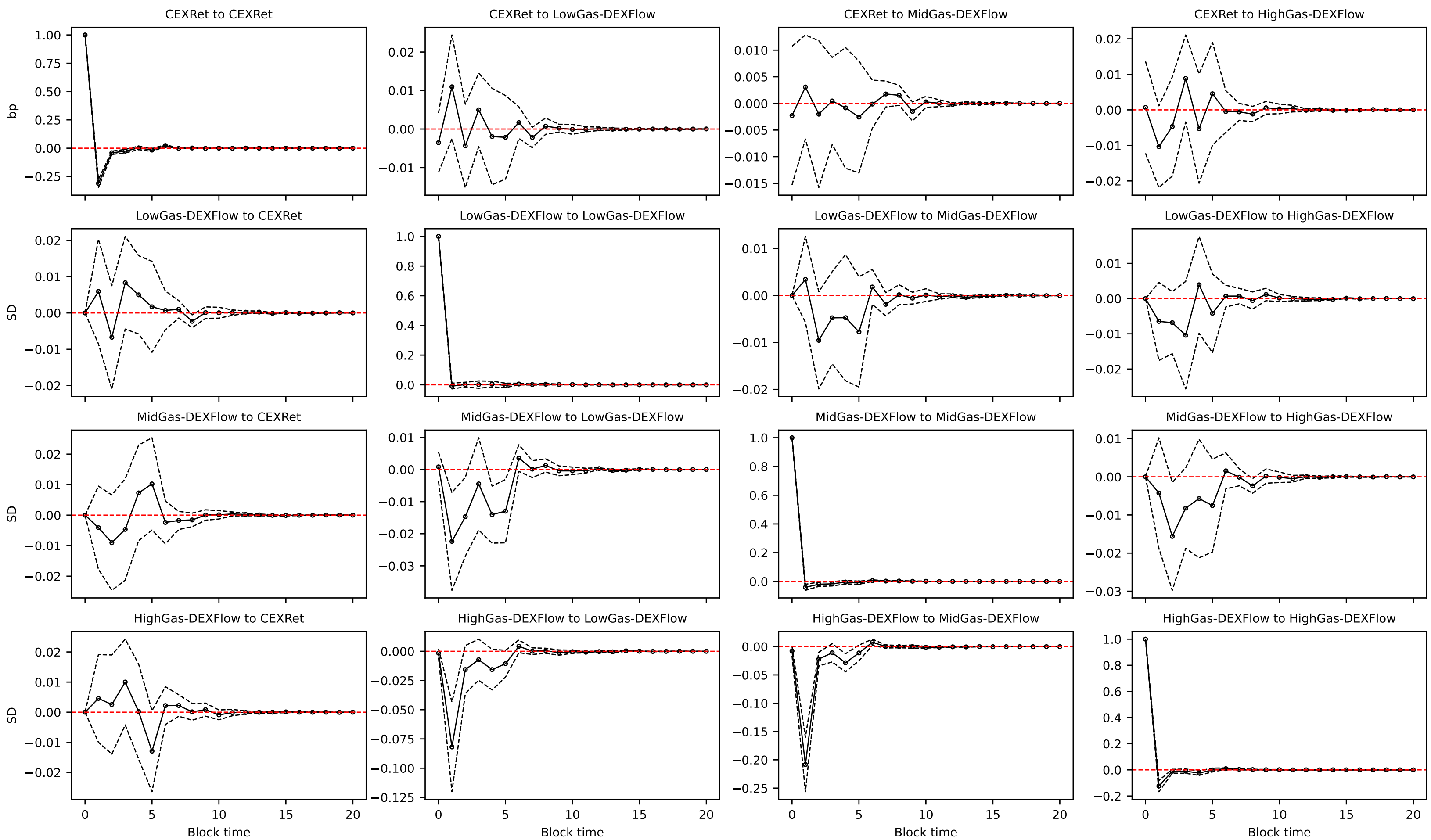
In addition to analyzing the cumulative impulse responses between the CEX return and DEX trade flows, it is useful to study full dynamics of the impulse responses so that we are able to see, for example, the speed of market price adjustments to shocks in high-blockchain-fee versus low-blockchain-fee DEX trade flow, and the persistency of DEX trade flows' responses to shocks in CEX return.

First, Panel (a) of Figure 3 plots the impulse response results between the return and trade flow variables for Stable-Stable token pairs. Consistent with the cumulative return impulse response results in Table 7, it shows that impulse responses of CEX return to DEX trades of all blockchain fee levels are insignificant over all periods. In addition, the impulse responses of DEX trade flows to CEX return, regardless of their blockchain fee levels, are statistically insignificant for all periods as well.

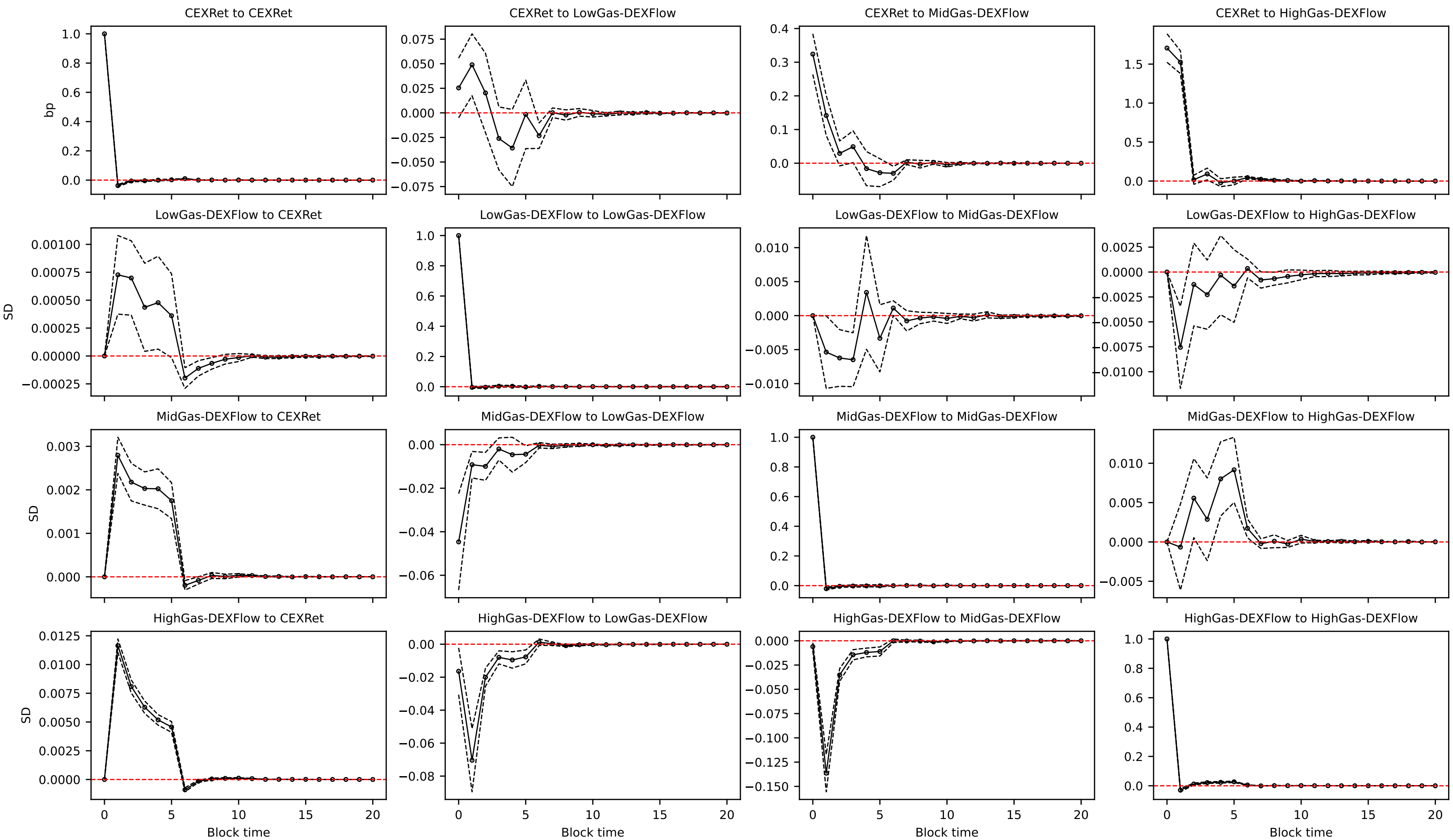
Second, Panel (b) and (c) of Figure 3 plot the impulse responses for NonStable-Stable and NonStable-NonStable token pairs respectively. It shows that while the return impulse response is positive and significant in the contemporaneous ($t = 0$) and the next block ($t = 1$), it turns insignificant from the second block ($t = 2$) onwards. It indicates that CEX return responds significantly and quickly to high-blockchain-fee DEX trade flow. Or, in other words, traders are able to learn the private information contained in the high-blockchain-fee DEX trade flow quickly and find a new equilibrium market price. In contrast, the impulse responses of high-blockchain-fee and mid-blockchain-fee trade flows to CEX return are statistically significant for around five blocks ($t = 1$ to $t = 5$). It shows that the response of DEX trade flows to public information is more sticky and takes several blocks of time.

Figure 3. Impulse response functions between CEX return and DEX trade flows with different gas price levels. This figure plots the impulse responses between the CEX return and DEX trade flows with different gas price levels over the horizon of 20 blocks. Impulse responses are obtained by estimating the structural VAR model specified in Equation 2. CEX return is in basis point and DEX trade flows are standardized and thus in standard deviations. The estimation is done for each pair-day and statistical inference is based on pair-day estimates. Dashed black lines represent 95% confidence bands.

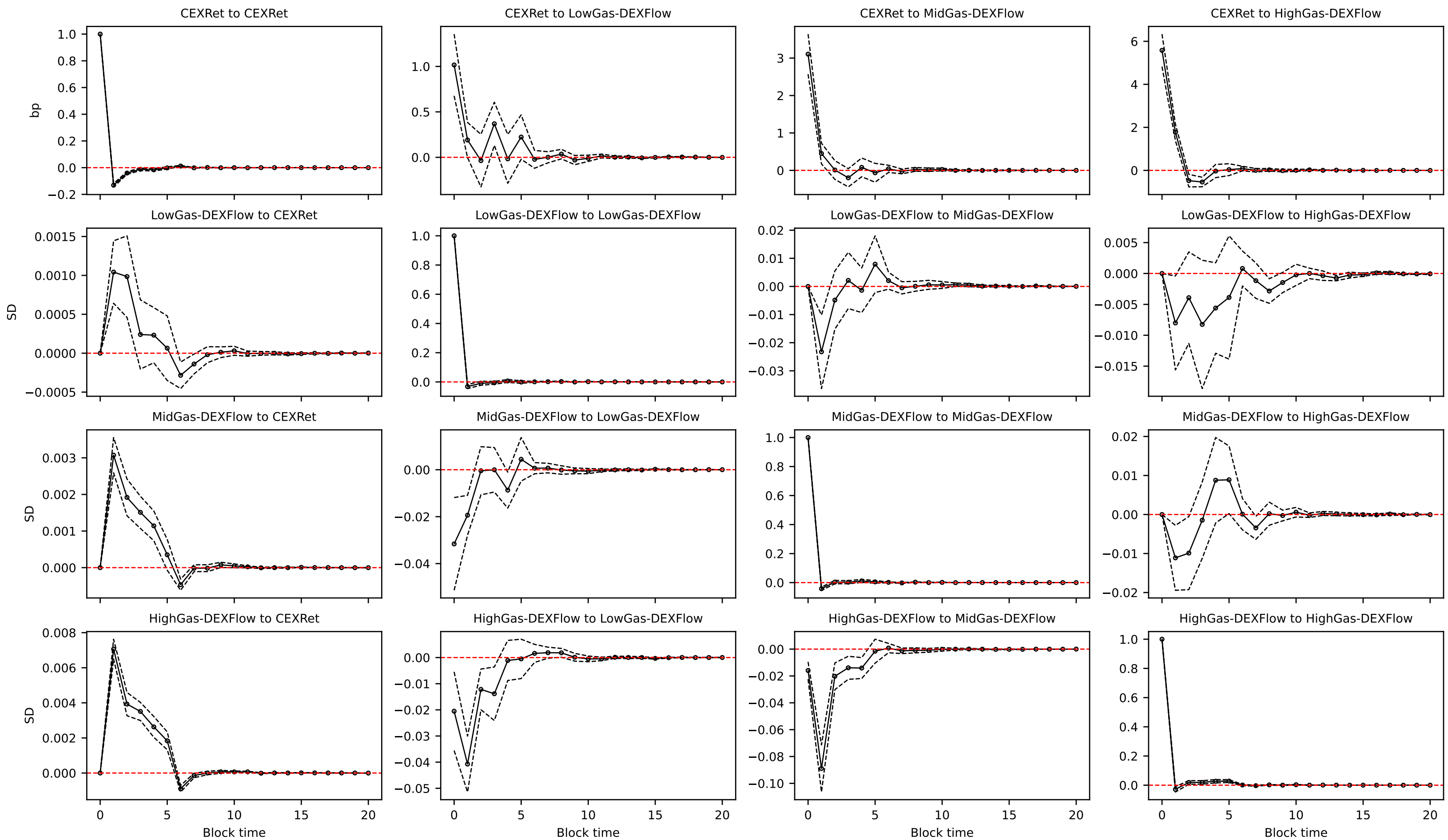
(a) Stable-Stable token pairs.



(b) NonStable-Stable token pairs.



(c) NonStable-NonStable token pairs.



5.5 Robustness: Structural VAR results with CEX trade flow included

For both informed traders trading on their private signals or arbitrageurs racing on public news, it is likely that they split their trades between CEXs and DEXs in order to reduce their price impacts. As a result, CEX trade flow and DEX trade flow can be positively correlated. As a robustness check, we include CEX trade flow in the structural VAR specification as in Equation 6. If, after controlling for CEX trade flow, DEX trade flow, especially the high-gas DEX trade flow, still has a large permanent impact on CEX price, we are more confident that DEX trade flow captures private information not contained in CEX trade flow.

In Table 7 and 8, we report permanent price impacts and information shares of DEX trade flows with different blockchain fee levels respectively, with CEX trade flow controlled. The results are qualitatively the same as the baseline: high-blockchain-fee DEX trade flow has a much larger permanent price impact and information share than mid- and low-blockchain-fee DEX trade flows. Actually, adding CEX trade flow only slightly reduces the economic magnitude of the permanent price impacts of high-blockchain-fee DEX trade flow. For example, for NonStable-Stable token pairs, the cumulative CEX return impulse response to one positive standard deviation shock in high-blockchain-fee DEX trade flow is 3.06 basis points when CEX trade flow is controlled, which is marginally smaller than 3.47 basis points when CEX trade flow is not controlled. In addition, the permanent price impact of high-blockchain-fee DEX trade flow is comparable to that of CEX trade flow: for NonStable-Stable (NonStable-NonStable) token pairs, one standard deviation of positive shock to the high-blockchain-fee DEX trade flow leads to an increase of about 3.06 (3.73) basis point in the CEX price, compared with 6.09 (5.07) basis points to the CEX trade flow.

5.6 Blockchain fees and information: Economic mechanisms

In the above sections, we have shown that high-blockchain-fee DEX trade flow not only contains more private information, but also responds more to public price innovations on CEXs compared

Table 7. Cumulative impulse responses between CEX return, CEX trade flow, and DEX trade flows with different gas price levels. This table reports the impulse responses between the CEX return, CEX trade flow, and DEX trade flows with different gas price levels, cumulative over 20 blocks. Impulse responses are obtained by estimating the structural VAR model specified in Equation 6. The estimation is done for each pair-day and statistical inference is based on pair-day estimates. Row variables are response variables and column variables are shock variables. CEX return is in basis point. CEX trade flow and DEX trade flows are standardized and thus in their standard deviations. *, ** and *** indicate significance levels at 1%, 5% and 10% respectively.

PairType	Variable	r_t^{CEX}	x_t^{CEX}	$x_t^{\text{LowGas-DEX}}$	$x_t^{\text{MidGas-DEX}}$	$x_t^{\text{HighGas-DEX}}$
Stable - Stable	r_t^{CEX}	0.55***	0.32***	0.01	0.0	−0.01
		(0.01)	(0.03)	(0.01)	(0.01)	(0.01)
	x_t^{CEX}	−0.2***	1.19***	0.01	0.0	0.0
		(0.03)	(0.07)	(0.02)	(0.02)	(0.02)
	$x_t^{\text{LowGas-DEX}}$	0.01	0.01	1.01***	−0.02*	−0.02*
		(0.02)	(0.02)	(0.03)	(0.01)	(0.01)
	$x_t^{\text{MidGas-DEX}}$	0.0	0.01	−0.06***	0.92***	−0.04***
		(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
	$x_t^{\text{HighGas-DEX}}$	0.0	0.01	−0.13***	−0.28***	0.85***
		(0.01)	(0.02)	(0.03)	(0.03)	(0.03)
NonStable - Stable	r_t^{CEX}	0.97***	3.73***	−0.01	0.4***	3.06***
		(0.01)	(0.28)	(0.04)	(0.08)	(0.17)
	x_t^{CEX}	0.0***	1.27***	0.0	0.0	0.09***
		(0.0)	(0.02)	(0.0)	(0.01)	(0.01)
	$x_t^{\text{LowGas-DEX}}$	0.0***	0.02***	1.0***	−0.02***	−0.02***
		(0.0)	(0.0)	(0.01)	(0.01)	(0.01)
	$x_t^{\text{MidGas-DEX}}$	0.01***	0.06***	−0.08***	0.97***	0.02***
		(0.0)	(0.01)	(0.01)	(0.01)	(0.01)
	$x_t^{\text{HighGas-DEX}}$	0.03***	0.22***	−0.13***	−0.22***	1.05***
		(0.0)	(0.01)	(0.01)	(0.01)	(0.01)
NonStable - NonStable	r_t^{CEX}	0.8***	5.07***	1.76***	3.18***	6.09***
		(0.01)	(0.44)	(0.32)	(0.37)	(0.55)
	x_t^{CEX}	0.0***	1.25***	0.02*	0.03***	0.1***
		(0.0)	(0.03)	(0.01)	(0.01)	(0.02)
	$x_t^{\text{LowGas-DEX}}$	0.0***	0.01	0.97***	−0.01	−0.04***
		(0.0)	(0.01)	(0.01)	(0.01)	(0.01)
	$x_t^{\text{MidGas-DEX}}$	0.01***	0.09***	−0.06***	0.98***	−0.01
		(0.0)	(0.01)	(0.01)	(0.01)	(0.01)
	$x_t^{\text{HighGas-DEX}}$	0.02***	0.19***	−0.09***	−0.17***	1.06***
		(0.0)	(0.02)	(0.01)	(0.02)	(0.02)

with low-blockchain-fee trade flow. In other words, both privately informed traders and public information arbitrageurs bid a high blockchain fee for their orders. Next, we provide possible economic channels and then use mempool data to test them.

Table 8. Robustness: Information shares of CEX trade flow and DEX trade flows with different gas price levels. This table reports the information shares of the CEX return, CEX trade flow and DEX trade flows with different gas price levels. Information shares are computed based on Equation 5. The estimation is done for each pair-day and statistical inference is based on pair-day estimates. Numbers in brackets are standard errors.

PairType Variable	Stable-Stable	NonStable-Stable	NonStable-NonStable
r_t^{CEX}	71.19	75.49	73.71
	(1.99)	(1.5)	(1.32)
x_t^{CEX}	26.91	15.51	12.48
	(1.93)	(1.5)	(1.28)
$x_t^{\text{LowGas-DEX}}$	0.62	0.21	1.74
	(0.13)	(0.03)	(0.33)
$x_t^{\text{MidGas-DEX}}$	0.74	0.45	3.27
	(0.16)	(0.07)	(0.48)
$x_t^{\text{HighGas-DEX}}$	0.54	8.35	8.8
	(0.13)	(0.4)	(0.68)

5.6.1 Channel #1: Competition among traders on the same or similar information

The first channel states that both privately informed traders and public information arbitrageurs bid high blockchain fees because they compete with each other on the same or similar information. Competition naturally arises among public information arbitrageurs as by nature they receive the same trading signal. For example, when there is a fundamental value shock to a token pair, prices on CEXs and DEXs might adjust asynchronously, creating cross-venue arbitrage opportunities. Another example of public information arbitrage is triangular arbitrage. If market is frictionless, prices of a group of linked token pairs such as ETH-USDT, ETH-USDC and USDC-USDT should be aligned. However, when there is a fundamental value shock to one token pair and prices of other token pairs adjust with a lag, cross-token-pair arbitrage opportunities might arise. The key point is that public information such as price differences across venues or linked token pairs is visible to all who exert monitoring effort, giving rise to competition.

As for competition on private information, it is more subtle. One might wonder why there is competition among privately informed traders in the first place due to its private nature. There are two possible causes: first, 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); second, there are "back-runners" (Yang and Zhu, 2020) or "predators" (Brunnermeier and Pedersen, 2005) who are not endowed with private signals but infer them from public signals such as order imbalance or blockchain fees attached to the trades in the context of DEXs.

There are two important implications from the first "trader competition" channel. The first implication is that we should observe competing traders, either for public information or private information, engage in a blockchain fee bidding game: they will keep increasing blockchain fees attached to their trades to outbid other traders and get their trades executed first in the next block. The second implication is that we should see the revenue of the competing traders be competed away.

5.6.2 Channel #2: Execution risk due to blockchain crowding

The second channel states that the reason why traders bid high blockchain fees is not due to competition with other traders. Rather, as the block space is limited, they do so to lower the execution risk and thus increase the expected profit on their private information. It is in particular true for informed traders who possess short-lived private information over, e.g, the next block, and public information arbitrageurs who can be viewed as a limit case of traders with extremely short-lived private information. In contrast, informed traders with long-lived private information or patient liquidity traders are willing to wait and thus to pay a low blockchain fee. We would like to note that in such a case, they are to some extent engaged with competition as well. The key difference is that now the competition is not with other traders, but with other users of the blockchain. For example, users who pay blockchain fees to transfer funds or bid high blockchain fees during events like initial coin offerings (ICOs).

The second "blockchain crowding" channel yields the opposite implications compared with the first "trader competition" channel. First, as there is no competition among the traders, we should not observe they engage in blockchain fee bidding competition, that is, consecutively increasing the blockchain fees attached to their orders. Instead, they will choose relatively high blockchain fees at the first place to increase the likelihood that their trades will be included by the validators in the next block. Second, given that there is no competition with other traders on the same or similar information, traders can potentially earn a positive profit on their information.

5.6.3 Identify "competition trades"

To test the two channels above, we first identify trades which are likely to be involved in competition among multiple traders ("competition trades"). The intuition is simple. When multiple traders compete either on public or private information, they are expected to increase the blockchain fees attached to their orders to outbid others and get their own orders executed first. Thus, we should see the final winning order, i.e., the executed trade, has multiple rounds of blockchain fee increase before it gets executed on chain. Specifically, we match executed trades with mempool orders based on the following three criteria:

1. *Matched orders must have the same submission address and nonce as the executed trade.* Recall that a trader on DEX has to attach a number called "nonce" to each of her orders. The most important property of 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 broadcasted 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 can not be executed before one with smaller nonce, which implies that, when a trader wants to modify her pending order, e.g., increase the blockchain fee, she needs to broadcast a new order with the same nonce as the pending one. So, the first criterion on submission address and nonce guarantees that the matched mempool orders are indeed the previous revisions of the final executed one.

2. *Gas price must increase from matched orders to the executed trade.* For both mempool orders and the executed trade, we observe the gas price attached to each of them. The second criterion requires that the gas price must increase from matched mempool orders to the final executed trade so that we capture the trades involved in gas fee bidding among traders.

3. *Matched orders and the executed trade must be in the same block.* We impose the last criterion on the time horizon because we believe gas bidding due to competition should happen within a fairly short time window. If the window is too long, it is more likely to result from patient liquidity traders who revise their blockchain fees after a long waiting time.

Table 9 reports the matching results. It shows that, surprisingly, across all eight token pairs in our sample, only a very small fraction of executed trades can be labeled as "competition trades". For example, only 0.64%, 0.78%, and 0.99% of executed trades for USDC-USDT, ETH-USDT and WBTC-ETH satisfy all three criteria and thus are "competition trades". Even if we drop the last criterion on the maximum time duration between the first matched mempool order and the final executed trade, which is subject to the researcher's choice, only 2.86%, 3.13%, and 3.20% of executed trades have a match for USDC-USDT, ETH-USDT and WBTC-ETH respectively. In other words, not only do most traders not engage in gas price bidding game, but also they rarely increase their gas price at all. It seems to be the case that they choose a gas price ex ante that they deem will be sufficient for getting their orders executed in a reasonable time window.

Table 9. Matching results between executed trades and mempool orders. This table shows the results of matching between executed trades and mempool orders. "All-Matched" shows the percentage of executed trades with at least one matched order, regardless of the last criteria on the time horizon. "Unmatched" shows the percentage of executed trades without any matched order, regardless of the last criteria on the time horizon. The "All-Matched" rate is further broken down based on the block distance between the first matched order and the executed trade.

PairType	Blocks Pair	0	1	2	3-5	6-10	10+	All-Matched	Unmatched
Stable-Stable	USDC-USDT	0.64	0.26	0.15	0.34	0.34	1.13	2.86	97.14
	DAI-USDT	0.63	0.35	0.19	0.33	0.42	1.46	3.39	96.61
NonStable-Stable	ETH-USDT	0.78	0.34	0.19	0.33	0.30	1.19	3.13	96.87
	ETH-USDC	0.92	0.35	0.19	0.33	0.29	1.19	3.27	96.73
	ETH-DAI	1.08	0.39	0.22	0.38	0.35	1.51	3.94	96.06
NonStable-NonStable	WBTC-ETH	0.99	0.31	0.15	0.33	0.28	1.15	3.20	96.80
	LINK-ETH	1.32	0.51	0.27	0.37	0.35	1.56	4.37	95.63
	AAVE-ETH	1.59	0.50	0.27	0.24	0.34	1.69	4.64	95.36

5.6.4 Information content of "competition trades" versus "non-competition trades"

Although the above matching results show that only a very small fraction of executed trades is likely to be involved in trader competition, they might play a disproportionately large role in price discovery and drive the structural VAR results we observe in the previous section. To formally test whether it is the case, we adopt the following two empirical strategies.

Excluding "competition trades" In the first approach, we construct DEX trade flows with different gas price levels excluding "competition trades" and then re-implement the structural VAR analysis. If high gas prices are the result of competition among privately informed traders and public information arbitrageurs, we should see both the impulse response of CEX return to the high-gas DEX trade flow shock and that of high-gas DEX trade flow to CEX return shock become either statistically insignificant or smaller in magnitude, after we exclude the "competition trades".

Table 10 reports the estimation results. Compared with the baseline results when "competition trades" are included (See Table 5), the results with "competition trades" excluded barely change. Again, we see that, for both NonStable-Stable and NonStable-NonStable token pairs, (1) the cumulative impulse response of CEX return to high-gas DEX trade flow shock is larger than that to mid- and low-gas DEX trade flows, and (2) the cumulative impulse response of high-gas DEX trade flow to CEX return shock is larger than that of mid- and low-gas DEX trade flows. In addition, the magnitudes of both impulse responses do not change much compared with the baseline. The results illustrate that our key results—high-gas-price DEX trade flow is more privately informed and more responsive to public information—are not driven by competition among traders.

Separating "competition trades" from "non-competition trades" In the second approach, we further separate "competition trades" from "non-competition trades" for DEX trades in each gas level and thus construct the following four different DEX trade flows: low-gas, competition trade flow ($x_t^{\text{LG-C-DEX}}$), low-gas, non-competition trade flow ($x_t^{\text{LG-NC-DEX}}$), high-gas, competition trade flow ($x_t^{\text{HG-C-DEX}}$), and high-gas, non-competition trade flow ($x_t^{\text{LG-NC-DEX}}$). Specifically, low-gas,

Table 10. Cumulative impulse responses between CEX return and DEX trade flows: Excluding "competition trades". This table reports the impulse responses between the return and trade flow variables, cumulative over 20 blocks. Impulse responses are obtained by estimating the structural VAR model specified in Equation 2. The estimation is done for each pair-day and statistical inference is based on pair-day estimates. Row variables are response variables and column variables are shock variables. CEX return is in basis point. DEX trade flows are standardized and thus in their standard deviations. *, ** and *** indicate significance levels at 1%, 5% and 10% respectively.

PairType	Variable	r_t^{CEX}	$x_t^{\text{LowGas-DEX}}$	$x_t^{\text{MidGas-DEX}}$	$x_t^{\text{HighGas-DEX}}$
Stable-Stable	r_t^{CEX}	0.59***	0.46	0.05	0.01
		(0.01)	(0.42)	(0.16)	(0.05)
	$x_t^{\text{LowGas-DEX}}$	0.01	1.0***	−0.02	−0.03***
		(0.01)	(0.04)	(0.02)	(0.01)
	$x_t^{\text{MidGas-DEX}}$	−0.01	−0.17**	0.89***	−0.07***
		(0.01)	(0.08)	(0.04)	(0.02)
	$x_t^{\text{HighGas-DEX}}$	0.01	−0.07	−0.25***	0.81***
		(0.01)	(0.06)	(0.03)	(0.02)
NonStable-Stable	r_t^{CEX}	0.96***	0.0	0.41***	3.25***
		(0.01)	(0.04)	(0.08)	(0.18)
	$x_t^{\text{LowGas-DEX}}$	0.0***	1.0***	−0.02**	−0.01***
		(0.0)	(0.01)	(0.01)	(0.01)
	$x_t^{\text{MidGas-DEX}}$	0.01***	−0.08***	0.97***	0.02***
		(0.0)	(0.01)	(0.01)	(0.01)
	$x_t^{\text{HighGas-DEX}}$	0.03***	−0.13***	−0.21***	1.06***
		(0.0)	(0.01)	(0.01)	(0.01)
NonStable-NonStable	r_t^{CEX}	0.81***	1.84***	3.16***	6.47***
		(0.01)	(0.34)	(0.35)	(0.53)
	$x_t^{\text{LowGas-DEX}}$	0.0***	0.97***	−0.01	−0.03***
		(0.0)	(0.01)	(0.01)	(0.01)
	$x_t^{\text{MidGas-DEX}}$	0.01***	−0.06***	0.99***	0.0
		(0.0)	(0.01)	(0.01)	(0.01)
	$x_t^{\text{HighGas-DEX}}$	0.02***	−0.08***	−0.15***	1.07***
		(0.0)	(0.01)	(0.02)	(0.02)

competition trade flow, $x_t^{\text{LG-C-DEX}}$, is constructed based on trades that satisfy both conditions below: (1) they have a gas price below the 50% quantile of all trades in the last 20 non-empty blocks; (2) they are identified as "competition trades".[9] Other three DEX trade flows are defined likewise. Then we re-implement the structural VAR estimation with the four DEX trade flows. If high gas

[9]Note that we only classify trades into two categories ($x_t^{\text{LowGas-DEX}}$, and $x_t^{\text{HighGas-DEX}}$) based on their gas price levels instead of three categories ($x_t^{\text{LowGas-DEX}}$, $x_t^{\text{MidGas-DEX}}$ and $x_t^{\text{HighGas-DEX}}$) as in our baseline specification above. We do so because of the small number of "competition trades" in our data. For the finer gas level classification, we end up with zero observation for certain trade flow, e.g., low-gas, competition trade flow, and thus makes the structural VAR estimation unfeasible.

prices are driven by competition, we should see (1) the impulse response of CEX return to high-gas, competition DEX trade flow shock larger than that to high-gas, non-competition trade flow shock, and (2) the impulse response of high-gas, competition DEX trade flow to CEX return shock larger than that of high-gas, non-competition trade flow.

In Table 11, we report the estimation results from our new structural VAR specification. It shows that the permanent price impact of non-competition DEX trade flow is much larger than that of competition DEX trade flow, especially for the high-gas category. For example, for NonStable-Stable token pairs, one standard deviation of a positive shock to the high-gas, non-competition DEX trade flow, $x_t^{\text{HG-NC-DEX}}$, leads to an increase of about 3.06 basis points in the CEX return while the same amount of positive shock to the high-gas, competition DEX trade flows, $x_t^{\text{HG-C-DEX}}$, increases CEX return by only 1.16 basis points. In addition, impulse response of high-gas, non-competition trade flow to CEX return shock is larger than that of high-gas, competition trade flow. Specifically, for NonStable-Stable token pairs, a positive shock of one basis point to CEX return leads to an increase of about 0.04 standard deviation in the high-gas, non-competition trade flow, compared with about 0.01 standard deviation in the high-gas, competition trade flow. The above results demonstrate that high-gas, non-competition trades are more privately informed and responsive to public information than high-gas competition trades. In other words, "trader compeittion" channel is not likely to drive our structural VAR results.

5.7 Trading revenue and blockchain fees

For both privately informed traders and public information arbitrageurs, blockchain fee is an extra source of transaction cost in addition to the immediate price impact. So, if they are willing to pay a higher blockchain fee, it must be the case that doing so leads to higher profits for them, given the fact that there is barely any competition as we have established above. To confirm our conjecture, we study the profitability of DEX trades with different blockchain fee levels.

Table 11. Cumulative impulse responses between CEX return and DEX trade flows: Separating "competition trades" from "non-competition trades". This table reports the impulse responses between the return and trade flow variables, cumulative over 20 blocks. $x_t^{\text{LG-C-DEX}}$ represents low-gas, competition trade flow. $x_t^{\text{LG-NC-DEX}}$ indicates low-gas, non-competition trade flow. $x_t^{\text{HG-C-DEX}}$ indicates high-gas, competition trade flow. $x_t^{\text{HG-NC-DEX}}$ means high-gas, non-competition trade flow. Impulse responses are obtained by estimating the structural VAR model. The estimation is done for each pair-day and statistical inference is based on pair-day estimates. Row variables are response variables and column variables are shock variables. The return variable is in basis point. Trade flow variables are standardized and thus in their standard deviations. *, ** and *** indicate significance levels at 1%, 5% and 10% respectively.

PairType	Variable	r_t^{CEX}	$x_t^{\text{LG-NC-DEX}}$	$x_t^{\text{LG-C-DEX}}$	$x_t^{\text{HG-NC-DEX}}$	$x_t^{\text{HG-C-DEX}}$
Stable - Stable	r_t^{CEX}	0.61***	0.02	0.01	0.0	0.0
		(0.03)	(0.02)	(0.02)	(0.02)	(0.01)
	$x_t^{\text{LG-NC-DEX}}$	0.0	0.91***	0.01	−0.03	0.04
		(0.02)	(0.03)	(0.03)	(0.02)	(0.03)
	$x_t^{\text{LG-C-DEX}}$	0.04	−0.04	0.98***	0.0	0.03
		(0.03)	(0.04)	(0.01)	(0.02)	(0.02)
	$x_t^{\text{HG-NC-DEX}}$	−0.01	−0.25***	0.02*	0.76***	−0.01
		(0.02)	(0.04)	(0.01)	(0.05)	(0.02)
	$x_t^{\text{HG-C-DEX}}$	0.02	−0.04	0.02	−0.08	0.99***
		(0.02)	(0.04)	(0.02)	(0.1)	(0.01)
NonStable - Stable	r_t^{CEX}	0.97***	0.11	0.09	3.06***	1.16***
		(0.01)	(0.07)	(0.07)	(0.2)	(0.16)
	$x_t^{\text{LG-NC-DEX}}$	0.0***	0.98***	0.0	0.0	0.01
		(0.0)	(0.01)	(0.01)	(0.01)	(0.0)
	$x_t^{\text{LG-C-DEX}}$	0.0	−0.01	1.0***	−0.02***	0.0
		(0.0)	(0.01)	(0.01)	(0.0)	(0.0)
	$x_t^{\text{HG-NC-DEX}}$	0.04***	−0.22***	0.01	1.0***	0.05***
		(0.0)	(0.02)	(0.01)	(0.01)	(0.01)
	$x_t^{\text{HG-C-DEX}}$	0.01***	−0.06***	0.0	−0.04***	1.04***
		(0.0)	(0.01)	(0.0)	(0.01)	(0.01)
NonStable - NonStable	r_t^{CEX}	0.83***	2.45***	1.37***	7.03***	1.34***
		(0.02)	(0.5)	(0.46)	(0.78)	(0.47)
	$x_t^{\text{LG-NC-DEX}}$	0.0***	1.01***	0.03	−0.05***	0.0
		(0.0)	(0.02)	(0.02)	(0.02)	(0.01)
	$x_t^{\text{LG-C-DEX}}$	0.0***	−0.06**	1.03***	−0.05***	0.01
		(0.0)	(0.03)	(0.02)	(0.01)	(0.01)
	$x_t^{\text{HG-NC-DEX}}$	0.02***	−0.13***	0.0	1.05***	0.03
		(0.0)	(0.02)	(0.02)	(0.03)	(0.02)
	$x_t^{\text{HG-C-DEX}}$	0.0***	−0.06***	−0.01	−0.03	1.03***
		(0.0)	(0.01)	(0.02)	(0.02)	(0.01)

5.7.1 A simple revenue metric

We use a simple metric to compute the revenue of a DEX trade:

$$\text{Revenue}_t = \text{RPI}_t - \text{RES}_t - \text{BlockchainFee}_t, \tag{7}$$

where t indexes Uniswap trades. RPI_t is the relative price impact, which serves as a proxy for the gross trading revenue. RES_t is the relative effective spread and measures the transaction cost of the trade resulting from both its (immediate) price impact along the Uniswap pricing curve and the Uniswap transaction fee of 30 basis points. BlockchainFee_t is the relative blockchain fee defined as the dollar blockchain fee divided by the dollar size of the trade.

In addition, we define the first two components of our revenue metric, RPI_t and RES_t, below:

$$\text{RPI}_t = \frac{d_i\,(\text{Mid}_{t+\Delta_t} - \text{Mid}_t)}{\text{Mid}_t}, \tag{8}$$

$$\text{RES}_t = \frac{d_i\,(p_t - \text{Mid}_t)}{\text{Mid}_t}. \tag{9}$$

where Mid_t represents the prevailing "midquote" *just before* trade t. $\text{Mid}_{t+\Delta_t}$ represents the prevailing "midquote" Δ_t after the trade. d_i is the trade direction indicator. Note that in CEXs with a limit order book, midquote is simply the average of the best bid and ask. As we do not have quotes in the Uniswap, midquote is computed as the ratio of the amount of two tokens in the pool, y/x, i.e., the hypothetical price for an infinitesimal trade. p_t is the transaction price. In LOB market, p_t is computed as the volume-weighted average price (VWAP) of the trade as one incoming market order can hit multiple limit orders with different prices, i.e., "walk the book". In contrast, we don't have the same issue in the AMM, so p_t is simply computed as the ratio of the amount of two tokens swapped, i.e., $\Delta y/\Delta x$, which, without transaction fees, should be approximately equal to $y/(x - \Delta x)$.

5.7.2 Revenue of DEX trades with different blockchain fee levels

After defining our revenue metric, we are ready now to compute it and its three components for all trades in our sample. Table 12 reports their summary statistics by token pair and gas fee level. There are three main takeaways. First, across all eight token pairs and gas fee levels, revenue of the majority of DEX trades is negative. The result is perhaps not surprising. For uninformed traders which make up a large share of all traders, they trade the token pairs to meet their liquidity needs or to realize their private values. So, they are wiling to incur a loss in their trades to get them executed. Essentially, they pay a service fee to the liquidity providers in the pool.

Second, for NonStable-Stable and NonStable-NonStable token pairs, there is a much higher fraction of positive-revenue DEX trades in the high-blockchain-fee category than in the low- and mid-blockchain-fee categories. For example, for both token pairs of ETH-USDT and WBTC-ETH, when sorted by revenue in an ascending order, a low- or mid-blockchain-fee DEX trade at the 90% quantile has a negative revenue while a high-blockchain-fee DEX trade at 90% quantile has a positive revenue. In other words, high-blockchain-fee DEX trades are more likely to be profitable compared with low- and mid-blockchain-fee DEX trades.

Third, virtually all DEX trades in the Stable-Stable token pairs have a negative revenue. For example, for both USDC-USDT and DAI-USDT, even trades at the extreme tail such as 95% quantile have a negative revenue. As argued above, there is hardly any private or public information contained in the Stable-Stable token pairs, thus it is expected that most trades in the token pair are for liquidity reasons and thus unlikely to become profitable. To present visual evidence on the revenue analysis, Figure 4 plots the revenue distribution by gas fee level. It can be clearly seen that as we move from trades in the low- and mid- to the high-blockchain-fee category, a larger fraction of DEX trades have a positive revenue.

Table 12. Summary statistics of the revenue metric. This table reports the summary statistics of our trading revenue metric and its three components respectively. Specifically, the trading revenue metric is defined as:

$$\text{Revenue}_t = \text{RPI}_t - \text{RES}_t - \text{BlockchainFee}_t$$

, where RPI_t is relative price impact. RES_t is relative effective spread. BlockchainFee_t is relative blockchain fee computed as the dollar blockchain fee divided by the dollar size of the trade. All variables are in basis points.

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

Pair	GasLevel	Component	N	5%	10%	25%	50%	75%	90%	95%
USDC-USDT	LowGas	RPI	15512	-37.79	-26.37	-10.24	4.94	21.24	39.16	50.43
		RES	15512	30.08	30.13	30.26	30.68	31.82	34.17	37.10
		GasFee	15512	9.93	18.76	49.29	140.08	404.02	1061.10	1953.37
		Rev	15512	-1983.20	-1087.62	-430.79	-168.07	-79.29	-44.43	-28.65
	MidGas	RPI	33266	-37.83	-26.88	-9.91	6.19	22.99	40.65	52.98
		RES	33266	30.10	30.16	30.36	30.98	32.47	35.63	39.75
		GasFee	33266	9.08	16.78	42.75	117.55	340.66	887.93	1705.57
		Rev	33266	-1729.56	-916.30	-367.64	-145.98	-72.50	-40.91	-25.52
	HighGas	RPI	17261	-38.09	-26.25	-8.75	7.82	25.78	44.89	58.20
		RES	17261	30.15	30.27	30.66	31.71	33.91	39.59	45.89
		GasFee	17261	9.22	15.44	39.36	96.48	250.32	684.78	1352.19
		Rev	17261	-1377.11	-711.73	-276.81	-126.63	-67.97	-37.17	-20.24
DAI-USDT	LowGas	RPI	7105	-49.76	-35.59	-12.91	7.90	29.42	50.16	63.54
		RES	7105	30.11	30.19	30.50	31.44	34.01	39.29	44.43
		GasFee	7105	14.04	25.10	63.39	183.53	546.44	1481.97	3025.50
		Rev	7105	-3077.90	-1496.92	-572.21	-208.94	-95.21	-52.14	-32.52
	MidGas	RPI	14973	-46.96	-33.20	-11.59	9.88	31.85	53.45	67.75
		RES	14973	30.16	30.28	30.76	32.10	35.40	42.59	49.10
		GasFee	14973	13.33	23.12	59.04	154.30	448.97	1244.70	2415.85
		Rev	14973	-2435.78	-1259.46	-473.06	-180.74	-90.28	-49.24	-28.84
	HighGas	RPI	7985	-45.04	-31.26	-9.37	12.96	35.90	59.92	78.40
		RES	7985	30.25	30.48	31.25	33.27	38.01	47.69	56.93
		GasFee	7985	14.32	22.49	52.54	137.36	364.50	1027.85	2048.14
		Rev	7985	-2061.96	-1055.96	-388.69	-162.70	-83.08	-43.35	-20.89

5.7.3 Do "competition trades" constitute the majority positive-revenue DEX trades?

Although "competition trades" constitute a rather small fraction of all trades (See Table 9), they might constitute a large fraction of positive-revenue DEX trades. To see whether it is true, we compute the percentage of "competition trades" versus "non-competition trades" for both positive-revenue and negative-revenue DEX trades. Table 13 reports the results. It shows that although for positive-revenue DEX trades, a larger fraction of them comes from "competition trades" compared with negative-revenue DEX trades, the percentage remains rather low. Take DEX trades in the token pair of ETH-USDT as an example. Out of all positive-revenue trades, only 1.60% of trades are competition trades. In other words, the vast majority of positive-revenue trades remain to be

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

Pair	GasLevel	Component	N	5%	10%	25%	50%	75%	90%	95%
ETH-USDT	LowGas	RPI	261367	-116.30	-77.24	-31.84	1.10	34.37	79.86	118.81
		RES	261367	30.01	30.02	30.09	30.14	30.32	30.79	31.38
		GasFee	261367	8.32	16.75	49.43	154.05	468.55	1242.21	2329.24
		Rev	261367	-2363.69	-1276.65	-508.94	-197.80	-80.91	-27.41	6.64
	MidGas	RPI	667625	-115.60	-75.95	-31.02	1.41	34.74	80.66	121.14
		RES	667625	30.01	30.03	30.10	30.18	30.45	31.13	32.23
		GasFee	667625	6.11	12.91	37.97	112.35	348.84	968.76	1857.24
		Rev	667625	-1890.12	-1004.70	-394.57	-158.34	-66.83	-17.17	18.99
	HighGas	RPI	293587	-109.91	-71.70	-27.60	5.22	41.68	91.92	136.61
		RES	293587	30.03	30.08	30.15	30.40	31.24	35.50	39.88
		GasFee	293587	2.46	4.60	19.86	65.89	191.18	559.09	1079.14
		Rev	293587	-1114.72	-598.77	-244.19	-109.08	-40.76	15.05	59.88
ETH-USDC	LowGas	RPI	229913	-123.97	-83.29	-34.75	1.18	37.64	85.34	127.53
		RES	229913	30.01	30.02	30.08	30.14	30.33	30.84	31.59
		GasFee	229913	7.51	15.96	49.40	154.81	459.86	1201.32	2275.59
		Rev	229913	-2311.89	-1234.57	-501.99	-200.25	-81.52	-23.51	13.06
	MidGas	RPI	584806	-123.97	-82.24	-33.73	1.98	38.56	88.35	132.43
		RES	584806	30.01	30.03	30.10	30.19	30.51	31.43	33.05
		GasFee	584806	4.33	9.99	33.69	107.05	336.11	903.46	1692.12
		Rev	584806	-1727.67	-941.11	-383.61	-154.92	-62.05	-8.58	32.38
	HighGas	RPI	259592	-118.47	-77.82	-29.85	6.33	45.81	100.18	148.67
		RES	259592	30.04	30.09	30.17	30.46	31.65	36.43	39.62
		GasFee	259592	2.21	3.59	13.92	59.63	187.57	545.29	1057.69
		Rev	259592	-1095.92	-589.20	-243.60	-105.26	-33.30	26.52	74.48
ETH-DAI	LowGas	RPI	114421	-125.69	-82.46	-32.77	1.59	37.32	87.57	132.62
		RES	114421	30.00	30.01	30.08	30.14	30.36	31.15	32.65
		GasFee	114421	5.53	12.75	46.35	166.06	555.62	1642.26	3488.11
		Rev	114421	-3517.76	-1674.85	-596.51	-211.02	-78.74	-21.90	17.97
	MidGas	RPI	276004	-124.97	-80.61	-30.85	2.86	39.06	91.77	137.99
		RES	276004	30.01	30.02	30.10	30.19	30.60	32.43	35.98
		GasFee	276004	2.94	6.77	28.56	110.50	391.19	1219.73	2573.99
		Rev	276004	-2615.18	-1252.30	-436.41	-158.58	-57.98	-3.18	41.23
	HighGas	RPI	125957	-114.47	-72.82	-24.66	9.03	50.20	108.79	160.72
		RES	125957	30.02	30.08	30.16	30.55	33.20	39.38	43.32
		GasFee	125957	1.63	2.58	7.67	49.15	203.58	674.95	1495.18
		Rev	125957	-1523.08	-713.64	-257.94	-96.28	-22.27	40.58	92.62

"non-competition trades".

To sum up, results from both the structural VAR results and the following revenue analysis favor the second channel as opposed to the first: traders bid high blockchain fees in order to lower the execution risk of their orders and thus realize a profit from the information. It is not likely that they bid high blockchain fees in order to win the competition from other traders.

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

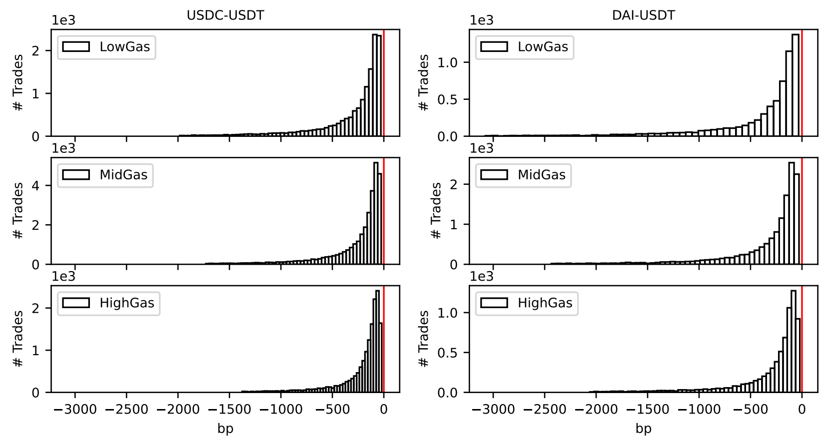
Pair	GasLevel	Component	N	5%	10%	25%	50%	75%	90%	95%
WBTC-ETH	LowGas	RPI	29383	-73.87	-47.49	-17.24	1.80	22.39	53.82	80.97
		RES	29383	30.01	30.03	30.11	30.19	30.56	31.99	34.11
		GasFee	29383	2.63	6.24	26.34	106.92	401.95	1254.80	2594.36
		Rev	29383	-2630.46	-1284.96	-436.11	-145.92	-58.30	-22.86	0.94
	MidGas	RPI	65471	-73.55	-46.32	-15.44	3.99	27.59	62.07	93.24
		RES	65471	30.02	30.06	30.13	30.31	31.27	34.42	37.52
		GasFee	65471	1.55	2.85	11.74	63.52	266.62	943.94	2134.96
		Rev	65471	-2166.65	-974.61	-305.03	-104.84	-39.54	-2.35	29.73
	HighGas	RPI	32540	-65.52	-38.58	-9.23	10.78	40.22	82.60	119.19
		RES	32540	30.06	30.11	30.25	31.13	34.21	38.41	41.36
		GasFee	32540	1.32	2.14	4.79	21.45	121.67	497.28	1132.20
		Rev	32540	-1167.05	-535.44	-166.87	-57.87	-13.90	30.47	67.32
LINK-ETH	LowGas	RPI	17207	-134.39	-86.75	-32.53	2.54	39.16	92.00	139.33
		RES	17207	30.05	30.10	30.19	30.49	31.42	34.15	37.56
		GasFee	17207	5.29	10.06	31.33	98.86	316.25	905.84	1744.66
		Rev	17207	-1775.67	-939.25	-369.41	-147.77	-58.62	-5.51	36.21
	MidGas	RPI	38151	-131.81	-80.34	-27.56	6.66	46.67	105.48	161.86
		RES	38151	30.07	30.12	30.28	30.92	33.42	39.41	44.46
		GasFee	38151	3.05	5.32	14.85	58.29	212.81	680.46	1418.33
		Rev	38151	-1453.14	-733.25	-277.22	-106.44	-33.60	27.95	84.61
	HighGas	RPI	18384	-106.42	-63.94	-14.62	22.15	71.07	140.14	203.05
		RES	18384	30.13	30.25	31.00	34.17	39.06	44.61	49.64
		GasFee	18384	3.08	4.34	7.68	19.48	85.34	360.85	847.79
		Rev	18384	-896.66	-421.37	-154.77	-49.52	9.90	78.80	139.72
AAVE-ETH	LowGas	RPI	7260	-197.29	-131.80	-53.59	5.52	68.95	154.31	227.07
		RES	7260	30.10	30.15	30.35	31.17	34.29	42.26	49.62
		GasFee	7260	6.19	11.91	37.99	138.28	458.67	1194.60	2307.88
		Rev	7260	-2345.24	-1239.36	-509.94	-210.04	-64.11	30.21	103.30
	MidGas	RPI	16861	-189.73	-121.94	-40.24	20.92	94.55	194.24	275.49
		RES	16861	30.13	30.20	30.75	33.68	43.36	53.17	64.78
		GasFee	16861	3.92	6.60	14.90	59.11	290.04	930.24	1905.84
		Rev	16861	-1945.75	-962.48	-360.62	-123.10	-7.47	98.01	181.15
	HighGas	RPI	7920	-157.61	-94.87	-19.04	45.35	120.83	226.63	303.93
		RES	7920	30.29	30.80	34.41	43.12	50.10	59.46	68.32
		GasFee	7920	5.16	7.03	10.89	22.68	83.46	432.10	1060.93
		Rev	7920	-1102.89	-499.44	-172.42	-44.55	41.05	145.42	224.54

6 Conclusion

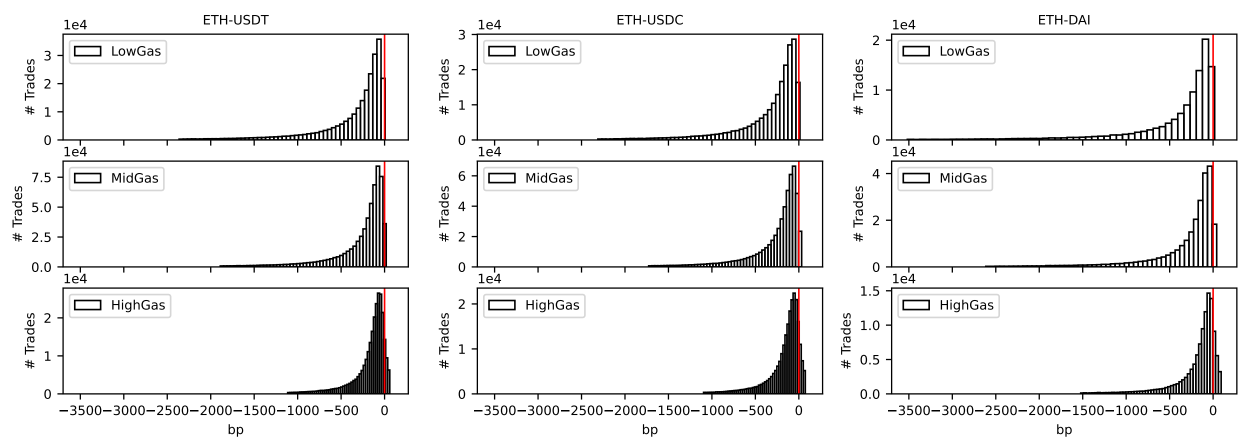
Decentralized exchanges (DEXs) have gained a significant market share in crypto trading since their inception. Unlike centralized exchanges (CEXs) which continuously execute incoming transactions based on their arrival time, DEXs process transactions in batches and prioritize their executions based on blockchain fees bid by their submitters. Thus, blockchain fee is a central element and an important choice variable of traders in DEX trading. In this paper, we study the information

Figure 4. Revenue distribution of DEX trades by gas fee level. This figure plots, for each token pair, the revenue distribution of DEX trades gas price level. The vertical red line indicates zero revenue.

(a) Stable-Stable token pairs.



(b) NonStable-Stable token pairs.



content of blockchain fees. Using a structural vector-autoregressive structural (structural VAR) model, we show that, compared with low-blockchain-fee trades, high-blockchain-fee trades not only reveal more private information, but also respond more to public price innovations on CEXs. We further test possible economic channels with a unique dataset of Ethereum mempool orders and find that informed traders or public information arbitrageurs do not bid high blockchain fee due to competition among them. Rather, it is likely that they do so to avoid the execution risk of their orders due to blockchain crowding. Our results demonstrate that blockchain fees play a conducive role in the price discovery process of crypto trading and market price efficiency.

(c) NonStable-NonStable token pairs.

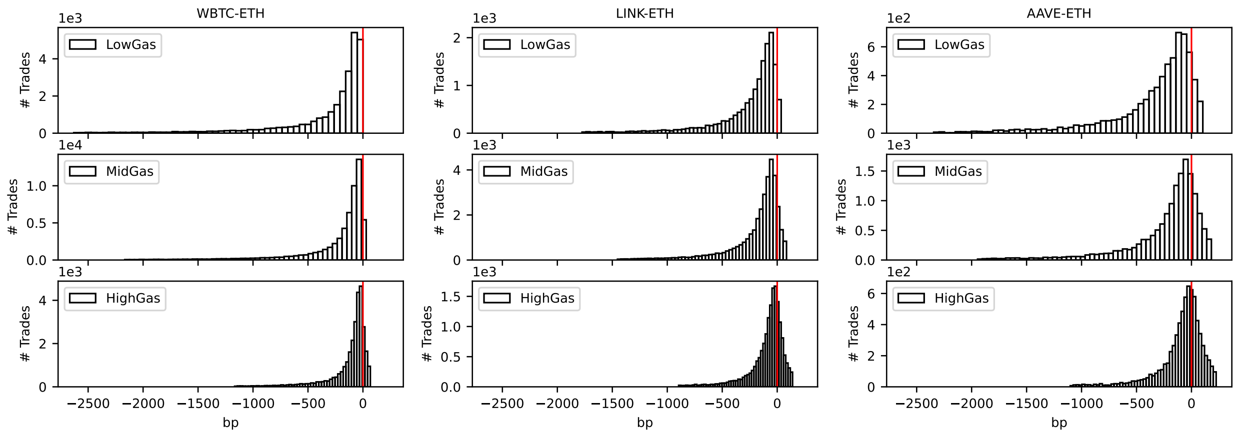


Table 13. Trade revenue: "Competition trades" versus "non-competition trades". This table reports the percentage of "competition trades" (i.e., trades likely involved in trader competition) versus the percentage of "non-competition trades" (i.e., trades not likely involved in trader competition) for the positive-revenue and negative-revenue traders respectively.

PairType	Pair	Revenue	Competition trades	Non-competition trades
Stable-Stable	USDC-USDT	% Negative	0.38	99.62
		% Positive	2.98	97.02
	DAI-USDT	% Negative	0.33	99.67
		% Positive	0.00	100.00
NonStable-Stable	ETH-USDT	% Negative	0.42	99.58
		% Positive	1.60	98.40
	ETH-USDC	% Negative	0.45	99.55
		% Positive	1.81	98.19
	ETH-DAI	% Negative	0.50	99.50
		% Positive	2.39	97.61
NonStable-NonStable	WBTC-ETH	% Negative	0.44	99.56
		% Positive	2.06	97.94
	LINK-ETH	% Negative	0.83	99.17
		% Positive	1.92	98.08
	AAVE-ETH	% Negative	0.89	99.11
		% Positive	1.34	98.66

A Robustness

Below we conduct a series of robustness checks.

A.1 Gas level classification

In our baseline structural VAR specification, we include (signed) DEX trade flows with high-, mid-, and low-gas price level respectively. Specifically, high-gas (low-gas) DEX trade flow is computed based on trades with a gas price above 75% (below 25%) quantile of the gas prices of all trades in the past 20 blocks on a rolling window basis.[10] On the one hand, a too short window makes our quantile estimates noisy. For example, if we only use trades in the current block to implement the classification, two trades with very similar gas prices will fall into different categories. On the other hand, a too long window might include trades with gas prices too distant to reflect the current crowding level of the blockchain. Thus, we set the window length to be 20 in the baseline results to strike a balance.

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 A1 reports the estimation results of the cumulative return impulse responses based on DEX trade flows from the two alternative gas level classification. It shows that the results are largely unchanged compared with the baseline results in Table 5.

A.2 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 included in the structural VAR specification. Table A2 report the return impulse responses when the number of lags is set to 10 and 20 respectively. It shows that the results are qualitatively the same as the

[10]See Section 4 for details of our classification scheme.

Table A1. Cumulative impulse responses between CEX return and DEX trade flows: Gas price level classification based on a rolling window of alternative numbers of blocks. This table reports the impulse responses between the CEX return and DEX trade flows with different gas price levels based on alternative classification rule, cumulative over 20 blocks. Impulse responses are obtained by estimating the structural VAR model specified in Equation 2. The estimation is done for each pair-day and statistical inference is based on pair-day estimates. Row variables are response variables and column variables are shock variables. CEX return is in basis point. DEX trade flows are standardized and thus in their standard deviations. *, ** and *** indicate significance levels at 1%, 5% and 10% respectively.

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

PairType	Variable	r_t^{CEX}	$x_t^{\text{LowGas-DEX}}$	$x_t^{\text{MidGas-DEX}}$	$x_t^{\text{HighGas-DEX}}$
Stable-Stable	r_t^{CEX}	0.6***	0.42	0.07	0.02
		(0.01)	(0.42)	(0.11)	(0.04)
	$x_t^{\text{LowGas-DEX}}$	0.0	0.97***	−0.01	−0.03***
		(0.01)	(0.02)	(0.02)	(0.01)
	$x_t^{\text{MidGas-DEX}}$	0.0	−0.13*	0.9***	−0.07***
		(0.01)	(0.07)	(0.03)	(0.02)
	$x_t^{\text{HighGas-DEX}}$	0.01	−0.02	−0.22***	0.81***
		(0.01)	(0.1)	(0.03)	(0.02)
NonStable-Stable	r_t^{CEX}	0.97***	0.03	0.52***	3.38***
		(0.01)	(0.05)	(0.08)	(0.19)
	$x_t^{\text{LowGas-DEX}}$	0.0***	1.01***	−0.01	−0.01*
		(0.0)	(0.02)	(0.01)	(0.01)
	$x_t^{\text{MidGas-DEX}}$	0.01***	−0.08***	0.97***	0.05***
		(0.0)	(0.02)	(0.01)	(0.01)
	$x_t^{\text{HighGas-DEX}}$	0.03***	−0.13***	−0.22***	1.06***
		(0.0)	(0.01)	(0.02)	(0.01)
NonStable-NonStable	r_t^{CEX}	0.81***	1.96***	2.83***	6.73***
		(0.01)	(0.34)	(0.36)	(0.56)
	$x_t^{\text{LowGas-DEX}}$	0.0***	0.97***	−0.03**	−0.04***
		(0.0)	(0.01)	(0.01)	(0.01)
	$x_t^{\text{MidGas-DEX}}$	0.01***	−0.06***	0.98***	0.03*
		(0.0)	(0.02)	(0.01)	(0.01)
	$x_t^{\text{HighGas-DEX}}$	0.02***	−0.08***	−0.15***	1.05***
		(0.0)	(0.01)	(0.02)	(0.02)

baseline results.

(b) Gas level classification based on a rolling window of 5 blocks.

PairType	Variable	r_t^{CEX}	$x_t^{\text{LowGas-DEX}}$	$x_t^{\text{MidGas-DEX}}$	$x_t^{\text{HighGas-DEX}}$
Stable-Stable	r_t^{CEX}	0.59***	0.71	0.2	0.04
		(0.01)	(0.57)	(0.15)	(0.05)
	$x_t^{\text{LowGas-DEX}}$	0.01	0.92***	−0.04**	−0.04***
		(0.01)	(0.06)	(0.02)	(0.02)
	$x_t^{\text{MidGas-DEX}}$	0.0	−0.06	0.88***	−0.1***
		(0.01)	(0.05)	(0.02)	(0.02)
	$x_t^{\text{HighGas-DEX}}$	0.01	0.04	−0.26***	0.86***
		(0.01)	(0.13)	(0.04)	(0.02)
NonStable-Stable	r_t^{CEX}	0.97***	0.12**	0.73***	3.29***
		(0.01)	(0.05)	(0.09)	(0.19)
	$x_t^{\text{LowGas-DEX}}$	0.0***	1.0***	−0.03***	−0.01
		(0.0)	(0.01)	(0.01)	(0.01)
	$x_t^{\text{MidGas-DEX}}$	0.02***	−0.07***	0.97***	0.08***
		(0.0)	(0.01)	(0.01)	(0.01)
	$x_t^{\text{HighGas-DEX}}$	0.03***	−0.12***	−0.21***	1.03***
		(0.0)	(0.01)	(0.02)	(0.01)
NonStable-NonStable	r_t^{CEX}	0.81***	1.66***	3.6***	6.54***
		(0.01)	(0.32)	(0.36)	(0.55)
	$x_t^{\text{LowGas-DEX}}$	0.0***	0.97***	−0.03***	−0.04***
		(0.0)	(0.01)	(0.01)	(0.01)
	$x_t^{\text{MidGas-DEX}}$	0.01***	−0.06***	0.97***	0.06***
		(0.0)	(0.01)	(0.01)	(0.02)
	$x_t^{\text{HighGas-DEX}}$	0.02***	−0.1***	−0.12***	1.03***
		(0.0)	(0.01)	(0.02)	(0.01)

Table A2. Cumulative impulse responses of CEX return and DEX trade flows with different gas price levels: Alternative number of lags in the structural VAR specification. This table reports the impulse responses between the CEX return and DEX trade flow variables based on alternative numbers of lags included in the structural VAR estimation, cumulative over 20 blocks. Impulse responses are obtained by estimating the structural VAR model specified in Equation 2. The estimation is done for each pair-day and statistical inference is based on pair-day estimates. Row variables are response variables and column variables are shock variables. CEX return is in basis point. 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.

PairType	Variable	r_t^{CEX}	$x_t^{\text{LowGas-DEX}}$	$x_t^{\text{MidGas-DEX}}$	$x_t^{\text{HighGas-DEX}}$
Stable-Stable	r_t^{CEX}	0.51***	0.33	0.06	0.03
		(0.01)	(0.3)	(0.13)	(0.04)
	$x_t^{\text{LowGas-DEX}}$	0.0	0.9***	−0.08***	−0.07***
		(0.01)	(0.03)	(0.02)	(0.01)
	$x_t^{\text{MidGas-DEX}}$	0.0	−0.2***	0.83***	−0.11***
		(0.01)	(0.08)	(0.04)	(0.02)
	$x_t^{\text{HighGas-DEX}}$	0.01	−0.12***	−0.3***	0.71***
		(0.01)	(0.04)	(0.03)	(0.02)
NonStable-Stable	r_t^{CEX}	0.96***	0.0	0.54***	3.42***
		(0.01)	(0.07)	(0.11)	(0.21)
	$x_t^{\text{LowGas-DEX}}$	0.0***	1.01***	−0.02***	−0.03***
		(0.0)	(0.02)	(0.01)	(0.01)
	$x_t^{\text{MidGas-DEX}}$	0.02***	−0.09***	0.97***	0.0
		(0.0)	(0.02)	(0.01)	(0.01)
	$x_t^{\text{HighGas-DEX}}$	0.05***	−0.16***	−0.27***	1.07***
		(0.0)	(0.02)	(0.02)	(0.01)
NonStable-NonStable	r_t^{CEX}	0.77***	1.88***	3.5***	6.89***
		(0.01)	(0.41)	(0.44)	(0.65)
	$x_t^{\text{LowGas-DEX}}$	0.0***	0.97***	0.01	−0.05***
		(0.0)	(0.02)	(0.02)	(0.02)
	$x_t^{\text{MidGas-DEX}}$	0.01***	−0.07***	0.95***	−0.01
		(0.0)	(0.02)	(0.01)	(0.02)
	$x_t^{\text{HighGas-DEX}}$	0.03***	−0.1***	−0.2***	1.05***
		(0.0)	(0.02)	(0.02)	(0.02)

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

PairType	Variable	r_t^{CEX}	$x_t^{\text{LowGas-DEX}}$	$x_t^{\text{MidGas-DEX}}$	$x_t^{\text{HighGas-DEX}}$
Stable-Stable	r_t^{CEX}	0.4***	0.26	0.06	0.11
		(0.02)	(0.25)	(0.12)	(0.08)
	$x_t^{\text{LowGas-DEX}}$	−0.02	0.76***	−0.24***	−0.13***
		(0.02)	(0.16)	(0.09)	(0.05)
	$x_t^{\text{MidGas-DEX}}$	0.0	−0.29***	0.75***	−0.22***
		(0.02)	(0.1)	(0.05)	(0.03)
	$x_t^{\text{HighGas-DEX}}$	0.02	−0.22***	−0.29***	0.68***
		(0.02)	(0.04)	(0.1)	(0.05)
NonStable-Stable	r_t^{CEX}	0.94***	0.04	0.57***	3.28***
		(0.01)	(0.11)	(0.14)	(0.2)
	$x_t^{\text{LowGas-DEX}}$	0.01***	1.01***	−0.02*	−0.05***
		(0.0)	(0.02)	(0.01)	(0.01)
	$x_t^{\text{MidGas-DEX}}$	0.03***	−0.09***	0.95***	−0.02**
		(0.0)	(0.02)	(0.01)	(0.01)
	$x_t^{\text{HighGas-DEX}}$	0.08***	−0.19***	−0.35***	1.05***
		(0.0)	(0.02)	(0.02)	(0.01)
NonStable-NonStable	r_t^{CEX}	0.71***	1.32***	3.89***	6.86***
		(0.02)	(0.54)	(0.6)	(0.61)
	$x_t^{\text{LowGas-DEX}}$	0.0***	0.95***	−0.03	−0.06***
		(0.0)	(0.02)	(0.02)	(0.02)
	$x_t^{\text{MidGas-DEX}}$	0.02***	−0.11***	0.94***	0.0
		(0.0)	(0.02)	(0.02)	(0.02)
	$x_t^{\text{HighGas-DEX}}$	0.04***	−0.14***	−0.25***	1.0***
		(0.0)	(0.02)	(0.03)	(0.03)

References

Aoyagi, Jun and Yuki Ito (2021). “Coexisting Exchange Platforms: Limit Order Books and Automated Market Makers”. In: *SSRN Electronic Journal* February, pp. 1–53.

Back, Kerry, C. Henry Cao, and Gregory A. Willard (2000). “Imperfect competition among informed traders”. In: *Journal of Finance* 55.5, pp. 2117–2155.

Barbon, Andrea and Angelo Ranaldo (2021). “On The Quality Of Cryptocurrency Markets”. In: *SSRN Electronic Journal* July.

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

Biais, Bruno, Thierry Foucault, and Sophie Moinas (2015). “Equilibrium fast trading”. In: *Journal of Financial Economics* 116.2, pp. 292–313.

Block, The (2022). *Exchange archives*.

Brunnermeier, Markus K. and Lasse H. Pedersen (2005). “Predatory trading”. In: *Journal of Finance* 60.4, pp. 1825–1864.

Budish, Eric B., Peter Cramton, and John J. Shim (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 (2021). “The Adoption of Blockchain-based Decentralized Exchanges”. In: pp. 1–73.

Easley, David and Maureen O’Hara (1987). “Price, trade size, and information in securities markets”. In: *Journal of Financial Economics* 19.1, pp. 69–90.

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.

Foucault, Thierry, Johan Hombert, and Ioanid Roşu (2016). "News Trading and Speed". In: *Journal of Finance* 71.1, pp. 335–382.

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

— (1991b). "The Summary Informativeness of Stock Trades: An Econometric Analysis". In: *Review of Financial Studies* 4.3, pp. 571–595.

Hendershott, Terrence J. and Albert J. Menkveld (2014). "Price pressures". In: *Journal of Financial Economics* 114.3, pp. 405–423.

Hoffmann, Peter (2014). "A dynamic limit order market with fast and slow traders". In: *Journal of Financial Economics* 113.1, pp. 156–169.

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

Jovanovic, Boyan and Albert J. Menkveld (2016). "Middlemen in Limit-Order Markets".

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. (1985). "Continuous Auctions and Insider Trading". In: *Econometrica* 53.6, p. 1315.

O'Hara, Maureen, Chen Yao, and Mao Ye (2014). "What's not there: Odd lots and market data". In: *Journal of Finance* 69.5, pp. 2199–2236.

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

Parlour, Christine A. and Alfred Lehar (2021). "Decentralized Exchanges". In: *SSRN Electronic Journal* 2022.

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