Price Discovery on Decentralized Exchanges

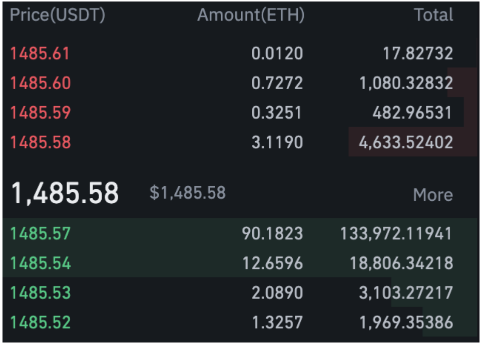
Agostino Capponi[1], Ruizhe Jia[1], Shihao Yu[1]

[1]Columbia University, IEOR

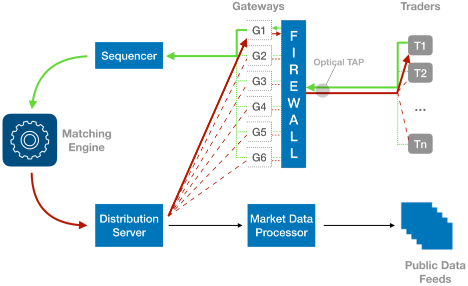
Aug 18, 2023 EFA

Centralized exchanges (CEXs)

- central limit order book (CLOB) + continuous execution + ex-post transparency of market orders
- Traders compete on speed (Budish, Cramton, and Shim, 2015).



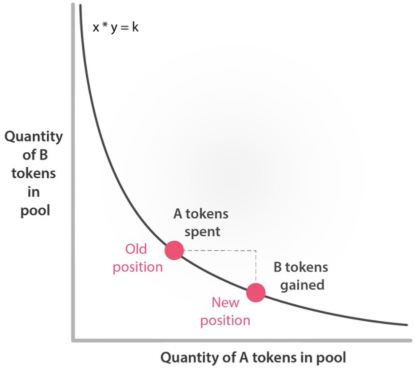
Price(USDT)	Amount(ETH)	Total
1485.61	0.0120	17.82732
1485.60	0.7272	1,080.32832
1485.59	0.3251	482.96531
1485.58	3.1190	4,633.52402
1,485.58	$1,485.58	More
1485.57	90.1823	133,972.11941
1485.54	12.6596	18,806.34218
1485.53	2.0890	3,103.27217
1485.52	1.3257	1,969.35386

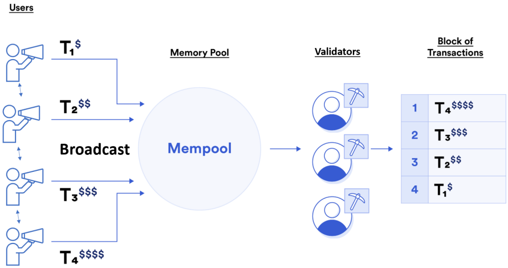


Binance order bookLSE co-location schematic. Source: Aquilina, Budish, and O'Neill (2021)

Decentralized exchanges (DEXs)

- on-chain ⇒ automated market makers (AMM) + batch execution in descending order of priority fees + ex-ante transparency of market orders
- Traders compete on **priority fees**





Uniswap bonding curveEthereum schematic

Paper in a nutshell

- How does price discovery realize on DEXs?
  - Do priority fees reveal private information?
    - Execution speed/competition versus information leakage (Yang and Zhu, 2020)
  - If so, why do informed traders bid high fees?
    - Execution risk versus competition on private information
  - How do they bid?
    - English open-cry or other forms of bidding strategy

- Main results:
  - High-fee DEX trade flow more informative → priority fees reveal trades' **private** information, contributing to price discovery.
  - Informed traders bid high fees not only to avoid execution risk but also to **compete with each other**.
  - Competition among informed traders mostly follows the **jump bidding** strategy.

Literature

- Public trade characteristics and private information
  - Size (Easley and O'Hara, 1987; O'Hara, Yao, and Ye, 2014); execution venue (Barclay, Hendershott, and McCormick, 2003)
  - **Contribution**: identify priority fee as a new public signal on DEXs
- Informed trading and price discovery
  - Competitive informed traders (Holden and Subrahmanyam, 1992; Foster and Viswanathan, 1996; Back, Cao, and Willard, 2000); impatient informed traders (Caldentey and Stacchetti, 2010; Kaniel and Liu, 2006); liquidity timing (Collin-Dufresne and Fos, 2015).
  - **Contribution**: analyze competition among informed traders on DEXs
- DEXs
  - Impact of priority fees on DEX liquidity (Park, 2021; Capponi and Jia, 2021; Barbon and Ranaldo, 2022; Hasbrouck, Rivera, and Saleh, 2022; Lehar, Parlour, and Zoican, 2022; Foley, O'Neill, and Putniņš, 2023); Impact of trust in DEXs on prices (Han, Huang, and Zhong, 2022)
  - **Contribution**: analyze the role of priority fees in price discovery

Data

- The four most traded crypto pairs (ETH-USDT, WBTC-ETH, LINK-ETH, AAVE-ETH) between November 18, 2020, and February 10, 2021
- **Executed trades** on the largest CEX (Binance) and the largest DEX (Uniswap)
  - Both: timestamp (to 1s), trade size, trade direction
  - Uniswap: hash, submission address, nonce, gas price, block number
- **Binance order-book updates**
  - Used to build order books and calculate midquote returns
  - Timestamped to the precision of 1s
- **Ethereum mempool orders**
  - Trades and their order history linked by submission address + nonce
  - Think of nonce as trader-specific order ID

Methodology: structural VAR Model

- A structural VAR model (Hasbrouck, 1991a; Hasbrouck, 1991b):

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

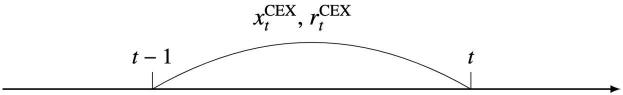
  - where

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

  - $x_t^{\text{HighFee-DEX}}$: the signed flow of high-fee DEX trades in block t, $>$ Q75 of a rolling window of past 20 non-empty blocks. $x_t^{\text{LowFee-DEX}}$: Q25.
  - A: structural matrix capturing the *contemporaneous* relations between the endogenous variables. To be specified

Timestamp convention

- DEXs run in *block* time while CEXs run in *continuous* time
- Need to harmonize the two clocks:
  - r_t^{CEX} and x_t^{CEX} are aggregated based on block time



- x_t^{DEX}. Note that DEX orders submitted during the block time interval of $(t-1, t]$ and picked by validators are all executed at the same block time t, in descending order by their fees.

Two information measures

- **Permanent price impact (PPI)**: cumulative impulse responses of CEX return to trade shocks:

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

- **Information share (IS)**: normalized PPI weighted by its innovation variance

$$\text{IS}_k = \frac{[\Theta_k^b(1)]^2 \sigma_{\varepsilon_k}^2}{\sigma_w^2} \tag{1}$$

- Follow Hasbrouck (1995) to resolve the contemporaneous relations, i.e., use Cholesky decomposition of the error variance to obtain lower and upper bounds

Do informed traders bid high fees?

- Key results:
  - PPI of high-fee DEX trade flow: 4.27 - 8.16 bps; low-fee: 0.41 - 0.94 bps
  - Price discovery on DEX realizes as traders bid high fees to trade their private information.

Variable	$x^{LowFee\text{-}DEX}$		$x^{MidFee\text{-}DEX}$		$x^{HighFee\text{-}DEX}$		$\Delta^{HighFee\ -\ LowFee}$
	LB	UB	LB	UB	LB	UB	LB - UB
ETH-USDT	-0.1^{**}	-0.04	0.27^{***}	0.57^{***}	2.62^{***}	4.15^{***}	2.66^{***}
	(0.05)	(0.05)	(0.07)	(0.1)	(0.31)	(0.42)	(0.31)
LINK-ETH	0.37^{***}	0.71^{***}	1.86^{***}	3.45^{***}	4.43^{***}	7.76^{***}	3.72^{***}
	(0.14)	(0.16)	(0.25)	(0.34)	(0.43)	(0.54)	(0.43)
WBTC-ETH	0.17^{*}	0.31^{***}	0.84^{***}	1.51^{***}	3.24^{***}	4.59^{***}	2.93^{***}
	(0.1)	(0.11)	(0.14)	(0.2)	(0.32)	(0.41)	(0.32)
AAVE-ETH	1.16^{***}	2.74^{***}	4.67^{***}	9.81^{***}	6.71^{***}	15.91^{***}	3.97^{***}
	(0.34)	(0.39)	(0.49)	(0.6)	(0.52)	(0.7)	(0.63)
Pooled	0.41^{***}	0.94^{***}	1.93^{***}	3.86^{***}	4.27^{***}	8.16^{***}	3.33^{***}
	(0.1)	(0.12)	(0.17)	(0.27)	(0.22)	(0.37)	(0.22)

Information share of DEX trade flows

- Key result:
  - IS of high-fee DEX trade flow: 6.73 - 18.06%; low-fee: 0.55 - 0.81 %

Variable	r^{CEX}		$x^{LowFee\text{-}DEX}$		$x^{MidFee\text{-}DEX}$		$x^{HighFee\text{-}DEX}$		$\Delta^{HighFee - LowFee}$
	LB	UB	LB	UB	LB	UB	LB	UB	LB - UB
ETH-USDT	84.92	93.58	0.15	0.17	0.28	0.59	5.98	14.37	5.81***
	(1.29)	(0.77)	(0.02)	(0.03)	(0.04)	(0.09)	(0.76)	(1.27)	(0.76)
LINK-ETH	76.72	90.73	0.52	0.72	1.63	4.18	7.06	18.55	6.34***
	(1.69)	(0.96)	(0.08)	(0.1)	(0.23)	(0.45)	(0.84)	(1.49)	(0.83)
WBTC-ETH	79.34	88.83	0.56	0.81	1.08	2.49	9.45	17.42	8.63***
	(1.41)	(1.07)	(0.08)	(0.12)	(0.18)	(0.33)	(1.02)	(1.3)	(1.03)
AAVE-ETH	67.54	91.03	0.96	1.55	3.05	9.47	4.63	22.05	3.08***
	(1.26)	(0.62)	(0.14)	(0.21)	(0.44)	(0.78)	(0.43)	(1.19)	(0.51)
Pooled	77.09	91.08	0.55	0.81	1.52	4.17	6.73	18.06	5.92***
	(0.79)	(0.44)	(0.05)	(0.07)	(0.14)	(0.3)	(0.4)	(0.67)	(0.41)

Robustness

- Three main robustness checks:
  1. Controlling for CEX trade flows ◂ Robustness CEX trade flow
  2. Controlling for the trade size ◂ Robustness trade size
  3. Use unstandardized trade flows ◂ Robustness unstandardized trade flows
- Key results qualitatively the same.

Why do informed traders bid high fees?

- **Execution risk due to blockchain congestion**
  - Other blockchain users for non-DEX related activities (e.g., NFT auctions, ICO issuance)
  - **Implication**: Fee bids by informed traders just sufficiently high to get into the block
- **Competition among informed traders**
  - Information leakage (e.g., "back-runners" Yang and Zhu (2020))
  - Or multiple traders receive correlated private signals (See, e.g., Holden and Subrahmanyam, 1992; Foster and Viswanathan, 1996; Back, Cao, and Willard, 2000)
  - **Implication**: Fee bids by informed traders way higher than those of non-DEX transactions in the same block, because informed traders' goal is to get first execution.

Identify competition trades

- Identify competition trades: a standard outliers detection approach.

$$fee > \text{Q75} + 1.5 \times \underbrace{\text{IQR}}_{\text{Q75 - Q25 based on all Txns}}$$

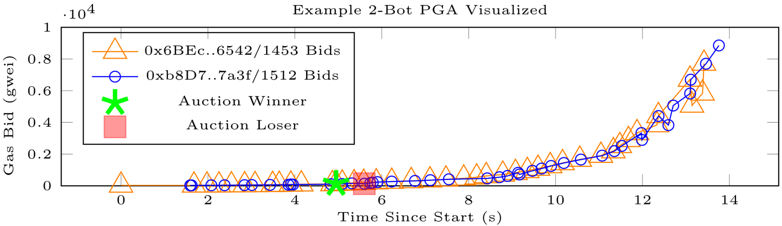
- Competition on private information? Public? Or uninformed liquidation trades?

- Exclude competition trades and re-implement the structural VAR. PPI of high-fee DEX trade flow drops significantly: **4.27** - **8.15 to 2.83** - **5.36** bps

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

How do informed traders bid?

- How do informed traders compete?
- Priority gas auctions (PGAs)? Arbitrageurs competitively bid up their blockchain fees (Daian et al., 2020)



Is PGA a dominant bidding strategy?

- Impose three minimal requirements for a PGA trade:
  1. At least **one** matched order with the same submission address and nonce.
  2. It has a higher gas price than its matched order(s).
  3. All matched orders must arrive at the mempool between $t-5$ and t.
- Surprisingly, even for competition trades, the fraction of PGA trades is very small (<25%).

TokenPair	ExplicitCompetition ExcessiveGas	Non-PGA trades	PGA trades
ETH-USDT	Other trades	97.68	2.32
	Excessively-high-fee trades	87.95	12.05
LINK-ETH	Other trades	95.06	4.94
	Excessively-high-fee trades	73.29	26.71
WBTC-ETH	Other trades	96.61	3.39
	Excessively-high-fee trades	84.79	15.21
AAVE-ETH	Other trades	95.39	4.61
	Excessively-high-fee trades	81.72	18.28

Jump bidding

- Informed traders bid an ex-ante high blockchain fee once, following a **jump bidding** strategy (Daniel and Hirshleifer, 1998; Avery, 1998).
- A signaling tool to communicate with other traders, discouraging competition.
- Jump bidding is a unique feature of competition on DEXs.
  - On CEXs, there is no counterpart to such a signaling tool
  - Informed traders can choose to use market orders to have a faster execution, but can not signal their high valuation.

Conclusion

- Unique trading mechanism of DEXs: orders processed in batches + execution priority determined by priority fees.
- High-fee DEX trade flow more informed than low-fee.
  - Price discovery on DEX partials realizes through traders bidding high fees to trade their private information.
- We test potential economic channels using a unique data set of Ethereum mempool orders.
  - Informed traders bid high blockchain fees not only to reduce execution risk arising from blockchain congestion but also to **compete with each other**.
  - The competition among informed traders follows a **jump bidding strategy**.

Thank You!

Robustness: Controlling for CEX trade flow

◂ Robustness overview

- Potential problem with the baseline specification: ◂ Baseline SVAR specification
  - Informed traders might split their trades between CEX and DEX. Thus CEX trade flow and DEX trade flow might be correlated.
- An alternative specification controlling for CEX trade flow:

$$y_t = \begin{pmatrix} r_t^{\text{CEX}} & x_t^{\text{CEX}} & x_t^{\text{LowFee-DEX}} & x_t^{\text{MidFee-DEX}} & x_t^{\text{HighFee-DEX}} \end{pmatrix} \quad (2)$$

Robustness: Controlling for CEX trade flow

Robustness overview

- Key results that high-fee DEX trade flow has a much larger permanent price impact than mid-fee and low-fee DEX trade flow remains.

Variable	x^{CEX}		$x^{LowFee-DEX}$		$x^{MidFee-DEX}$		$x^{HighFee-DEX}$		$\Delta^{HighFee - LowFee}$
	LB	UB	LB	UB	LB	UB	LB	UB	LB - UB
ETH-USDT	0.93***	3.89***	−0.1*	−0.03	0.24***	0.56***	2.53***	4.12***	2.56***
	(0.13)	(0.24)	(0.05)	(0.05)	(0.07)	(0.1)	(0.3)	(0.42)	(0.31)
LINK-ETH	0.52***	2.87***	0.36***	0.69***	1.87***	3.45***	4.44***	7.75***	3.75***
	(0.21)	(0.3)	(0.14)	(0.16)	(0.25)	(0.34)	(0.43)	(0.54)	(0.43)
WBTC-ETH	2.12***	5.15***	0.17*	0.3***	0.74***	1.44***	3.08***	4.53***	2.78***
	(0.31)	(0.54)	(0.1)	(0.11)	(0.13)	(0.19)	(0.29)	(0.39)	(0.31)
AAVE-ETH	1.09**	5.67***	1.1***	2.7***	4.68***	9.81***	6.69***	15.92***	3.98***
	(0.54)	(0.62)	(0.35)	(0.39)	(0.49)	(0.6)	(0.54)	(0.7)	(0.67)
Pooled	1.15***	4.38***	0.39***	0.92***	1.91***	3.84***	4.2***	8.14***	3.28***
	(0.17)	(0.23)	(0.1)	(0.13)	(0.17)	(0.27)	(0.22)	(0.37)	(0.23)

Robustness: Controlling for trade size

◂ Robustness overview

- Blockchain fee is a fixed cost and traders bid higher fees for larger trades as it is relatively cheaper. In addition, literature shows that larger trades are more informed (Easley and O'Hara, 1987). Thus, trade size might be a confounding factor.

- To alleviate the concern, we further group DEX trades by their trade size, in addition to fee. Specifically, we classify DEX trades into two size groups: a large-size group consisting of trades with a size above its 90% quantile and a small-size group consisting of trades with a size below its 90% quantile.

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

Robustness: Controlling for trade size

◂ Robustness overview

- Key results that high-fee DEX trade flow has a much larger permanent price impact than mid-fee and low-fee DEX trade flow remains.

Variable	$_x$S-L-DEX		$_x$S-M-DEX		$_x$S-H-DEX		$_x$L-L-DEX		$_x$L-M-DEX		$_x$L-H-DEX	
	LB	UB	LB	UB	LB	UB	LB	UB	LB	UB	LB	U
ETH-USDT	−0.02	0.0	−0.04	0.04	0.13***	0.31***	−0.09*	−0.03	0.27***	0.58***	2.63***	4.
	(0.05)	(0.05)	(0.05)	(0.05)	(0.04)	(0.05)	(0.05)	(0.05)	(0.07)	(0.1)	(0.31)	(0.
LINK-ETH	−0.11	−0.02	−0.02	0.06	0.0	0.04	0.37***	0.71***	1.87***	3.46***	4.42***	7.
	(0.14)	(0.13)	(0.14)	(0.15)	(0.13)	(0.14)	(0.15)	(0.17)	(0.25)	(0.35)	(0.44)	(0.
WBTC-ETH	0.27	0.41	0.78***	1.45***	1.72***	2.73***	0.0	0.17	0.18	0.58***	1.6***	2.
	(0.23)	(0.26)	(0.25)	(0.3)	(0.35)	(0.46)	(0.18)	(0.21)	(0.13)	(0.18)	(0.36)	(0.
AAVE-ETH	−0.16	0.45	0.68*	1.63***	1.16***	2.32***	1.18***	2.73***	4.59***	9.81***	6.57***	15.
	(0.37)	(0.4)	(0.37)	(0.36)	(0.47)	(0.49)	(0.34)	(0.39)	(0.49)	(0.59)	(0.5)	(0.
Pooled	−0.07	0.11	0.2	0.58***	0.44***	0.88***	0.49***	1.14***	2.25***	4.6***	4.54***	9.
	(0.13)	(0.14)	(0.14)	(0.14)	(0.17)	(0.18)	(0.13)	(0.16)	(0.22)	(0.33)	(0.26)	(0.

Robustness: Use Raw Level trade flows

◂ Robustness overview

- Key results that high-fee DEX trade flow has a much larger permanent price impact than mid-fee and low-fee DEX trade flow remains.

Variable	$x^{\text{LowFee-DEX}}$		$x^{\text{MidFee-DEX}}$		$x^{\text{HighFee-DEX}}$		$\Delta^{\text{HighFee - LowFee}}$
	LB	UB	LB	UB	LB	UB	LB - UB
ETH-USDT	-0.01^{**}	0.0	0.01^{***}	0.03^{***}	0.08^{***}	0.13^{***}	0.09^{***}
	(0.01)	(0.01)	(0.0)	(0.0)	(0.01)	(0.01)	(0.01)
LINK-ETH	0.08^{***}	0.15^{***}	0.15^{***}	0.28^{***}	0.29^{***}	0.51^{***}	0.14^{***}
	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)	(0.04)
WBTC-ETH	0.01	0.02^{**}	0.02^{***}	0.04^{***}	0.09^{***}	0.13^{***}	0.07^{***}
	(0.01)	(0.01)	(0.0)	(0.0)	(0.01)	(0.01)	(0.01)
AAVE-ETH	0.24^{***}	0.57^{***}	0.38^{***}	0.83^{***}	0.54^{***}	1.26^{***}	-0.03
	(0.07)	(0.08)	(0.04)	(0.05)	(0.04)	(0.06)	(0.09)
Pooled	0.08^{***}	0.18^{***}	0.14^{***}	0.3^{***}	0.25^{***}	0.52^{***}	0.07^{***}
	(0.02)	(0.03)	(0.01)	(0.02)	(0.02)	(0.03)	(0.03)

Daily trading summary statistics on Uniswap and Binance

Data description

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

Summary statistics of return and trade flow variables

		N	Mean	SD	Min	50%	Max
ETH-USDT	$r_t^{\mathbf{CEX}}$	370291	0.03	10.27	-476.61	0.00	368.22
	$x_t^{\mathbf{CEX}}$	370291	-0.32	221.19	-7370.94	0.11	10152.33
	$x_t^{\mathbf{DEX}}$	370291	0.15	40.76	-3111.34	0.04	2154.22
	$x_t^{\mathbf{LowFee\text{-}DEX}}$	370291	-0.03	10.29	-2345.49	0.00	1241.70
	$x_t^{\mathbf{MidFee\text{-}DEX}}$	370291	-0.06	21.37	-1897.53	0.00	2147.57
	$x_t^{\mathbf{HighFee\text{-}DEX}}$	370291	0.23	33.18	-3498.28	0.00	2217.48
WBTC-ETH	$r_t^{\mathbf{CEX}}$	81892	-0.05	9.12	-269.32	0.00	245.93
	$x_t^{\mathbf{CEX}}$	81892	-0.02	9.93	-395.21	0.00	1991.97
	$x_t^{\mathbf{DEX}}$	81892	-0.25	56.17	-2750.21	0.22	2331.24
	$x_t^{\mathbf{LowFee\text{-}DEX}}$	81892	0.07	15.87	-475.92	0.00	698.13
	$x_t^{\mathbf{MidFee\text{-}DEX}}$	81892	0.07	36.64	-2750.21	0.00	726.66
	$x_t^{\mathbf{HighFee\text{-}DEX}}$	81892	-0.40	39.13	-771.15	0.00	2331.24

Summary statistics of return and trade flow variables (Cont.)

		N	Mean	SD	Min	50%	Max
LINK-ETH	r_t^{CEX}	72951	-0.07	16.10	-494.76	0.00	467.55
	x_t^{CEX}	72951	-0.47	16.73	-2047.56	0.00	432.04
	x_t^{DEX}	72951	-0.08	22.57	-1187.08	0.00	652.36
	$x_t^{\text{LowFee-DEX}}$	72951	-0.04	5.32	-202.07	0.00	161.16
	$x_t^{\text{MidFee-DEX}}$	72951	-0.10	14.47	-1187.08	0.00	652.36
	$x_t^{\text{HighFee-DEX}}$	72951	0.06	16.11	-432.35	0.00	541.94
AAVE-ETH	r_t^{CEX}	42975	0.14	29.89	-509.77	0.00	582.37
	x_t^{CEX}	42975	-0.31	10.83	-676.27	0.00	239.78
	x_t^{DEX}	42975	0.14	19.59	-417.79	0.10	374.95
	$x_t^{\text{LowFee-DEX}}$	42975	0.07	5.51	-150.28	0.00	225.81
	$x_t^{\text{MidFee-DEX}}$	42975	0.02	12.78	-417.79	0.00	192.39
	$x_t^{\text{HighFee-DEX}}$	42975	0.05	13.75	-221.06	0.00	374.95

Implementation details

◂ Baseline specification

- Model estimated at block-by-block frequency
- We set the number of lags in the structural VAR model to 5. In the Appendix, we change the number of lags included in the structural VAR model to 10 and 20, and show that estimation results remain qualitatively the same.
- As the base currency varies across token pairs, to ease comparison and aggregation across token pairs, we standardize all trade flow variables such that they have zero mean and unit variance.
- Hence, the impulse responses reported below should be interpreted as permanent price impacts in basis points per standard deviation increase in the trade flow.