Deviations from Tradition: Stylized Facts in the Era of DeFi

Daniele Maria Di Nosse[1], Federico Gatta[1], Fabrizio Lillo[1,2], Sebastian Jaimungal[3,4]

[1]Scuola Normale Superiore, Italy
[2]Università di Bologna, Italy
[3]University of Toronto
[4]Oxford-Man Institute for Quantitative Finance

federico.gatta@sns.it

SCUOLA NORMALE SUPERIORE

Summary

1. Introduction
2. Maximal Extractable Value
3. Time Series Analysis

Introduction

A Quick Recap

- Ethereum blockchain kickstarted **Decentralized Finance** (DeFi) thanks to **smart-contracts** (2015).
- **Decentralized Exchanges** (DEXs) replace Centralized Exchanges (CEXs) in DeFi using **Automated Market Makers**.
- **Uniswap v3**, on Ethereum, is the most popular DEX.

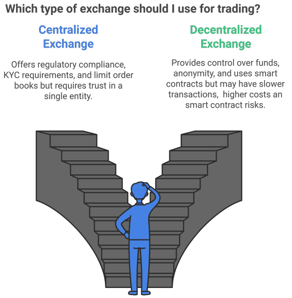

Figure 1: Centralized vs Decentralized Ex.

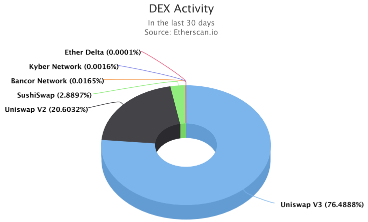

Figure 2: Distribution of activities between 23/03/2025 to 22/04/2025 in the major DEXs.

Decentralized Exchanges Organization

- DEXs are organized in **pools** -i.e., exchange desk with two **tokens**: X and Y.
- Key role played by token **reserves** R_X and R_Y.
- **Transaction** price determined according to a **constant product** rule.
- (Marginal) **Price** defined as the **reserves ratio**.

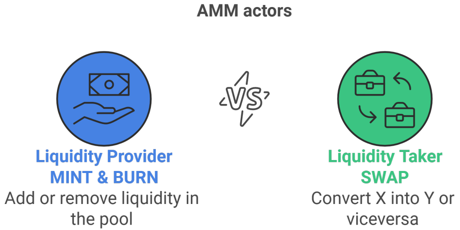

Figure 3: Main actors and operations in Decentralized Exchanges.

Uniswap v3 introduced the **concentrated liquidity**. It is possible to add liquidity within a **certain price range** defined as $[1.0001^i, 1.0001^j)$ with $i, j \in \mathbb{Z}$ called **ticks**. The interval $j - i$ is known as the **tick range**.

SCUOLA NORMALE SUPERIORE

What is in this work? Why?

We aim to understand how the new market framework introduced by blockchain based platforms like DEXs changes the statistical properties of the price and liquidity. Thus, we perform an analysis of these quantities at very high frequencies.

Motivation

CEX - Limit Order Book	**DEX - AMM**
Mid price	Marginal price as $S := \frac{R_X}{R_Y}$
Transactions validated continuously in time.	Lumpy order flow: blocks sparsely validated.
All the orders can change the price.	All and only the swaps move the price.
Submission transaction order matches the execution transaction order.	Time is not sequential: submission and execution orders can differ.
Loans always require collateral.	Flashloans allow retail traders to perform sophisticated strategies without upfront capital.

Possible applications

- Reference study for researchers and practitioners.
- Starting point for **generative models for AMM**.

The Framework

We focus on the Uniswap v3 market in the period between 2023-01-01 and 2024-12-31.

- By far, the **most traded** DEX.
- **Galore of data** for significant analysis.

Data about 15 pairs distributed across 24 pools and clustered into 4 groups:

- **Normal** - Involving a volatile token and a stablecoin -e.g. `USDC-WETH`.
- **Volatile** - Two volatile tokens -e.g. `WBTC-WETH`.
- **Stable** - Made up of two stablecoins -e.g. `USDC-USDT`.
- **Synthetic** - Defined by different representations of the same asset, mainly related to liquid staking -e.g. `wstETH-WETH`.

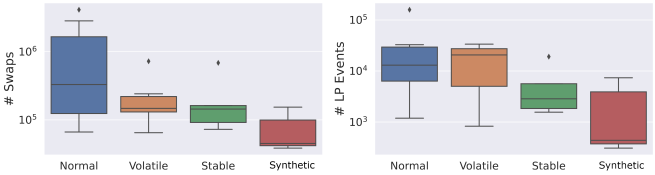

Figure 4: Amount of available data. The pools are divided into clusters

Measuring Time

We consider three timescales.

- **Event-time**: the time increases by one when a new event is recorded in the blockchain. Specifically, it makes sense to divide **swap-time** and **mint/burn-time**.
- **Tick-time**: the time increases by one when the price tick changes. The price tick is defined as $\lfloor \log_{1.0001} S \rfloor$; a tick is said to be "active" if some LP uses it as the bound of a position. Crossing an active tick could change the liquidity profile and increase the gas fee.

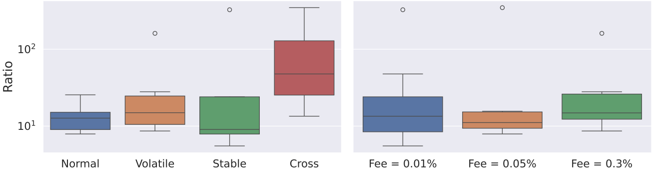

Figure 5: Average number of swaps for crossing an active tick.

- **Block-time**: the time increases by one when a new block is added to the blockchain.

Intraday Patterns

We study the dynamics of the number of events during each day. The pipeline is:

- For every day, compute the daily percentages of the **Modulated Realized Variance** over a time window of 30 minutes.
- **Regress onto** the first five **Fourier** basis **coefficient**.
- Study the **average** representation in the Fourier space over all the days.

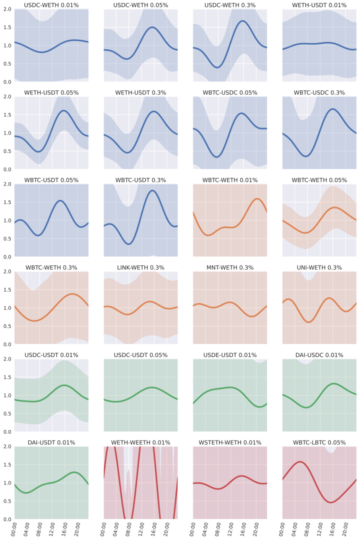

Figure 6: Market activity. The lineplot is for the mean; the shadow area is for $\pm\frac{1}{2}$ standard deviations.

Intraday Patterns

We study the dynamics of the number of events during each day. The pipeline is:

- For every day, compute the daily percentages of the **Modulated Realized Variance** over a time window of 30 minutes.
- **Regress onto** the first five **Fourier** basis **coefficient**.
- Study the **average** representation in the Fourier space over all the days.

Some pools show a common pattern characterized by a **decrease in activity around 08 am** (UTC time) and an **increase between 03 and 05 pm**.

Question:

Discretionary trading or correlation with other markets?

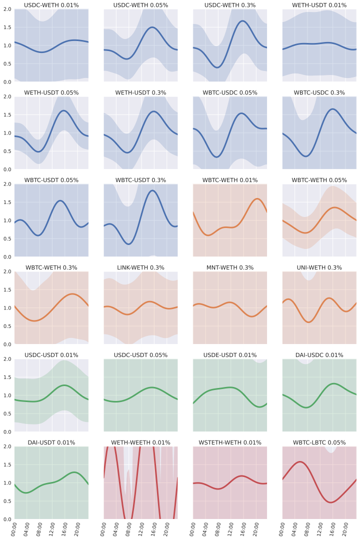

Figure 6: Market activity. The lineplot is for the mean; the shadow area is for $\pm\frac{1}{2}$ standard deviations.

Maximal Extractable Value

Behind the Scenes - From Users to Blockchain

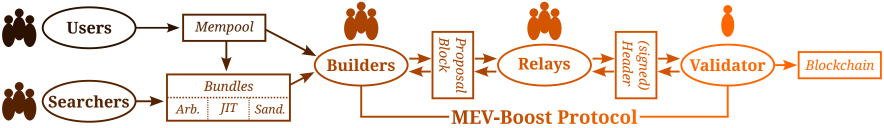

Figure 7: From Agents to Blockchain.

- The agent's **order** is received by an Ethereum node and **stored** in a **mempool**.
- The **searchers find profitable** transactions from the mempool to be added to the block -economic arguments.
- The **block builders** collect info from the searchers and build a block proposal.
- The **validator** collects the block proposal, **chooses** the best bid, and **adds** it to the **blockchain**.

Maximal Extractable Value (MEV)

MEV is the **profit** that can be made by **manipulating** the **transactions** in the mempool.

Some strategies:

- **Arbitrage** – Exploiting price mismatches between different pools or exchanges.
- **Sandwich Attacks** – Front running: Placing trades around a large swap to profit from its price impact.
- **Just-in-Time (JIT) Liquidity** – Temporarily providing liquidity before a large trade to earn fees, then withdrawing it immediately.
- **Mixed Attacks** – Made up of a sandwich attack surrounding a JIT.

Figure 8: Focus on Sandwich attacks and JIT liquidity.

MEV Impact - Price

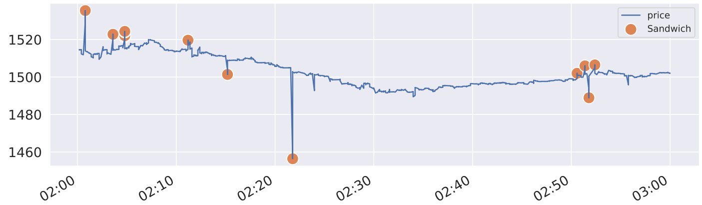

Figure 9: USDC-WETH 0.05% price (swap-time) on March 11, 2023, between 2 and 3 am (UTC).

MEV Impact - Price

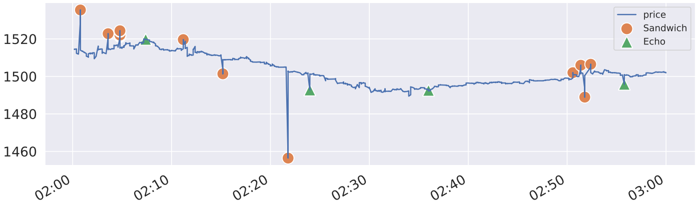

Figure 10: USDC-WETH 0.05% price (swap-time) on March 11, 2023, between 2 and 3 am (UTC).

Echo Swaps

Pair of consecutive (swap or event-time) swaps, placed by the same agent within the same block and with opposite signs.

Question

What is the economic driver of Echo swaps?
Main intuition: part of **exogenous MEV strategies**.

MEV Impact - Pools Activity

Why does it make sense to focus on MEV strategies?

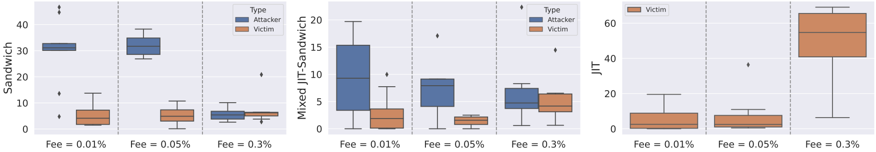

Figure 11: Volume percentage linked to MEV strategies -attackers (blue) and victims (orange).

MEV Impact - Pools Activity

Why does it make sense to focus on MEV strategies?

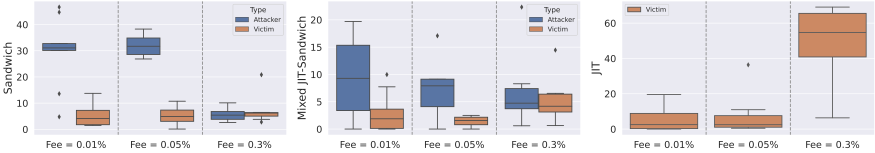

Figure 11: Volume percentage linked to MEV strategies -attackers (blue) and victims (orange).

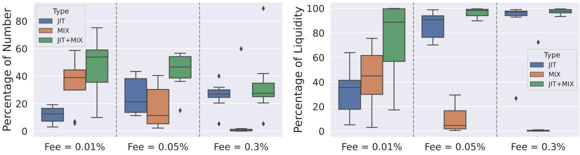

Figure 12: Impact of MEV strategies on LP events -number (left) and liquidity minted (right). Shown both JIT (blue), mixed attacks (orange) and their sum (green).

Time Series Analysis

Returns Autocorrelation

Peculiar returns **AutoCorrelation Function** (ACF) computed at swap-time.

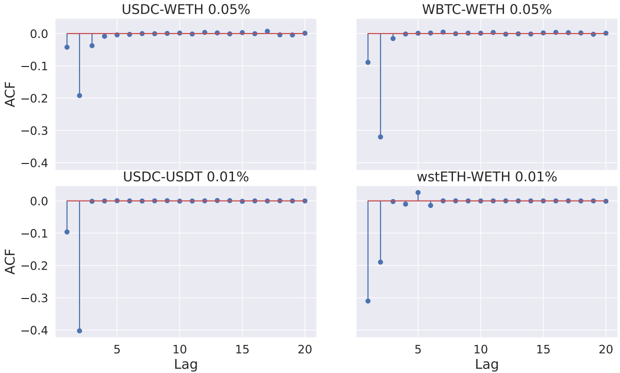

Figure 13: Returns ACF at swap-time.

Strong negative autocorrelation in **swap-time** at **lags 1**, **2**, and **3** for half of the pools. This dynamics **disappears** in **block-time**.

Returns Autocorrelation - Explaining Peaks

- In **TradFi**, negative peak at lag one due to the **bid-ask bounce.** Don't hold for DEXs.
- **Lag 2 peaks** stem from **sandwich** attacks.

Clean MEV strategies:

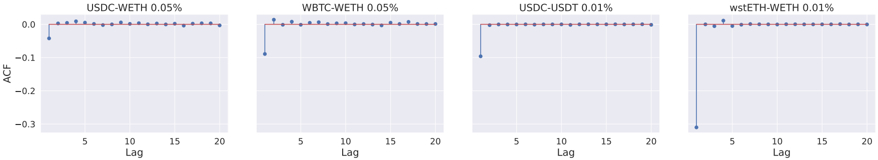

Figure 14: Returns ACF at swap-time -conditional on the absence of sandwich and echo swaps.

Returns Autocorrelation - Explaining Peaks

- In **TradFi**, negative peak at lag one due to the **bid-ask bounce**. Don't hold for DEXs.
- **Lag 2 peaks** stem from **sandwich** attacks.

Clean MEV strategies:

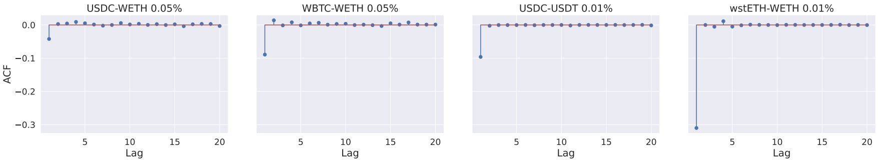

Figure 14: Returns ACF at swap-time -conditional on the absence of sandwich and echo swaps.

How can we explain the ACF at lag 1?

Reverse arbitrage (Capponi and Jia - 2025)

Arbitrage opportunity generated by the price **impact of noisy traders**.

Naive test - **conditional ACF** on:

- **abs return**$_{t-1}$ $>$ **pool fee:** recovered the negative part.
- **abs return**$_{t-1}$ $\leq$ **pool fee:** positive.

Long Memory Time Series

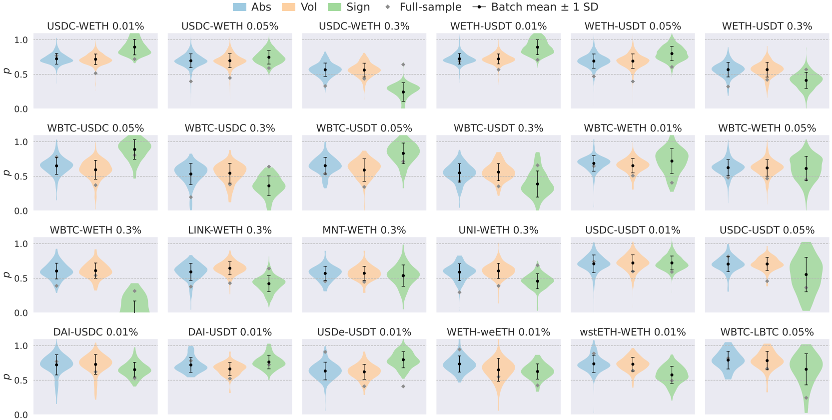

Figure 15: p coefficients from the power-law relationship $ACF(L) \sim A \cdot L^{-p}$.

Abs is for the absolute returns; **Vol** is for the volume; **Sign** is the trade direction. The distributions are obtained via non-overlapping time windows.

Why long memory in the sign ACF?

CEXs usually display long memory time series. Main hypotheses:

- **Meta-orders** (Lillo and Farmer - 2004). This is the most respected explanation for CEXs. However, its validity for DEXs is in doubt as, due to the different market structure, meta-orders are probably not profitable (Angeris et al. - 2021).
- **Herd effect** (LeBaron and Yamamoto - 2007). That is, strong persistence can be due to a follower-leader mechanism. Most reliable, as cryptocurrencies are characterized by many strong price trends attributable to herding effects.

Decomposing the ACF (Tóth et al. - 2015)

Key idea: decompose the ACF into two components. One accounts for swaps from the same wallet; the other for swaps from different wallets.

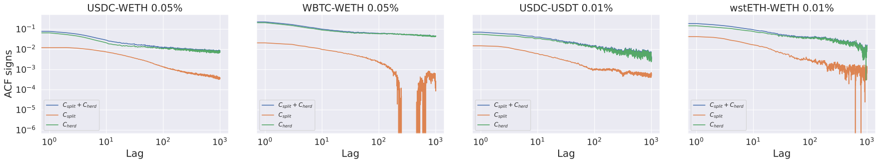

Figure 16: ACF (blue line) decomposition into split (orange) and herd (green) components.

LPs Behavior against Volatility

Cross-pool study of the LP decision as a function of the pool marginal price volatility.

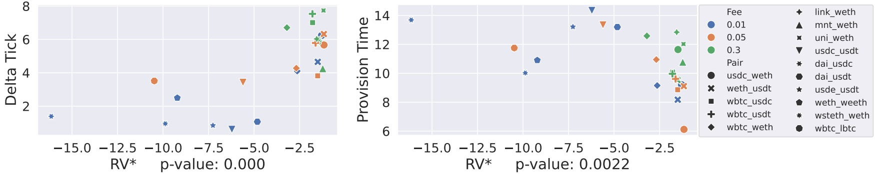

Figure 17: Realized Variance vs range size (left) and time length (right).

- p-value is for the **Spearman correlation test** with permutation.
- The realized variance is computed by removing outlier returns originating from microstructure.
- The range size and provision time are **normalized by** the amount of **liquidity provided**.
- **JITs** are **excluded** from the computation.

Transition Probabilities

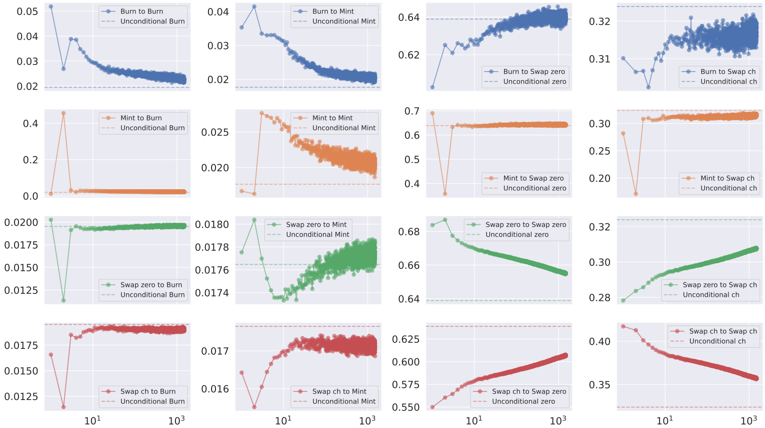

Figure 18: USDC-WETH 0.05% fee - Transition probabilities $\mathbb{P}(e_{t+l} = E | e_t = E')$, with $E, E' \in \{Burn, Mint, Swap\ zero, Swap\ ch\}$. *Swap zero* and *Swap ch* are for swap events without or with tick change.

Transition Probabilities - Without Attackers

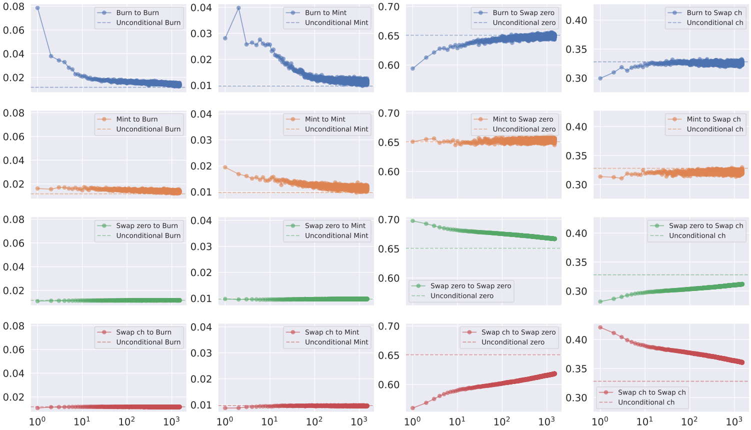

Figure 19: USDC-WETH 0.05% fee - Transition probabilities (sandwich, mix, and JIT attacks removed) $\mathbb{P}(e_{t+l} = E | e_t = E')$, with $E, E' \in \{Burn, Mint, Swap\ zero, Swap\ ch\}$. *Swap zero* and *Swap ch* are for swap events without or with tick change.

Conclusion

Summary

To summarize, the analysis is carried out on 15 pairs distributed across 24 pools.

Some other findings in the paper:

- **Market activity** shows a recurrent **intraday pattern**: low activity around 8 a.m.; high activity around 4 p.m., UTC time.
- The series of **trade signs** is generally **long-memory**. As the opposite of TradFi, this is due to the **herd effect**.
- **Log-volume** and **log-variance** are typically linearly related.
- The active liquidity exhibits **droughts** leading to artificial, **abnormal price return**.
- **Transition probabilities** are strongly impacted by JIT events. After cleaning, the diagonal effect is clearly visible.

Thanks for your attention!