Measuring CEX-DEX Extracted Value and Searcher Profitability: The Darkest of the MEV Dark Forest

Fei Wu
King's College London, UK

Danning Sui
Flashbots, San Francisco, CA, USA

Thomas Thiery
Ethereum Foundation, Lisbon, Portugal

Mallesh Pai
Rice University, Houston, TX, USA
Tempo Labs, Houston, TX, USA
Paradigm, Houston, TX, USA

Abstract

This paper provides a comprehensive empirical analysis of the economics and dynamics behind arbitrages between centralized and decentralized exchanges (CEX-DEX) on Ethereum. We refine heuristics to identify arbitrage transactions from on-chain data and introduce a robust empirical framework to estimate arbitrage revenue without knowing traders' actual behaviors on CEX. Leveraging an extensive dataset spanning 19 months from August 2023 to March 2025, we estimate a total of 233.8M USD extracted by 19 major CEX-DEX searchers from 7,203,560 identified CEX-DEX arbitrages. Our analysis reveals increasing centralization trends as three searchers captured three-quarters of both volume and extracted value. We also demonstrate that searchers' profitability is tied to their integration level with block builders and uncover exclusive searcher-builder relationships and their market impact. Finally, we correct the previously underestimated profitability of block builders who vertically integrate with a searcher. These insights illuminate the darkest corner of the MEV landscape and highlight the critical implications for Ethereum's decentralization.

2012 ACM Subject Classification Applied computing → Electronic commerce

Keywords and phrases Decentralized Finance, Maximal Extractable Value, CEX-DEX arbitrages

Digital Object Identifier 10.4230/LIPIcs.AFT.2025.26

Related Version *Full Version*: `arXiv:2507.13023`

Supplementary Material *Dataset*: `https://dune.com/rig_ef/cex-dex-dash`

Funding The work is supported by The Latest in DeFi Research (TLDR) Fellowship 2024 funded by Uniswap Foundation.

Acknowledgements We appreciate Gaussian Process, Kevin Pang, Bill Zhang, Tom Zhao, Romain Butteaud, Burak Öz and Max Resnick for for helpful discussions and comments. We also thank Alexander Tesfamichael from Ultra Sound Money and Žan Knafelc from Eden Network for their help with data collection.

1 Introduction

Over recent years, the focus of study around Maximal Extractable Value (MEV) extraction has broadened beyond on-chain atomic strategies to a richer landscape of more sophisticated strategies across different venues. Early studies primarily investigated MEV extracted solely from the blockchain's internal state and mempool, termed atomic MEV [19, 44, 45]. Subsequent research has turned attention towards non-atomic MEV, strategies that leverage

external price information alongside on-chain execution, most notably arbitrage between centralized exchanges (CEX) and decentralized exchanges (DEX) [42, 34, 62]. These CEX-DEX arbitrages have emerged as an economically substantial type of MEV.

CEX-DEX arbitrageurs (searchers) capitalize on temporary price discrepancies arising from asynchronous price discovery across venues. While centralized exchanges provide high liquidity and near-instantaneous trade execution, decentralized exchanges experience liquidity constraints and inherent latency due to blockchain settlement delays. These structural frictions regularly produce significant, short-lived deviations in asset prices on DEXes from their fair value, offering lucrative arbitrage opportunities for sophisticated market participants.

Despite high entry barriers such as capital requirements, low-latency infrastructure, inventory risk, and uncertainty of block inclusion, CEX-DEX arbitrage remains remarkably profitable [48, 17]. Such trades often take up less than 2% of block space but contribute more than 15% of the total block value [34, 43]. Consequently, under Ethereum's Proposer-Builder Separation (PBS) framework, block builders with privileged access to such arbitrage flows benefit from a decisive advantage in winning block auctions. At the time of writing, three builders `beaverbuild`, `Titan`, and `rsync` dominate the Ethereum builder market, two of which vertically integrate their own CEX-DEX searchers [43, 34, 58]. Such vertical integration raises important concerns for Ethereum's decentralization and security: it fosters economies of scale that strengthen dominant players, enables monopoly pricing that causes proposer loss [58], and increases vulnerability to censorship and commitment attacks [53, 33, 30].

Although theoretical models [40, 39, 41] and prior empirical studies [34, 43, 58, 17] have illuminated the existence and broader impact of CEX-DEX arbitrage on the Ethereum builder market, the economic details remain unclear. Accurately identifying such transactions on-chain and estimating realized revenue without knowing searchers' off-chain behaviors is particularly challenging, rendering CEX-DEX arbitrage arguably "the darkest part of the MEV dark forest."

This paper provides a rigorous empirical investigation into value extraction, searcher profitability, and market structure effects of CEX-DEX arbitrage on Ethereum. Integrating DEX trades data, CEX pricing, and MEV-Boost data over 19 months from August 2023 to March 2025 and quantifying *who* earns *what*, our results supply the missing inputs needed to reason about decentralization guarantees. Our contributions are summarized as follows:

1. We refine existing heuristics for identifying CEX-DEX arbitrage transactions [34, 43], significantly expanding coverage and accuracy.
2. Without directly observing the CEX leg of searchers' CEX-DEX arbitrages, we develop an empirical framework to infer the likely CEX execution by tracking when each searcher's information advantage maximizes and begins to erode, thereby inferring from each on-chain swap a realized arbitrage revenue. This data-driven proxy matches theoretical price-impact intuition and enables consistent comparisons between CEX-DEX searchers.
3. We demonstrate that token liquidity influences how searchers extract value and hedge: high-liquidity tokens enable searchers to place large orders at tight spreads with minimal market impact, whereas low-liquidity tokens present wider spreads but restrict them to smaller trades that incur larger price impact.
4. We estimate that 19 major searchers extracted a total value of 233.8M USD from 7,203,560 CEX-DEX arbitrages during the observed period and uncover an increasing centralization trend, with the three leading searchers capturing approximately three-quarters of the total arbitrage volume and extracted value.
5. Our results reveal a clear profit-sharing pattern tied to searchers' integration levels with block builders. Neutral searchers, distributing flow among multiple builders, retain higher profit margins. Conversely, exclusive searchers maintain lower margins, sharing

most of their revenue with affiliated builders, sometimes even operating at negative net profit. Additionally, we find that the searcher-builder exclusivity deal is likely to mutually reinforce competitive positions for both the searcher and the builder.

6. By accounting for searcher profits, we correct previously understated estimates of integrated builder profits and subsidies and better illuminate their profitability.

2 Background and Related Works

Ethereum, Proposer-Builder Separation, and MEV-Boost. Ethereum advances in discrete 12-second slots. At each slot, a randomly selected validator – the beacon proposer – publishes a new block. To enhance validator decentralization and censorship resistance, Proposer-Builder Separation (PBS) allows proposers to outsource block construction tasks to specialized entities called *builders*. PBS is currently implemented via MEV-Boost [27], an out-of-protocol solution where builders compete in MEV-Boost auctions by submitting blocks alongside bids to trusted intermediaries known as *relays*. Builders submit bids for slot n starting from the start of slot $n-1$, funded by tips and payments alongside transactions from users and MEV searchers. At the start of slot n, the proposer signs the highest-bid block header from any relay to which the proposer is connected, after which the winning relay publishes the block.

Maximal Extractable Value and its supply chain. Maximal Extractable Value (MEV) refers to profits obtainable through strategic transaction ordering, inclusion, or censorship within a block. MEV extraction can rely solely on internal blockchain state (atomic MEV) [19, 45, 61, 44], or incorporate external market information (non-atomic MEV), such as prices from centralized exchanges and decentralized exchanges on other blockchains [34, 62, 42].

Under PBS, MEV extraction is typically performed by specialized searchers. The competitive nature of MEV extraction gives rise to a structured MEV supply chain: searchers share with builders part of their extracted value to ensure prioritized inclusion of their transactions. Builders subsequently share a portion of these MEV revenues with proposers through competitive bids in MEV-Boost auctions. Such intense competition incentivizes *vertical integration*, where entities optimize MEV extraction by simultaneously controlling upstream (searcher-level strategies) and downstream (builder and relay infrastructure) activities, thus securing a strategic advantage in capturing valuable MEV opportunities.

CEX-DEX Arbitrages. CEX-DEX arbitrages represent a prominent category of non-atomic MEV. Typically, these arbitrages involve two complementary trades: one executed on-chain via a decentralized exchange, and the other executed off-chain via a centralized exchange, in this order. Figure 1 illustrates an example of a CEX-DEX operation. First, searchers identify price discrepancies between assets listed on a DEX and a CEX. Based on the observed price differences, searchers form an *ex-ante expectation* of the extractable value (i.e., arbitrage revenue) once hedged. They then initiate the arbitrage by submitting their DEX trade to builders, offering a fraction of their expected arbitrage revenue as payments to the builder to secure prioritized block inclusion. Once the DEX trade is confirmed on-chain, searchers swiftly execute an offsetting hedge trade on the CEX to realize the profit.

Related Works. [40, 39, 41] introduced the concept of Loss-Versus-Rebalancing (LVR), modeling losses of DEX liquidity providers due to informed arbitrageurs, establishing a theoretical benchmark for CEX-DEX arbitrage revenue. [18] characterized the extractable value by informed searchers and showed that MEV extraction grows sophisticated with more

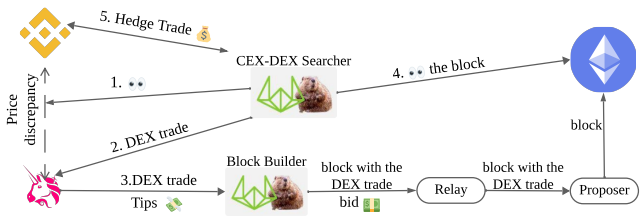


Figure 1 An example of a CEX-DEX arbitrage operation.

informed trading, notably CEX-DEX arbitrages. The authors further elaborated on the risks and entry barriers of CEX-DEX arbitrage operation and observed that CEX-DEX arbitrage primarily involves trading high-liquidity tokens [17]. [16] provided initial insights into CEX-DEX arbitrages and searcher-builder relationships, and estimated searcher profits from trades in the Uniswap V2 WETH-USDC pool. Expanding this, [47] investigated CEX-DEX arbitrage activities in all Uniswap V2 and V3 pools, highlighting searcher-builder integration and economies of scale in the Ethereum builder market. [34] empirically studied CEX-DEX arbitrages identified through comprehensive heuristics they introduce, uncovering a strong correlation between arbitrage volumes and token price volatility, and demonstrating builders integrated with CEX-DEX searchers perform better in MEV-Boost auctions during volatile periods. The work also provided evidence for vertical integration between specific searchers and builders, such as `beaverbuild` and `rsync`. Similarly, [32] revealed how builders operated by high-frequency trading firms leverage access to CEX-DEX arbitrage flows to gain competitive advantages in block building. [43] empirically demonstrated that builders with exclusive providers, such as CEX-DEX searchers, enjoy increased market share and profitability. Similar to [34], the work found that CEX-DEX arbitrage flow contributes to a prominent share of the block value. [57, 56] conducted simulations and empirical game-theoretic analysis to show builders' strategic advantages in block auctions with exclusive access to high-value orderflow. [58] further analyzed how vertical integration adversely impacts proposer revenue, block optimality, and censorship resistance. Lastly, longitudinal studies by [54, 33] provided empirical overviews of Ethereum's evolving builder market, highlighting centralization and censorship pressures under PBS.

3 Identifying CEX-DEX Arbitrage Transactions

CEX-DEX arbitrage transactions are typically more challenging to identify than other types of MEV transactions [61, 45, 44], primarily because only the DEX leg of such arbitrages is explicitly observable on-chain. Nonetheless, by exploiting certain properties of the transaction and behavioral regularities of CEX-DEX MEV searchers, we can infer whether a given DEX trade can be part of a CEX-DEX arbitrage. Specifically, we assume that CEX-DEX arbitrage transactions are private and prioritized trades in the block executed by specialized MEV searchers, who generally do not interact with regular users.

Our methodology builds upon and extends previous approaches [34, 43, 16], refining existing heuristics and introducing additional filtering mechanisms to detect CEX-DEX arbitrage transactions from the broader set of DEX trades. We implement these detection methods using Dune Analytics [2], leveraging established datasets including [9, 35, 26, 25].

A significant advancement in our approach is the expansion beyond the constraints of previous research, which typically limited CEX-DEX trades to transactions that contained exactly one on-chain swap (or two ERC-20 token transfers) [34, 43]. Our investigation reveals

that certain CEX-DEX searchers often execute multiple swaps across different DEX pools within a single transaction to optimize on-chain liquidity utilization.[1] These transactions do not fit into the traditional heuristic of one swap and two ERC-20 token transfers. Our methodology accommodates these multi-swap transactions by aggregating sequential swaps and reconstructing the final effective trading token pair with corresponding aggregate traded amounts. Specifically, we apply the following set of formulated heuristics:

Heuristic 1. The transaction is *private*, i.e., it is not observed in any public mempool [26] prior to its on-chain inclusion.

Heuristic 2. At least one of the swaps included in the transaction is the first swap executed in the respective direction and DEX pool for the given block.

Heuristic 3. The transaction is not categorized by existing frameworks [5, 35] or our own Dune query as any type of atomic MEV activity (e.g., sandwich attacks or atomic arbitrages). The transaction does not include a liquidation event.

Heuristic 4. The transaction is not a backrun transaction associated with an Orderflow Auction (OFA) bundle.

Heuristic 5. The transaction is not submitted to any known router smart contract identified in [25], nor to any trading bot that is labeled in [24] or controls an Externally Owned Account (EOA) with an Ethereum Name Service (ENS) name attached.

Heuristic 6. The transaction contains no ERC-721 token transfer and settles as a swap between two tokens that are both listed on a major centralized exchange (e.g., Binance) after all intermediate swaps.

Heuristic 1 and 2 ensure transaction prioritization within blocks and favorable execution pricing in at least one DEX pool. Heuristics 3, 4, and 5 eliminate potential confusion with other MEV activities or retail trades executed by non-MEV bots. In particular, Heuristic 5 recognizes that bot contracts controlling EOAs with attached ENS names typically interact with regular users, suggesting they likely function as telegram bots or trading front-ends rather than specialized CEX-DEX MEV bots. For Heuristic 6, we cross-verified ERC-20 contract addresses for 287 Binance-listed tokens collected from Etherscan [3] and CoinMarketCap [1] with the Dune DEX trades dataset [9], ensuring that we filtered out meme tokens that share the same symbols with major tokens but are not listed on a CEX.

Applying the above criteria, we curate an empirical dataset comprising 8,723,233 transactions that are likely to be CEX-DEX arbitrages across multiple DEXes [9] from block 17866488 to block 21998438 spanning the period from August 8, 2023, to March 8, 2025 with their detailed information, including searcher contract address, fees, payments to the builder, trade volume, and the precise tokens and amounts traded. To the best of our knowledge, this constitutes the most comprehensive empirical dataset to date for analyzing CEX-DEX arbitrages in terms of observation duration, transaction count, and token coverage. We open source the Dune query for reproducibility [49].

3.1 Current CEX-DEX Arbitrage Landscape

Utilizing the curated dataset, we analyze temporal trends in CEX-DEX arbitrage activity over the observed period. To enhance readability and streamline the visualization, we assign intuitive labels to the 23 major searcher entities identified in the dataset with the highest

[1] Consider for example searcher contract address: `0x767c8bb1574bee5d4fe35e27e0003c89d43c5121`, with sample transaction: `0x0488cad7726ed6947be6af09d17b012fd7c986cb32a9d59cd7285b9a4e0926be`.

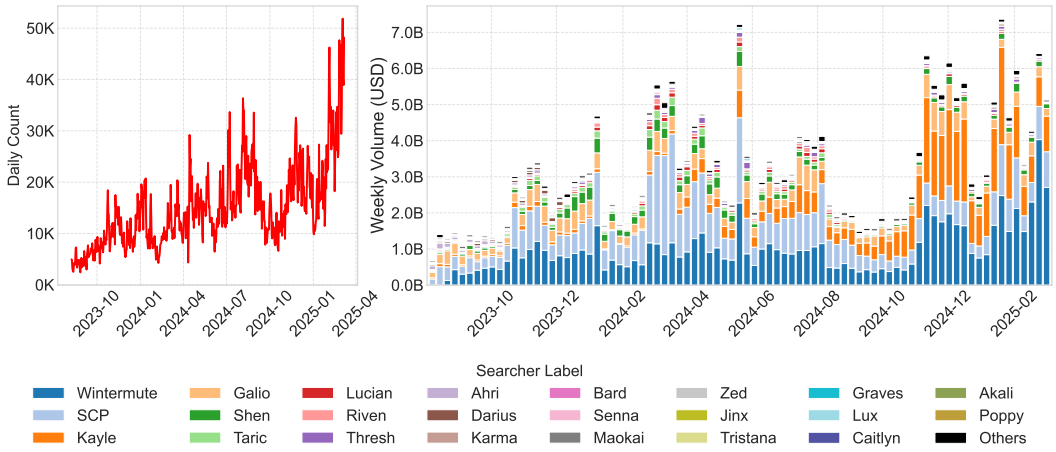


Figure 2 Daily count and weekly volume of detected CEX-DEX trades between August 2023 and March 2025.

total volume in place of their contract addresses, except for `Wintermute` and `SCP`, whose identities are retained as-is. Our subsequent analysis will focus on these 23 searcher entities, who collectively contribute to 99% of total volume.

Figure 2 presents the daily CEX-DEX transaction counts (left panel) alongside the weekly volumes of each searcher entity (right panel). Note that all trades flagged by our heuristics are shown; transactions later excluded by the stricter profitability filter (cf. Section 4.1), which roughly contribute to roughly 15% of the total count and 7% of the total volume, are still included here for landscape context. The daily count of CEX-DEX arbitrage transactions exhibits a clear upward trend over the 19-month period, despite occasional fluctuations. Notably, daily transaction counts increased by 7.2 times by Q1 2025 compared to Q3 2023.

Despite the overall increase in CEX-DEX arbitrage activity, we observe reduced diversity of successful searchers over time. Before October 2024, 23 labeled searchers collectively accounted for nearly 99% of the total CEX-DEX arbitrage volume on-chain. This landscape transformed substantially afterward, with just 14 labeled searchers remaining active and successful. The figure further reduced to 11 by Q1 2025, with three leading searchers capturing 90% of the total volume. Among all searchers, `Wintermute` and `SCP` have maintained dominant positions throughout the entire period. Interestingly, `Kayle`, initially a relatively minor searcher, began to rise significantly in influence around June 2024 and emerged as a leading searcher.

4 Estimating CEX-DEX Arbitrage Revenue and Searcher Profit

We next proceed to introduce our methodology of estimating CEX-DEX arbitrage revenue and searcher profit. We assume that searchers first establish token inventory positions through on-chain DEX trades before subsequently hedging these accumulated positions on off-chain CEX venues. This sequential execution assumption stems from the inherently higher execution uncertainty on-chain compared to centralized venues. When multiple searchers target the same arbitrage opportunity in a specific DEX pool, the probability of any searcher's on-chain transaction being included in a block is uncertain. Consequently, searchers initiate off-chain hedging operations *only after* receiving confirmation that their on-chain trades have been successfully executed, thereby avoiding potential unhedged inventory risk [17].

In addition, we assume that searchers execute the hedge for *both directions* on CEX within a short timeframe or even instantaneously after confirming successful on-chain trade. Concretely, if a searcher buys BTC and sells ETH on the DEX, they will – within a very short window – sell the newly acquired BTC and repurchase ETH on the CEX. This behavior is economically rational given the inherent volatility risk of token inventory left unhedged, particularly for low-liquidity tokens that exhibit higher price volatility [17].

Since searchers' actual hedging details on CEX are unobservable, we present a methodology to estimate their realized arbitrage revenues based on empirical observations.

4.1 Searcher-Specific Optimal Execution Horizons

A key challenge in estimating realized revenue is identifying the timing when searchers typically hedge their DEX positions off-chain. Searchers exhibit notable variations in the timing of their observations of DEX trade confirmations and subsequent CEX hedge, contrasting with standardized `slot time` reported by platforms such as Dune [2] and Etherscan [3]. These timing variations stem from several factors:

1. **Block confirmation:** Searchers observe the block and on-chain transaction executions at varying times due to differences in network connectivity and block propagation path.
2. **Information asymmetry:** Searchers possess different levels of market information (also known as *alpha* signal) affecting their execution timing decisions.
3. **Execution infrastructure:** Variations in the infrastructure of different searchers result in different processing speed. Certain integrated searchers have privileged access to the block building process and can execute their strategy closer to the actual block time.

To capture these searcher-specific timing characteristics, we analyze each searcher's historical trades by examining when their information advantage maximizes, reflected in the maximum achievable spread. To facilitate fair comparisons across different trade sizes, we define the *gross return* (GR) at time t (the *markout horizon*) as the markout revenue per unit of trade volume, i.e., the spread captured per dollar deployed in the arbitrage.

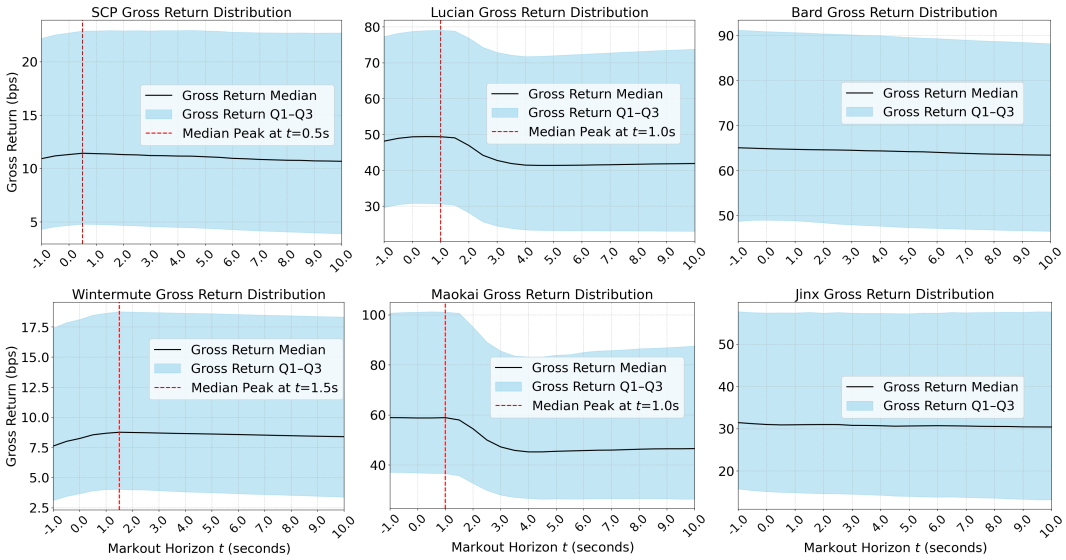
Formally, consider a CEX-DEX arbitrage transaction i with volume V_i that purchases x amount of token A and sells y amount of token B on the DEX, net of liquidity provider fees. Let $P_A(t)$ and $P_B(t)$ represent the respective CEX USD prices at markout horizon t. The markout revenue (MR) and gross return (GR) achievable by flattening the DEX-acquired inventory on CEX at time t can be estimated as:

$$\mathrm{MR}_i(t) = x \cdot P_A(t) - y \cdot P_B(t) - \text{CEX taker fees},^{2} \quad \text{and } \mathrm{GR}_i(t) = \frac{\mathrm{MR}_i(t)}{V_i}. \tag{1}$$

We value all tokens using USD to ensure consistent and accurate pricing by leveraging the superior liquidity of USDT pairs on Binance. To derive accurate USD prices for each token at specific markout horizons, we collect Binance historical quote data for USDT pairs from Tardis.dev [51] and employ the *mid-price* at each relevant markout timestamp.

▶ Remark 1. We assume 1 USDT = 1 USD, as most tokens are priced in USDT on Binance. In contrast, other USD stablecoins (e.g., USDC) may not equal 1 USD due to potential arbitrage opportunities between stablecoins.

[2] We assume CEX-DEX searchers operate at the highest user tier on Binance, thereby benefiting from the lowest applicable fee rate of 0.01725% [12].



(a) Pattern 1: `SCP` & `Wintermute`. **(b)** Pattern 2: `Lucian` & `Maokai`. **(c)** Pattern 3: `Bard` & `Jinx`.

Figure 3 Three empirically observed patterns of gross return distributions across all labeled searchers. Each *column* corresponds to one pattern that contains two representative searchers that exhibit that pattern. The black line is the median, the shaded band the interquartile range, and the dashed red line indicates the markout horizon where the median gross return peaks. $t = 0$ second denotes the `slot time`. The labels of the x-axis and y-axis are shared between plots.

We then calculate the gross return for all trades across a time window spanning from `slot time` $-$ 1 second to `slot time` $+$ 10 seconds, measured at an interval of 0.5 seconds. This interval range was chosen to represent searchers' typical submission time of DEX trades to builders, suggested by builders' bidding timing in the MEV-Boost auctions [8], and a reasonable upper bound for immediate hedging activity following DEX execution without introducing further noise caused by unrelated trades on CEX.

We exclude trades whose token price data is not recorded on Tardis on certain dates [50] and transactions whose markout revenue persistently fails to cover the base fees throughout the interval, indicative of inventory adjustments rather than genuine arbitrage opportunities.[3]

The results reveal distinct temporal patterns in gross return across different searchers. Figure 3 presents the gross return distributions for all trades of 6 representative searchers over the interval of $t \in [-1, 10]$, where $t = 0$ denotes the `slot time`, and reveals 3 different patterns that span all 23 labeled searchers. We here summarize the three patterns observed:

Pattern 1 – Rising or plateau then gentle decay. Gross return climbs (or maintains) at the peak then tapers off gradually (Figure 3a). Representative searchers: `SCP`, `Wintermute`.

Pattern 2 – Rising or plateau then abrupt drop. Gross return climbs (or maintains) at the peak, followed by a rapid decline and a flat tail (Figure 3b). Representative searchers: `Lucian`, `Maokai`.

Pattern 3 – Flat or Minimal decline. Gross return remains flat or declines slightly across the interval with no discernible peak (Figure 3c). Representative searchers: `Bard`, `Jinx`.

For simplicity, we refer to searchers whose gross return behaviors follow Pattern 1, 2, or 3 as Pattern 1, 2, or 3 searchers, respectively.

[3] We observe searchers pay much lower tips to builders for these non-arbitrage transactions. For instance, `SCP` pays on average only 0.0021 ETH for these transactions compared to 0.0081 ETH for arbitrage transactions; similarly, `Wintermute` pays 0.0011 ETH versus 0.0089 ETH for arbitrage transactions.

Pattern 1&2. 19 out of 23 labeled searchers follow the first two patterns and exhibit a characteristic *optimal execution horizon*, identified by their (median) gross return rising to or maintaining at its maximum before declining. These patterns align with theoretical expectations: before hedge execution begins, the spread typically increases as the trader's information advantage manifests in favorable price movements. Once hedging commences, gross return declines as the trade's market impact pushes prices against their position. For instance, we observe that the median gross return of `SCP`'s trades reaches its maximum 0.5 seconds after the `slot time`, while the optimal horizon for `Lucian` and `Maokai` occurs at 1.0 second, for `Wintermute` at 1.5 seconds.

We interpret the *optimal execution horizon* as the most plausible empirical proxy for when searchers initiate their CEX hedge – the instant at which their information advantage is maximized and before the profit opportunity begins to erode due to delaying execution and market impact. Accordingly, we assume each searcher initiates their hedge at their optimal execution horizon empirically identified from all their historical trades. Although a single horizon cannot capture the exact timing of every trade, it summarizes the dominant behavior across all historical trades for each searcher: the cross-trade median peaks are sharp, and the interquartile bands around them are flat, indicating the estimate is not driven by outliers.

Formally, For each searcher j we define the optimal execution horizon t_j^* as the markout horizon that maximizes the cross-trade median of gross return:

$$t_j^* = \arg\max_{t \in \mathcal{T}} \ \text{Median}_{\,i \in \mathcal{I}_j} \big\{ \text{GR}_i(t) \big\},$$

where $\mathcal{T} = \{-1.0, -0.5, 0, \ldots, 10.0\}$ seconds is set of markout horizons around the `slot time` and $\mathcal{I}_j$ denotes the full set of historical trades of searcher j. If multiple markout horizons attain the same maximum before the median gross return decreases, we select the largest t. Given t_j^*, the revenue (i.e., extractable value) and net *profit and loss* (PnL) for any trade $i \in \mathcal{I}_j$ are estimated as

$$\widehat{\text{EV}}_i = \text{MR}_i\big(t_j^*\big) - \text{base\_fees}_i, \ \text{and} \ \widehat{\text{PnL}}_i = \widehat{\text{EV}}_i - \text{builder\_tips}_i, \tag{2}$$

where builder tips include priority fees and coinbase transfer.[4] The *profit margin* is thus:

$$\text{PM}_i = \frac{\widehat{\text{PnL}}_i}{\widehat{\text{EV}}_i}, \widehat{\text{EV}}_i > 0.$$

We compute the profit margin only for trades whose estimated revenue is positive. Trades with $\widehat{\text{EV}}_i \leq 0$ are labeled as "N/A" for margin calculations and excluded from profit margin analysis, but they are retained in all other analyses related to PnL directly.[5]

Our approach to inferring CEX-DEX searchers' optimal execution horizon builds on methodologies from informed trading and market microstructure literature. Specifically, to identify this timing for each searcher, we adapt strategies from [10, 37], who empirically recover hidden execution timing of informed traders from observable post-trade price movements, by observing how CEX price movements affect the spread of their trades around the `slot time`.

[4] For calculation, we convert the price unit of base fees, priority fees, and coinbase transfer from ETH to USD using the ETH-USDT mid-price at the `slot time`.

[5] For example, consider a trade with an estimated revenue of -1 USD, for which the searcher pays 10 USD to the builder. This results in a net PnL of -11 USD. Although the trade incurs a loss, computing the profit margin would yield an illogical value of +1100%, which motivates its exclusion from margin-based analysis.

Moreover, our data-driven framework aligns with theoretical predictions by [7, 11], suggesting each trader possesses a unique, liquidity-driven optimal execution horizon characterized by a distinct "peak-then-decay" pattern in returns (cf. Section 5.2). While our method necessarily provides a simplified approximation due to unobservable hedging price-impact costs, it nonetheless yields a consistent, empirically grounded upper-bound estimate of extracted value and PnL across searchers.

Pattern 3. However, for four searchers – `Bard`, `Jinx`, `Tristana` and `Lux` – their gross return curve is essentially flat (or drifts only marginally downward) across the $[-1, 10]$ seconds window, so no clear "peak-then-decay" point emerges. The most plausible interpretation is that these searchers hold their inventory and do not hedge their positions on CEX within this window: if they were unwinding inventory promptly, the resulting price impact on the CEX leg should depress the gross return. More explanations can be found in Section 5.2.

While extending the window further than +10 seconds might eventually reveal such a decline, that comes at the cost of substantial noise – price moves unrelated to the original trade. As our revenue and PnL estimation relies on identifying an optimal execution horizon, it cannot be applied reliably to these searchers. We therefore exclude these four searchers from subsequent revenue and PnL calculations, whose trades together account only for 2.7% of total transactions detected; they remain in the descriptive landscape figures but are omitted from revenue and PnL analyses.

5 Token Liquidity as the Driver

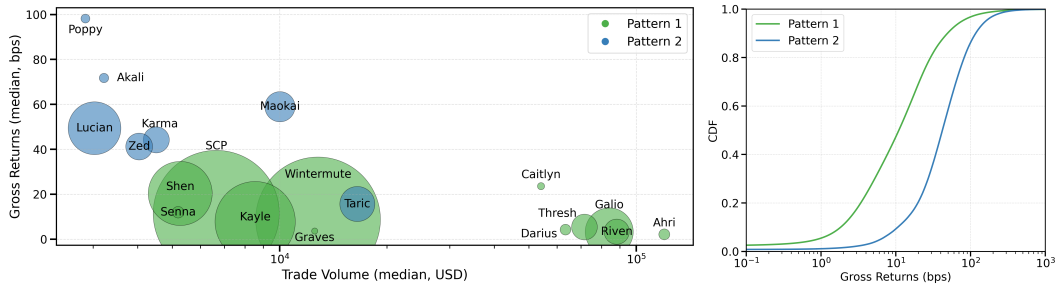
In this section, we demonstrate that the distinct patterns of the gross return curve observed across searchers are associated with the difference in the token liquidity they trade, which shapes searchers' value extraction regimes and subsequent hedge execution.

5.1 Liquidity and Value Extraction Regimes

As shown in Figures 3, Pattern 2 searchers earn markedly higher gross returns compared to Pattern 1 searchers. To highlight the role of token liquidity, we group tokens into "Major" – large-cap assets and commonly traded stablecoins (WETH, WBTC, USDT, USDC, TUSD, FDUSD, BUSD, DAI) – and "ALT", the remaining long-tail assets, and compare the trading pairs between these searchers.

We find that Pattern 1 searchers such as `SCP` and `Wintermute` predominantly engage with highly liquid major tokens characterized by deep order books. In contrast, Pattern 2 searchers – typically smaller and less competitive – tend to focus more on lower-liquidity ALT tokens. Prices of these assets show higher volatility and suffer significant slippage even with modest trade volumes. This suggests that differences in token liquidity may influence both the achievable spread and the revenue strategies adopted by different searchers. We illustrate the empirical relationship between trade size and gross return in Figure 4.

From Figure 4a, we observe that Pattern 1 searchers frequently execute trades exceeding 10,000 USD, leveraging the liquidity depth of major tokens to deploy substantial capital without significant price impact. However, they only earn single-digit basis-point gross returns. Figure 4b confirms that nearly 90% of their trades earn less than 20 bps gross return. Despite lower spreads per trade, these searchers accumulate substantial total revenue by leveraging large-scale execution. This pattern is particularly evident for leading searchers.



(a) (b)

Figure 4 a) Median trade volume and gross returns for each searcher, with bubble size indicating searchers' total estimated revenue and color encoding their execution pattern. **b)** Cumulative Distribution Function (CDF) of gross returns across all trades between Pattern 1 and 2 searchers.

Conversely, Pattern 2 searchers face constrained liquidity conditions, typically restricting median trade sizes below 5,000 USD to avoid significant self-induced slippage. Nonetheless, limited liquidity frequently creates greater mispricing opportunities for long-tail assets, resulting in markedly higher gross returns ranging from 20 to 120 bps per trade. Figure 4b confirms that 50% of these searchers' trades exceed 40 bps in gross return.

This divergence reveals two distinct strategies to extract value: large-volume trade with narrow-spread targeting major tokens (Pattern 1) versus small-volume trade with wider-spread targeting long-tail tokens (Pattern 2). Interestingly, certain searchers exhibit hybrid strategies: `Maokai` and `Taric`, while categorized as Pattern 2, sustain relatively large trading volumes by maintaining significant activity in Major-Major pairs. `Shen` and `Senna`, categorized as Pattern 1, engage in notable ALT token trading, resulting in comparatively smaller volumes.

5.2 Liquidity and Hedge Execution

Variations in the speed and magnitude of spread closure after peaking between Pattern 1 and 2 observed in Figure 3 can now be better understood as the value extraction regimes above manifest mechanically in how quickly spreads collapse once the hedge is executed.

As described above, Pattern 1 searchers focus primarily on deeply liquid major tokens. Consequently, after the initial peak, their gross returns decline smoothly and gradually, indicating that their hedge executions generate limited market impact. Conversely, Pattern 2 searchers, trading primarily lower-liquidity ALT tokens, experience rapid, pronounced closures of spreads immediately after peak returns. The thin liquidity conditions mean their hedge executions will quickly push prices against their positions, collapsing the spread.

To quantitatively validate this link, we examine the correlation between token liquidity and how rapidly and significantly gross return declines. As is illustrated in Figure 5, we find a strong negative correlation between the proportion of major token trade and the percentage decline in median gross returns within 3 seconds after peaking for each searcher. This confirms that searchers predominantly trading major tokens experience slower and smaller gross return declines, reflective of reduced price impact. These findings align with the theoretical predictions in [7, 11].

This rapid and significant spread closing after the peak for Pattern 2 searchers strongly indicates that they prioritize immediate position neutralization despite significant price impact, as they face a trade-off in less liquid markets: extending hedge execution may

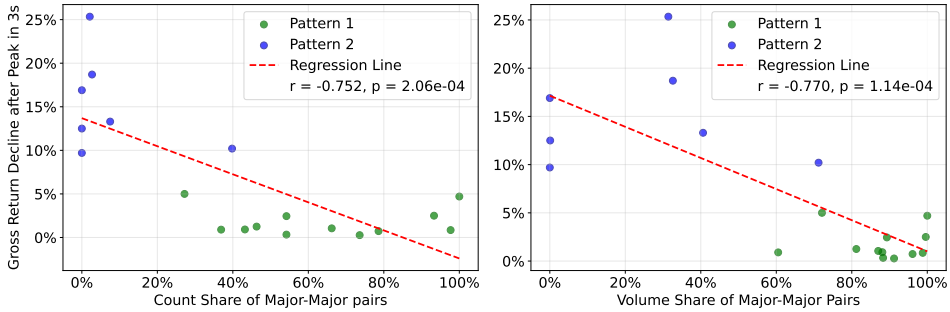


Figure 5 Correlation between the decline in median gross returns within 3 seconds after peak and the proportion of trade count (left panel) and trade volume (right panel) involving Major–Major token pairs for each searcher. Each scatter represents one searcher, with color indicating their pattern. The y-axis label is shared between the two subplots.

increase inventory risk compared to potential hedge cost savings. The observed flat tail of their gross return after the drop further confirms their emphasis on execution speed and inventory risk reduction over slippage minimization. In contrast, the muted, gradual spread compression for Pattern 1 searchers is consistent with deeper liquidity markets: even if hedging immediately with substantial volume, high liquid markets typically absorb size with only limited price impact. These searchers are further enabled to fragment large-size hedge trade into smaller pieces to minimize price impact at the cost of holding the inventory longer.

Finally, we note that Pattern 3 searchers exclusively trade ALT tokens, yet exhibit minimal to no spread closure within the observed interval. This reinforces our earlier interpretation that these searchers likely delay hedging or choose not to hedge within our observation window, as immediate hedging would otherwise generate pronounced declines in their gross returns similar to those of Pattern 2.

Overall, our analysis empirically highlights how token liquidity profiles shape distinct value extraction and hedging regimes for CEX-DEX searchers. While token liquidity appears central to explaining observed patterns, we acknowledge other factors not presented here – such as individual searcher strategies or risk preferences – may also influence these outcomes.

6 Extracted Value from CEX-DEX Arbitrages

Using the methodology outlined in Section 4, we analyze a total of 7,203,560 CEX-DEX arbitrages with a total volume of 241.7B USD executed by 19 labeled searchers between August 8, 2023 and March 8, 2025. These figures significantly surpass atomic MEV activities during the same timeframe. Figure 6 illustrates the cumulative CEX-DEX arbitrage revenue (i.e., extracted value) and weekly distribution of revenue among searchers over these 19 months. According to our estimation, a total of 233.8M USD is extracted by these 19 searchers through CEX-DEX arbitrages. At the time of writing, this figure is comparable to the total value extracted from atomic arbitrages *since The Merge* [4].

Beyond the revenue figures, our result highlights a structural shift towards market centralization. Both the volume of CEX-DEX arbitrages and the associated extracted value have become increasingly concentrated among a smaller subset of leading searchers. To quantify this trend, we compute the Herfindahl-Hirschman Index (HHI) for both trade volume and extracted value, revealing consistently rising concentration and a highly centralized market by the end of the observed period.

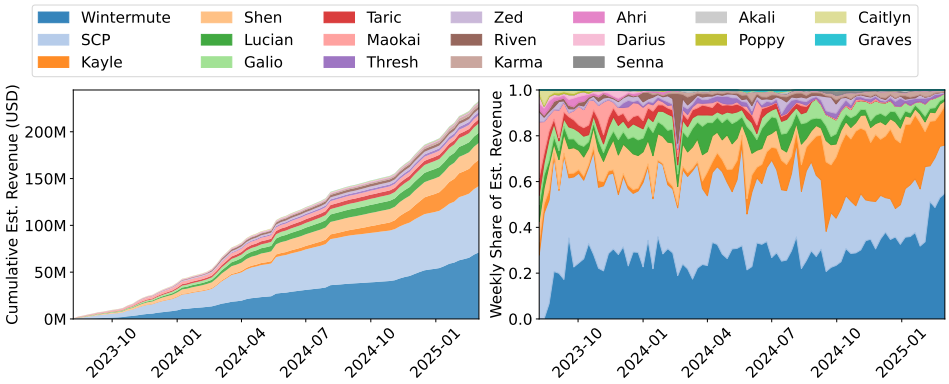


Figure 6 Cumulative CEX-DEX arbitrage revenue (left panel) and daily share of revenue (right panel) by 19 labeled searchers. A total of 233.8M USD is extracted from 7,203,560 CEX-DEX arbitrages by these 19 searchers between August 8, 2023 and March 8, 2025.

The two largest searchers, `Wintermute` and `SCP` consistently extracted over 50% of total revenue throughout the entire observation period. A notable turning point occurred around June 2024, as the previously smaller searcher `Kayle` rapidly expanded its market presence and altered the competitive landscape. By October 2024, we observe a sharp reduction in searcher participation, with only 12 out of 19 labeled searchers remaining active. The market thus resembles a clear trident, with three searchers – `Wintermute`, `SCP`, and `Kayle` – collectively capturing 170.8M USD, representing approximately 73% of the total cumulative extracted value. Furthermore, these top three players dominate arbitrage opportunities even more decisively, jointly accounting for around 90% of the extracted value by Q1 2025. In contrast, smaller searchers' volume and revenue share have been compressed over time, and some searchers eventually stopped their operations and retreated from the market.

7 Searcher-Builder Integration and Profitability

Searchers' net profitability ultimately depends on the fraction of revenue shared with the block builder. In this section, we examine searchers' integration level with block builders, and analyze its impact on searchers' PnL and profit margin.

7.1 Searcher-Builder Integration

We start by assessing the level of integration between each of the 19 labeled searchers and major block builders by measuring the proportion of CEX-DEX arbitrage volume that each searcher directs to specific builders. Figure 7 visualizes these relationships through color intensity – darker cells indicate higher volume proportions and thus stronger integration between two entities. It is well-established that searcher `SCP` and builder `beaverbuild` are vertically integrated, as do `Wintermute` and `rsync` [34, 43, 58].

Our analysis further reveals exclusivity patterns between specific searcher-builder pairs:

- `Thresh` and `Senna` direct 60% and 89% of their volume to `beaverbuild`, respectively.
- `Kayle` and `Graves` send 52% and 100% of their volume to `Titan`, respectively.

We refer to these searchers, who send more than 50% of their volume to one builder, as the *exclusive searcher* (including *integrated searchers* `SCP` and `Wintermute`) for that builder, and correspondingly label their primary recipient builder as their *exclusive builder*. All other searchers, who spread their trades evenly among multiple builders, we classify as

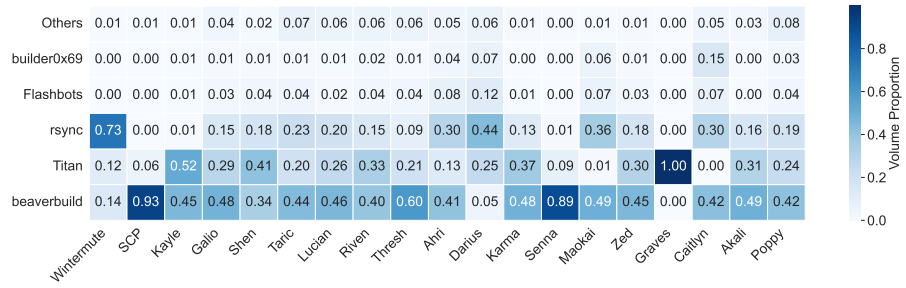
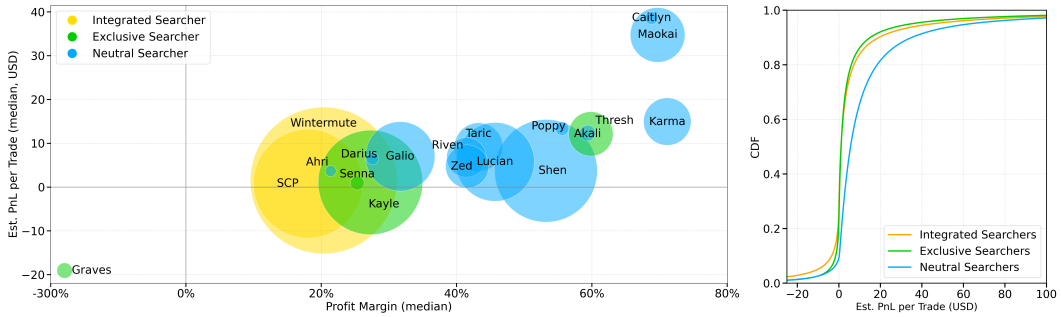


Figure 7 The proportion of CEX-DEX arbitrage volume from each labeled searcher (x-axis) in the blocks of top 5 builders (y-axis).



(a) **(b)**

Figure 8 a) Bubble plot illustrating searchers' median profit margin (x-axis) and median estimated PnL per trade (y-axis). The color shows the searcher's integration level. The bubble size represents the searcher's total PnL. Due to `Graves` having a negative total PnL, we show its absolute value. The negative tick of the x-axis (-300%) is not proportionally spaced. **b)** Cumulative Distribution Function (CDF) of estimated PnL across all trades between the three integration levels.

neutral searchers. Notably, although `Kayle` primarily splits trades between `Titan` (52%) and `beaverbuild` (45%), subsequent analysis will demonstrate a higher integration level with `Titan` than with `beaverbuild`.

7.2 Searcher Profitability Analysis

We then proceed to analyze the profitability of the searchers. In Figure 8a, we find that neutral searchers generally have much higher profitability than most exclusive searchers: their median profit margins cluster around 30% - 70%, whereas those of most exclusive searchers are within 20% - 40%. Neutral searchers' median PnL per trade often exceeds 10 USD, while those of most exclusive searchers rarely exceeds 5 USD. Figure 8b confirms the edge: 80% of exclusive and integrated searchers' trades net fewer than 10 USD, whereas nearly half of neutral searchers' trades exceed this threshold.

Table 1 summarizes the profitability status of the searchers. Despite being market leaders, `Wintermute`, `SCP`, and `Kayle` clear the largest notional volume and earn the highest total revenues, yet their profit margin is among the lowest and their median PnL per trade mostly hovers only a few dollars above zero. `SCP`'s total PnL is barely larger than that of neutral searcher `Shen`, whose trade volume is much lower. Moreover, `Graves`, who sends almost all their trades to `Titan`, exhibits negative profitability according to our estimation, indicating payments to the builder often exceed their arbitrage revenue. We report the medians rather

Table 1 Searcher profitability summary sorted by total volume.

Searcher	Total Volume [USD]	Total Est. Rev. [USD]	Total Builder Tips [USD]	Total Est. PnL [USD]	Median Trade Rev. [USD]	Median Trade PnL [USD]	Median Profit Margin
Wintermute	74.8B	71.4M	47.1M	24.3M	11.36	1.65	20.9%
SCP	63.5B	71.1M	57.4M	13.7M	7.34	0.86	18.5%
Kayle	41.0B	28.3M	16.0M	12.3M	6.05	1.15	27.9%
Galio	25.9B	9.9M	4.5M	5.4M	24.93	7.10	31.7%
Shen	12.3B	18.1M	6.3M	11.8M	8.42	3.77	53.2%
Taric	5.0B	4.4M	1.8M	2.6M	22.92	9.28	43.3%
Lucian	3.9B	10.9M	3.9M	7.0M	14.74	5.83	45.8%
Riven	3.1B	3.1M	1.3M	1.9M	31.42	8.16	42.5%
Thresh	2.7B	3.2M	1.0M	2.2M	34.12	13.65	59.7%
Ahri	1.7B	541.4K	399.0K	162.4K	24.08	4.30	22.8%
Darius	1.1B	497.0K	331.3K	165.6K	29.62	6.62	27.9%
Karma	674.8M	3.0M	517.5K	2.5M	22.82	15.02	71.2%
Senna	596.3M	503.9K	268.1K	235.8K	5.22	1.06	27.2%
Maokai	588.4M	4.4M	1.1M	3.3M	52.40	35.13	69.7%
Zed	580.2M	3.1M	1.1M	2.1M	13.99	4.83	41.8%
Graves	229.3M	173.4K	353.1K	-179.7K	8.77	-19.34	-272.1%
Caitlyn	96.1M	253.7K	67.6K	186.1K	69.02	39.62	68.8%
Akali	49.4M	378.6K	100.5K	278.1K	22.20	12.37	59.4%
Poppy	35.5M	371.7K	171.0K	200.7K	25.00	13.43	55.6%

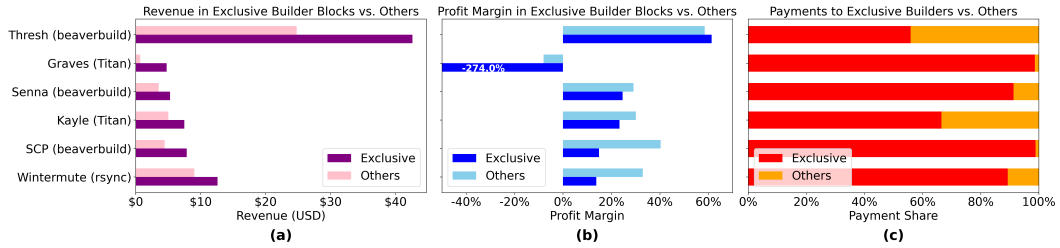


Figure 9 a) Revenue (median) of exclusive searchers' trades in their respective exclusive builders' blocks vs. in other blocks. **b)** Profit margin (median) of exclusive searchers' trades in their respective exclusive builders' blocks vs. in other blocks. **c)** Percentage of total payments from exclusive searchers to their exclusive builder vs. to other builders. The y-axis is shared between three plots and represents the exclusive pair of searcher and builder.

than the cumulative means for metrics of revenue per trade, PnL per trade, and profit margin to avoid the disproportionate influence of outlier trades that skew the means; the medians therefore represent the typical trade more faithfully.

Figure 9 offers deeper insights into these dynamics between exclusive searchers and their corresponding exclusive builders. In Figure 9a, we observe that exclusive searchers consistently send higher-revenue trades to their exclusive builder, compared to other builders. In Figure 9b, for most exclusive searchers, their trades in their respective exclusive builder's block consistently have lower median profit margins, indicating a higher proportion of revenue shared with the exclusive builder than other builders. For instance, both `Wintermute` and `SCP` transfer nearly 90% of their arbitrage revenue to their integrated builder `rsync` and `beaverbuild`, retaining lower profit margins slightly above 10%. While most `Grave`'s trades are not profitable, their trades in their exclusive builder `Titan`'s blocks have a median profit margin of -274.0%. In Figure 9c, as a result of submitting higher revenue trades and sharing greater revenue proportion, exclusive searchers favor their exclusive builder with substantially higher total payments than the amount paid to other builders.

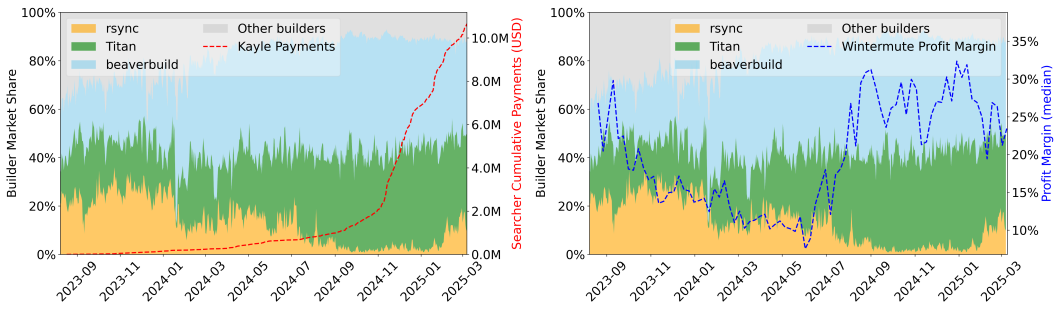


Figure 10 **a)** `Titan` market share and `Kayle`'s cumulative payments to `Titan`. The left y-axis shows the builder market share, and the right y-axis shows `Kayle`' cumulative payments to `Titan`. **b)** `rsync` market share and `Wintermute` profit margin. The left y-axis shows the builder market share, and the right y-axis shows `Wintermute`'s median profit margin.

Both `Wintermute` and `SCP` transfer nearly 90% of their arbitrage revenue directly to their integrated builder to gain an edge in the block building auction and seek a prioritized execution for their trades. `Graves` sends their trade exclusively to `Titan` with payments much higher than their revenue. While `Kayle` splits its trade volume almost evenly between `beaverbuild` and `Titan`, the above economics point to a much higher integration with `Titan`: trades landed in `Titan` blocks earn noticeably higher revenue but carry a markedly lower profit margin, and `Kayle`'s total payments to Titan are roughly double those paid to the rest builders combined. Similar patterns are also evident for `Senna` with `beaverbuild`.

`Thresh`, by contrast, appears only loosely affiliated with `beaverbuild` despite sending 60% of its volume there. The median margin `Thresh` keeps in `beaverbuild` blocks still exceeds 50% – in fact slightly higher than in other blocks – and its total payments to `beaverbuild` are only modestly larger than to builders, which is reasonable considering the success of `beaverbuild` as a builder [55]. This indicates that `Thresh` simply routes most of their trades to `beaverbuild` but is less likely to engage in an exclusive arrangement.

In summary, neutral searchers distribute trades across several builders and pay only the lowest necessary tips for block inclusion, thereby retaining the largest possible share of revenue. Exclusive searchers – likely bound by an exclusive orderflow deal with the builder – often direct their profitable trades to a single builder, willingly sharing a substantially larger fraction of their revenue and accepting slim or even negative PnL at the trade level.

7.3 Impact of Exclusive CEX-DEX Flow on Market Structures

Prior research has highlighted the strategic advantage builders gain by accessing high-value exclusive CEX-DEX flow [32, 34, 43, 57, 58]. Here, we extend this discussion by closely examining interactions between exclusive searcher-builder pairs and their impact on both the searcher and builder markets.

As previously described, `Kayle`, an exclusive searcher affiliated with builder `Titan`, quickly ascended to top-tier searcher status with the third-largest volume since June 2024. Simultaneously, as illustrated in Figure 10a, `Titan`'s market share grew by 15% till September 2024 with a substantial surge in tips from `Kayle`. To quantitatively assess this relationship, we conduct a correlation analysis between `Kayle`'s daily CEX-DEX volume share and `Titan`'s daily market share between June 1 to September 30, 2024.

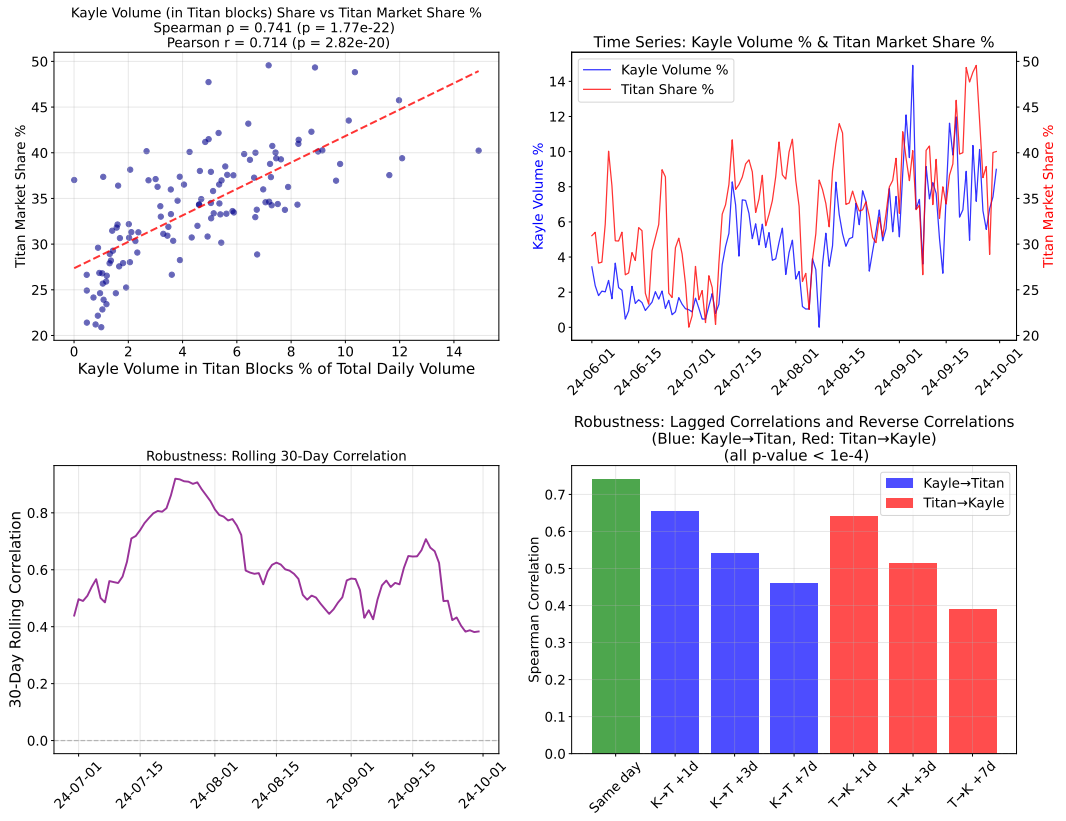


Figure 11 Correlation analysis of `Kayle`'s daily volume share in `Titan`'s blocks (as a percentage of total daily CEX-DEX volume) and `Titan`'s daily market share. Using percentage metrics effectively controls for overall fluctuations in daily total CEX-DEX volume.

We find a strong and robust mutual correlation between `Kayle`'s daily volume share in `Titan's` blocks and `Titan`'s daily market share during this period. Specifically, Figure 11 (top-left panel) shows a significant contemporaneous correlation (Spearman's $\rho = 0.74$, $p < 0.0001$), indicating that higher proportion of `Kayle`'s trade volume included in `Titan`'s blocks strongly coincide with higher share for `Titan` on the same day. The aligned trajectories shown in the time series comparison (top-right panel) further confirm their synchronized behavior. Robustness checks using rolling 30-day correlations confirm the consistency and significance of this relationship across the period analyzed (bottom-left panel).

We further investigate the lagged and reverse correlations at intervals of 1, 3, and 7 days (bottom-right panel). The results indicate a notable predictive component: higher daily volume share from `Kayle` to `Titan` positively correlates with increased market share for `Titan` in subsequent days. This effect is most pronounced at a 1-day lag (Spearman's $\rho = 0.655, p < 0.0001$), gradually decreasing over one week. One plausible explanation is that larger `Kayle` flow in one day brings larger tip revenue for Titan and enhances `Titan`'s bidding power in subsequent days. Moreover, we identify a reverse effect of similar magnitude, suggesting a strong feedback loop: higher market share for `Titan` subsequently increases `Kayle`'s market capture and flow directed back to `Titan`. Indeed, we observe a substantial surge in CEX-DEX volume share included in `Titan`'s blocks since June 2024.

This dynamic played a major role in accelerating centralization within the searcher market during this period and beyond. As shown in Figure 12, CEX-DEX volume became increasingly concentrated. This shift also had significant repercussions for competing builders

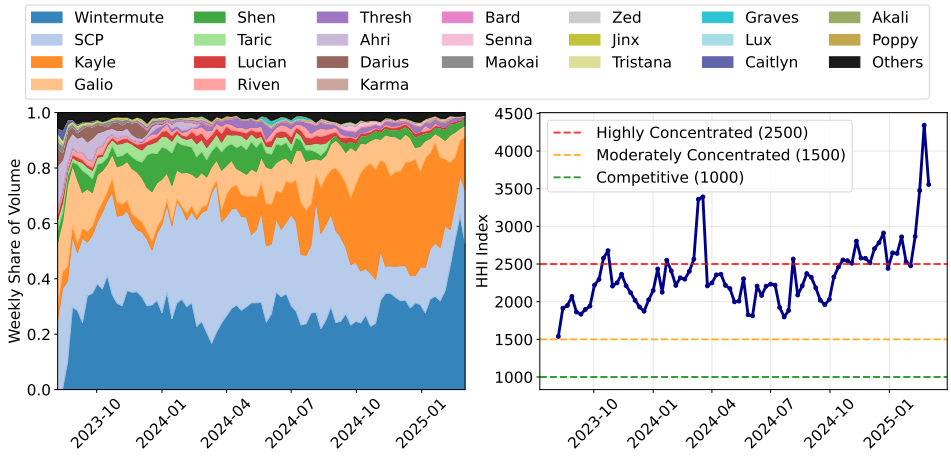


Figure 12 Weekly share of CEX-DEX volume by all the searchers, including Pattern 3 and other searchers, and HHI index of weekly CEX-DEX volume.

– particularly `rsync`. As a result of `Titan`'s expansion, in Figure 10b, we observe `rsync`'s market share fell below 5% by September 2024, effectively retreating from competition in the builder market. At the same time, `Wintermute`'s median profit margin quickly rebounded, indicating a strategic shift away from aggressively supporting `rsync` towards maximizing searcher profitability.

Interpretation and Implications. Combined with the analysis in Section 7.2, these findings provide strong quantitative evidence supporting a mutually reinforcing exclusivity partnership between `Kayle` and `Titan`. Indeed, [15] analyzed both winning and non-winning block contents submitted to the Agnostic Relay and confirmed that `Kayle`'s trades were exclusively seen in `Titan`'s blocks.

Searchers with exclusive builder partnerships, despite paying higher per-trade costs and retaining lower margins, likely gain prioritized inclusion, potentially capturing a greater share of arbitrage opportunities. Builders, in turn, leverage exclusive high-value flow to strengthen their competitiveness in block auctions, further empowering their searcher partners. While explicit off-chain rebates and agreements remain opaque, the observed patterns – higher payments, lower searcher margins, and increased market share – strongly indicate such arrangements. Collectively, these results underscore the significant centralizing pressure exerted by exclusive partnerships and vertical integration within the Ethereum searcher and builder markets.

8 Integrated Builder Profits Correction

Section 7.2 shows `SCP` and `Wintermute` only retain 10-15% of their arbitrage revenue and transfer nearly 90% to their integrated builders. However, `beaverbuild` and `rsync` are unlikely to pass this entire amount as their bid to the proposer; part of the searcher payments may be pocketed as builder surplus upon winning the block. In this section, we combine builders' on-chain profits with searcher-level PnL estimates to correct the oversight by prior research, where only builders' on-chain profit was mesaured [43, 58]. This reconciliation provides deeper insight into the profitability of integrated searcher-builder entities.

Table 2 Builder profitability summary.

Builder-Searcher	Total Blocks [#]	Total Bid Value [#]	Total Builder Profit [USD]	Builder Margin	Total Searcher PnL [USD]	Aggregated Profit [USD]	Agg. Profit Margin
beaverbuild - SCP	1,603,780	402.87M	66.68M	6.69%	11.05M	77.73M	7.92%
rsync - Wintermute	543,564	156.78M	9.65M	-2.24%	14.77M	24.42M	27.06%
Titan	1,206,806	266.80M	92.44M	5.85%	N/A	92.44M	5.85%

We collect the Ultra Sound bid adjustments data [46] and MEV-Boost payloads data [28], and analyze builder profitability from block 17866488 to block 21998438 spanning the period from August 8, 2023, to March 8, 2025. Table 2 summarizes the profitability of `beaverbuild`, `rsync`, and `Titan`, considering integrated searcher profits (if any). "Builder Margin" presents the cumulative mean of profit margin if only builder profit for each block is considered, while "Aggregated Profit Margin" presents the cumulative mean of profit margin if considering both builder profit and integrated searcher PnL for each block. Recall, given that the searcher PnL is an upper-bound estimate, the aggregated profit and margin are also upper-bound estimates.

Results show that `beaverbuild`'s builder profit is roughly 6 times their searcher profit, and the profit margin only increases by about 1% by incorporating their searcher profit. This confirms that while `SCP` transfers nearly 90% of their revenue to their builder, a sizable share never reaches the proposer and is instead retained by their builder.

Since `rsync` is less competitive as a builder, their builder profit is lower even with a slightly negative builder margin. However, if incorporating their searcher profit, we find higher profit retained at their searcher level than `SCP`, and the blocks they win are exceptionally lucrative with a 27.06% profit margin. This indicates that they compete selectively for the most profitable blocks and mainly profit as a searcher instead of as a builder.

Finally, `Titan`, despite not being vertically integrated with any searcher, earns the highest total profit. This is strong evidence that they have exclusive deals with high-value orderflow providers (e.g., searcher `Kayle`), and rebate part of their builder profit back to them.

We further refine builder subsidization measure in MEV-Boost auctions by including integrated searcher profits for `beaverbuild` and `rsync`. Prior studies measured subsidies with builders' on-chain profits only, suggesting builders subsidize blocks to maintain market presence, potentially offset by exclusive orderflow provider profits [43, 58, 52]. We define a block as "subsidized" only if both the builder's on-chain profit and aggregated profit are negative. If the builder profit is positive, we do not classify the block as subsidized, even if the aggregated profit is negative. As is shown in Table 3, after accounting for integrated searcher profits, the apparent subsidy burden for both builders shrinks by roughly 15% fewer blocks and slightly lower outflows. This suggests that the majority of block subsidizing behavior by `beaverbuild` and `rsync` is not directly driven by immediate coverage from their searcher profits in that block.

9 Discussion

Limitations of Our Approach. Although our heuristics expand the coverage of identified CEX-DEX arbitrage trades compared to prior methodologies [43, 34], certain trades may remain undetected, resulting in potentially underestimating searcher revenues and PnL. Given the resources available, our methodology represents the best feasible solution at present.

Table 3 Builder subsidy summary.

Builder	Subsidized Blocks Before Correction [#]	Subsidized Blocks After Correction [#]	Builder Subsidy Before Correction [USD]	Builder Subsidy After Correction [USD]
beaverbuild	498,388	426,335	-2.31M	-2.13M
rsync	179,925	153,713	-3.89M	-3.77M

Our revenue estimation method – based on empirically derived optimal execution horizons across historical trades for each searcher – introduces approximation errors by assuming uniform hedge execution timing across all trades for each searcher and neglecting the price impact of hedging, particularly relevant for low-liquidity tokens. Despite these limitations, our approach is grounded in microstructure literature and captures dominant execution patterns for each searcher, offering consistent upper-bound estimates that allow meaningful comparisons among searchers.

Finally, we do not account for potential off-chain rebates or Orderflow Auction refunds, which might underestimate actual searcher PnL and overstate builder profits, especially for exclusive searcher-builder partnerships.

Centralization in the searcher market. Previous studies have discussed the impact of vertical integration, exclusive orderflow, and builder market centralization, highlighting concerns such as censorship and proposer loss [43, 58, 34]. Our analysis pushes the centralization narrative one step upstream. Critically, we are the first to provide robust quantitative estimates of the value extracted through CEX-DEX arbitrages and shed empirical light on profit-sharing dynamics previously understood only qualitatively.

It appears that, just as arbitrages in traditional finance have been concentrated among a few specialist firms [48] – often because they purchase retail order flow (i.e., payment for order flow) or control clearing licenses [22, 36] – arbitrages at the TradFi and DeFi interface are similarly centralized by searcher-builder orderflow arrangements and vertical integration.

The two largest CEX-DEX searchers by volume and revenue – `SCP` and `Wintermute` – are each vertically integrated with leading block builders, `beaverbuild` and `rsync`, respectively. This vertical integration grants them a decisive strategic advantage in securing block inclusion for their arbitrage transactions. Further, we identify that the third-best searcher `Kayle` emerged as a significant market participant beginning around June 2024, likely following an exclusive arrangement with the second largest builder `Titan`. Collectively, these three searchers dominate approximately three-quarters of the total arbitrage volume and value extracted, reinforcing further centralization in both the builder and searcher markets.

In contrast, smaller searchers, despite maintaining higher margins per trade, have seen their overall volume and revenue share diminish over time. Given the already high entry barriers to CEX-DEX arbitrage, this further centralization significantly elevates barriers to entry, discouraging market participation by smaller entities and fostering economies of scale.

Recent research also highlights centralization concerns in DEX solver markets, where notably, `Wintermute` and `SCP` similarly dominate the solver markets on prominent platforms like CoWSwap, Uniswap X, and 1inchFusion [59]. Taken together, these observations suggest that the entire MEV supply chain – from solver to searcher to builder – is becoming increasingly consolidated around a few dominant market participants.

Mitigations and Challenges. One proposed mitigation discussed in prior works [39, 34, 31] involves shorter block times (e.g., EIP-7782 [6]) to reduce price movements between blocks and limit overall arbitrage opportunities. While this lowers the total value extractable from CEX-DEX arbitrages, it is less clear that shorter block times would meaningfully shift the market structure, given the sustained advantage held by dominant, integrated players.

Mechanisms designed to promote decentralization in the downstream builder market, such as Orderflow Auctions [13, 29] and BuilderNet [23], could enhance decentralization at the builder level. However, their impact on the upstream searcher market remains uncertain and warrants further study.

Lastly, mechanisms aimed at capturing, redistributing, or burning MEV at the application [60, 38] or protocol level [20, 21, 14] could reduce centralization pressures. Nevertheless, these approaches rely on accurate MEV oracles [58], which remain an open technical challenge.

10 Conclusion

In this work, we shed light on value extraction, profitability, and market dynamics associated with CEX-DEX arbitrage on Ethereum. Our findings highlight increasing centralization as three major searchers affiliated with top builders dominate CEX-DEX arbitrage opportunities, compressing smaller participants despite their higher per-trade margins. Exclusive searcher-builder arrangements further amplify these centralization pressures both downstream and upstream of the MEV supply chain, underscoring critical economic and strategic considerations for Ethereum's decentralization guarantees.

References

1 Coinmarketcap. URL: `https://coinmarketcap.com/`.
2 Dune analytics. URL: `https://dune.com/home`.
3 Etherscan.io. URL: `https://etherscan.io/`.
4 libmev. URL: `https://libmev.com/`.
5 Zeromev. URL: `https://zeromev.org/`.
6 Ben Adams and Dankrad Feist. Eip-7782: Reduce block latency. URL: `https://eips.ethereum.org/EIPS/eip-7782`.
7 Robert Almgren and Neil Chriss. Optimal execution of portfolio transactions. *Journal of Risk*, 3(2):5–39, 2000.
8 Data Always. Mev bid data explorer. URL: `https://dune.com/data_always/mev-bid-data-explorer?block_number_t72e94=22196045`.
9 Dune Analytics. dex.trades table on ethereum. URL: `https://docs.dune.com/data-catalog/evm/ethereum/curated-data/dex/dex-trades`.
10 Michael J. Barclay and Jerold B. Warner. Stealth trading and volatility: Which trades move prices? *Journal of Financial Economics*, 34(3):281–305, 1993.
11 Dimitris Bertsimas and Andrew W Lo. Optimal control of execution costs. *Journal of financial markets*, 1(1):1–50, 1998.
12 Binance. Spot trading fee rate. URL: `https://www.binance.com/fee/trading`.
13 MEV Blocker. Mev blocker. `https://mevblocker.io/`. accessed: 2025-05-14.
14 Jonah Burian, Davide Crapis, and Fahad Saleh. Mev capture and decentralization in execution tickets, 2024. `doi:10.48550/arXiv.2408.11255`.
15 Andrea Canidio and Vabuk Pahari. Becoming immutable: How ethereum is made, 2025. `doi:10.48550/arXiv.2506.04940`.
16 Colin Chan, Thomas Thiery, Julian Ma, and Barnabé Monnot. Empirical analysis of cross domain cex-dex arbitrage on ethereum. URL: `https://ethresear.ch/t/empirical-analysis-of-cross-domain-cex-dex-arbitrage-on-ethereum/17620`.

17 Eugene Chen, Alex Toberoff, Suraj Srinivasan, and Ankit Chiplunkar. A tale of two arbitrages. URL: https://frontier.tech/a-tale-of-two-arbitrages.

18 Ankit Chiplunkar and Stephane Gosselin. A new game in town. URL: http://frontier.tech/a-new-game-in-town.

19 Philip Daian, Steven Goldfeder, Tyler Kell, Yunqi Li, Xueyuan Zhao, Iddo Bentov, Lorenz Breidenbach, and Ari Juels. Flash boys 2.0: Frontrunning, transaction reordering, and consensus instability in decentralized exchanges. *arXiv preprint arXiv:1904.05234*, 2019. arXiv:1904.05234.

20 Francesco D'Amato. Committee-driven mev smoothing. URL: https://ethresear.ch/t/committee-driven-mev-smoothing/10408.

21 Justin Drake. Mev burn — a simple design. URL: https://ethresear.ch/t/mev-burn-a-simple-design/15590.

22 Darrell Duffie and Haoxiang Zhu. Does a central clearing counterparty reduce counterparty risk? *The Review of Asset Pricing Studies*, 1(1):74–95, 2011.

23 Flashbots. Buildernet. URL: http://buildernet.org/docs.

24 Flashbots. Dex trading bot table. URL: https://dune.com/queries/3375587.

25 Flashbots. Illuminating the order flow landscape. https://writings.flashbots.net/illuminate-the-order-flow. accessed: 2025-05-23.

26 Flashbots. Mempool dumpster. accessed: 2025-05-08. URL: https://github.com/flashbots/mempool-dumpster.

27 Flashbots. Mev-boost. https://docs.flashbots.net/flashbots-mev-boost/. accessed: 2025-05-20.

28 Flashbots. Mev-boost analytics - relay scan. https://www.relayscan.io/builder-profit?t=7d. accessed: 2025-05-14.

29 Flashbots. Mev-share. https://docs.flashbots.net/flashbots-protect/mev-share. accessed: 2025-05-14.

30 Elijah Fox, Mallesh Pai, and Max Resnick. Censorship resistance in on-chain auctions, 2023. doi:10.48550/arXiv.2301.13321.

31 Robin Fritsch and Andrea Canidio. Measuring arbitrage losses and profitability of amm liquidity, 2024. doi:10.48550/arXiv.2404.05803.

32 Tivas Gupta, Mallesh M Pai, and Max Resnick. The centralizing effects of private order flow on proposer-builder separation. In *5th Conference on Advances in Financial Technologies*, 2023.

33 Lioba Heimbach, Lucianna Kiffer, Christof Ferreira Torres, and Roger Wattenhofer. Ethereum's proposer-builder separation: Promises and realities. In *2023 ACM Internet Measurement Conference (IMC), Montreal, QC, Canada*, October 2023.

34 Lioba Heimbach, Vabuk Pahari, and Eric Schertenleib. Non-Atomic Arbitrage in Decentralized Finance . In *2024 IEEE Symposium on Security and Privacy (SP)*, pages 3866–3884, Los Alamitos, CA, USA, May 2024. IEEE Computer Society. doi:10.1109/SP54263.2024.00256.

35 hildobby. Atomic mev table. URL: https://dune.com/queries/3493305.

36 Edwin Hu and Dermot Murphy. Competition for retail order flow and market quality. *Available at SSRN 4070056*, 2022.

37 Robert A. Korajczyk and Dermot Murphy. Performance of algorithmic trading orders: Size, speed, and volume. *Review of Financial Studies*, 32(7):2829–2873, 2019.

38 Sorella Labs. Angstrom. URL: https://x.com/angstromxyz/status/1889725902851854520.

39 Jason Milionis, Ciamac C. Moallemi, and Tim Roughgarden. Automated market making and arbitrage profits in the presence of fees, 2023. arXiv:2305.14604.

40 Jason Milionis, Ciamac C Moallemi, Tim Roughgarden, and Anthony Lee Zhang. Automated market making and loss-versus-rebalancing. *arXiv preprint arXiv:2208.06046*, 2022.

41 Alex Nezlobin and Martin Tassy. Loss-versus-rebalancing under deterministic and generalized block-times, 2025. arXiv:2505.05113.

42 Alexandre Obadia, Alejo Salles, Lakshman Sankar, Tarun Chitra, Vaibhav Chellani, and Philip Daian. Unity is strength: A formalization of cross-domain maximal extractable value, 2021. arXiv:2112.01472.

43 Burak Öz, Danning Sui, Thomas Thiery, and Florian Matthes. Who wins ethereum block building auctions and why? In *6th Conference on Advances in Financial Technologies*, 2024.

44 Kaihua Qin, Liyi Zhou, Pablo Gamito, Philipp Jovanovic, and Arthur Gervais. An empirical study of defi liquidations: Incentives, risks, and instabilities. In *Proceedings of the 21st ACM Internet Measurement Conference*, pages 336–350, 2021. doi:10.1145/3487552.3487811.

45 Kaihua Qin, Liyi Zhou, and Arthur Gervais. Quantifying blockchain extractable value: How dark is the forest? In *2022 IEEE Symposium on Security and Privacy (SP)*, pages 198–214. IEEE, 2022. doi:10.1109/SP46214.2022.9833734.

46 Ultra Sound Relay. Ultra sound relay bid adjustment. URL: https://github.com/ultrasoundmoney/docs/blob/main/bid_adjustment.md.

47 Christoph Rosenmayr and Mateusz Dominiak. Statistical arbitrage on amms and block building on ethereum, 2024. URL: https://greenfield.xyz/2024/09/10/statistical-arbitrage-on-amms-and-block-building-on-ethereum-part-1/.

48 Andrei Shleifer and Robert W. Vishny. The limits of arbitrage. *The Journal of Finance*, 52(1):35–55, 1997.

49 Danning Sui and Fei Wu. Cex-dex transactions 202308-202503. URL: https://dune.com/queries/4931834.

50 Tardis.dev. Available token price data on binance. URL: https://api.tardis.dev/v1/exchanges/binance.

51 Tardis.dev. Binance spot historical data. URL: https://docs.tardis.dev/historical-data-details/binance.

52 Titan and Frontier Resarch. Builder dominance and searcher dependence. URL: https://frontier.tech/builder-dominance-and-searcher-dependence.

53 Anton Wahrstätter, Jens Ernstberger, Aviv Yaish, Liyi Zhou, Kaihua Qin, Taro Tsuchiya, Sebastian Steinhorst, Davor Svetinovic, Nicolas Christin, Mikolaj Barczentewicz, et al. Blockchain censorship. In *Proceedings of the ACM Web Conference 2024*, pages 1632–1643, 2024.

54 Anton Wahrstätter, Liyi Zhou, Kaihua Qin, Davor Svetinovic, and Arthur Gervais. Time to bribe: Measuring block construction market. *arXiv preprint arXiv:2305.16468*, 2023.

55 Toni Wahrstätter. mevboost.pics. URL: https://mevboost.pics/.

56 Fei Wu, Thomas Thiery, Stefanos Leonardos, and Carmine Ventre. From competition to centralization: The oligopoly in ethereum block building auctions. *arXiv preprint arXiv:2412.18074*, 2024. doi:10.48550/arXiv.2412.18074.

57 Fei Wu, Thomas Thiery, Stefanos Leonardos, and Carmine Ventre. Strategic bidding wars in on-chain auctions. In *2024 IEEE International Conference on Blockchain and Cryptocurrency (ICBC)*, pages 503–511, 2024. doi:10.1109/ICBC59979.2024.10634354.

58 Sen Yang, Kartik Nayak, and Fan Zhang. Decentralization of ethereum's builder market. In *2025 IEEE Symposium on Security and Privacy (SP)*, pages 1512–1530. IEEE, May 2025. doi:10.1109/sp61157.2025.00157.

59 Yuki Yuminaga, Dex Chen, and Danning Sui. Execution welfare across solver-based dexes, 2025. doi:10.48550/arXiv.2503.00738.

60 Mengqian Zhang, Sen Yang, and Fan Zhang. Rediswap: Mev redistribution mechanism for cfmms, 2024. doi:10.48550/arXiv.2410.18434.

61 Liyi Zhou, Kaihua Qin, Christof Ferreira Torres, Duc V Le, and Arthur Gervais. High-frequency trading on decentralized on-chain exchanges. In *2021 IEEE Symposium on Security and Privacy (SP)*, pages 428–445. IEEE, 2021.

62 Burak Öz, Christof Ferreira Torres, Christoph Schlegel, Bruno Mazorra, Jonas Gebele, Filip Rezabek, and Florian Matthes. Cross-chain arbitrage: The next frontier of mev in decentralized finance, 2025. arXiv:2501.17335.