Can Automated Market Makers Provide Price Discovery?

Luke Johnson[1]

Ester Félez-Viñas[1] , Tālis Putniņš[12]

[1]University of Technology Sydney, DFCRC

[2]Stockholm School of Economics in Riga

37th Australasian Finance and Banking Conference

11 December 2024

Motivation

Successful market: Liquidity and price discovery

- Liquidity providers (LPs) in automated market makers (AMM) let traders set price through a mathematical formula
- In LOB markets quotes provide the bulk of price discovery (Brogaard, Hendershott, and Riordan, 2019)
- LPs face increased adverse selection which theory suggests to be a driver of illiquidity (Kyle, 1985; Glosten and Milgrom, 1985)
    - AMMs should be illiquid and inefficient

Can AMMs provide price discovery?

This paper

Price discovery: Information → Prices

- When AMMs are the only market can they provide price discovery?
- Can AMMs lead price discovery?

AMM design creates constant arbitrage opportunities

- AMM price just follows the LOB market

Compare Uniswap and Binance (Largest AMM and LOB)

- Amount of information in prices (AMM only)
- Where that information enters first (AMM vs LOB)

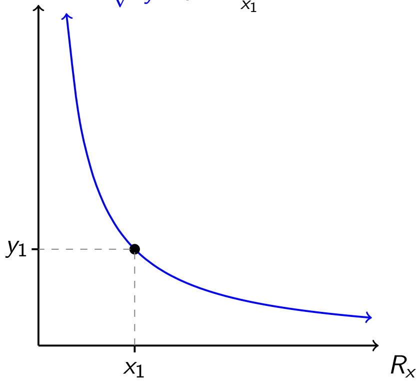
So what?

AMMs have potential to change the future of financial markets

- Facilitate the exchange of tokenized assets
  - ◇ Cross border FX payments
  - ◇ Project Mariana (Bank for International Settlements, 2023)
- More efficient for high volume, low volatility assets (Foley, O'Neill, and Putniņš, 2023)
- Save billions on transaction costs (Malinova and Park, 2023)
- Improve stability of liquidity (Adams et al., 2023)

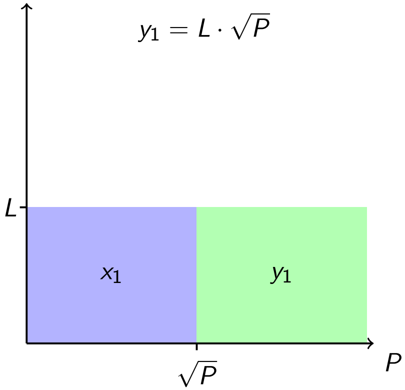
Constant product market maker (CPMM)

$$R_y \qquad L = \sqrt{xy} \qquad P = \frac{y_1}{x_1}$$

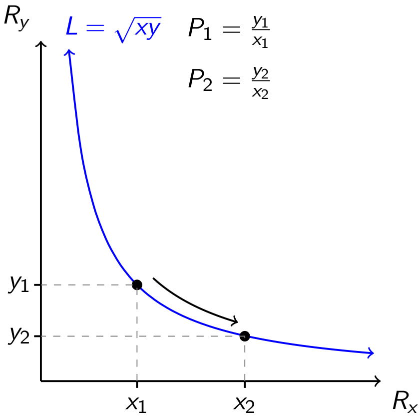


$$L \qquad x_1 = L \cdot \frac{1}{\sqrt{P}}$$

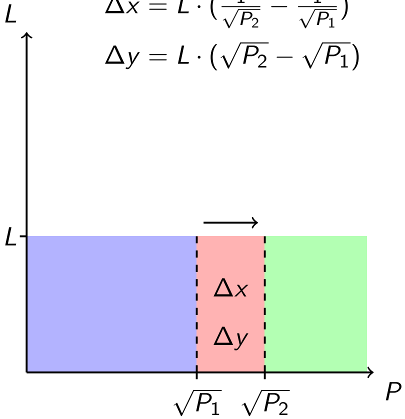
$$y_1 = L \cdot \sqrt{P}$$



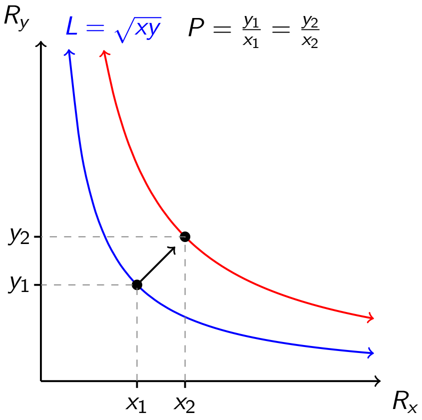
Swapping in a CPMM

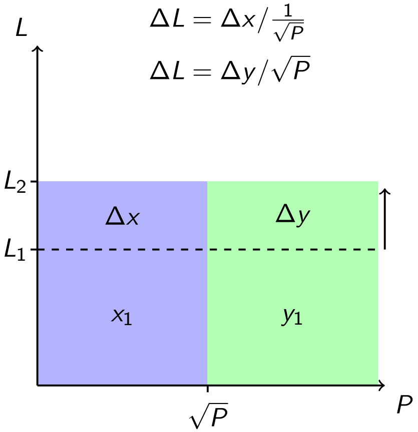


$$\Delta x = L \cdot \left(\frac{1}{\sqrt{P_2}} - \frac{1}{\sqrt{P_1}}\right)$$
$$\Delta y = L \cdot \left(\sqrt{P_2} - \sqrt{P_1}\right)$$

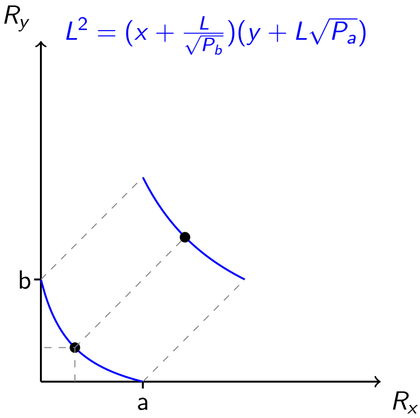


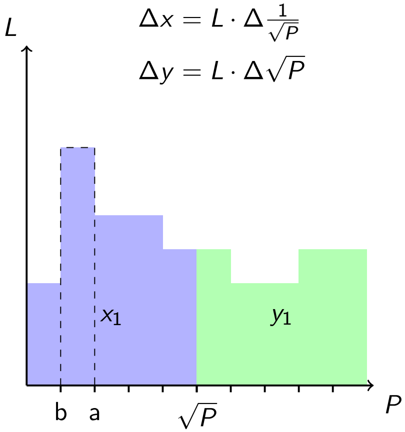
Changing liquidity in a CPMM





Concentrated liquidity





Can AMMs provide price discovery?

When AMMs are the only market can they provide price discovery?

- Measure amount of information in prices
- Compare sample of assets that trade only on AMM and pairs that trade on AMM and LOB
- Matched sample based on market, year, month and nearest neighbor by volume up to 20% threshold

Measuring the amount of information in prices

Variance Decomposition (Brogaard et al., 2022)

- Innovations in price that can be explained by sources of market, private or public information
- Asset returns: $r_t = \mu + \theta_{r_m}\varepsilon_{r_m,t} + \theta_x\varepsilon_{x,t} + \theta_r\varepsilon_{r,t} + \Delta s_t$
- Structural VAR with market returns (r_m), signed dollar volume (x) and returns (r) 12 lags hourly frequency
  - θ's from cumulative impulse response function
  - ε's from the VAR residuals
- *MarktInfoShare*, *PrivateInfoShare*, *PublicInfoShare*, and *NoiseShare*

Variance decomposition shares

Panel A. AMM and LOB Summary Statistics	Mean	Standard Deviation	p10	Median	p90
MarketInfoShare (%)	7.44	14.86	0.04	2.24	17.58
PrivateInfoShare (%)	42.35	25.03	4.99	45.22	73.88
PublicInfoShare (%)	5.45	9.37	0.21	2.14	11.72
NoiseShare (%)	44.76	25.11	17.59	39.67	88.37
Panel B. AMM Only Summary Statistics					
MarketInfoShare (%)	5.64	11.67	0.05	1.23	14.77
PrivateInfoShare (%)	45.02	21.87	11.47	48.85	71.57
PublicInfoShare (%)	4.80	5.73	0.21	2.99	12.47
NoiseShare (%)	44.53	19.82	20.07	43.34	72.32

Testing if AMMs can provide price discovery?

To test if AMMs can provide price discovery we use the following regression model:

$$VarShare_{i,t} = \alpha_{i,t} + \beta_1 AMMOnly_{i,t} + \beta_2 Liquidity_{i,t} + \beta Controls_{i,t} + \epsilon_{i,t}$$

Where *VarShare* is a measure of variance shares such as the *NoiseShare*, *MarketInfoShare*, *PrivateInfoShare*, or the *PublicInfoShare*.

Focus on the *NoiseShare*

- If an AMM on its own cant provide price discovery the noise share should be higher (less information in price)

Variance shares regressions

Dependent variable:	*NoiseShare*	*MarketInfoShare*	*PrivateInfoShare*	*PublicInfoShare*
AMM Only	−0.036**	−0.004	0.039**	0.002
	(−2.26)	(−0.42)	(2.09)	(0.31)
EffectiveSpread	0.092***	0.002	−0.103***	0.009
	(5.51)	(0.27)	(−6.74)	(0.95)
$Depth_{[1.5\%]}$	0.004	0.014*	−0.020**	0.002
	(0.76)	(1.84)	(−2.22)	(0.56)
Volume	−0.035***	0.014**	0.026**	−0.006
	(−4.59)	(2.14)	(2.50)	(−1.53)
RealizedVolatility	0.052***	−0.024***	−0.017	−0.012*
	(4.61)	(−3.52)	(−1.35)	(−1.95)
ConcLiquidity	0.121***	0.027	−0.123***	−0.025***
	(4.76)	(1.53)	(−4.07)	(−2.64)
Controls	Y	Y	Y	Y
Observations	522	522	522	522
Adjusted R^2	39.3%	20.3%	20.8%	6.4%

Significance: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Can AMMs provide price discovery?

No economically significant difference in the *NoiseShare* between AMM only listings and AMM and LOB listings

- Significantly less (<5%) *NoiseShare* in AMM only pairs
- Listings are not random, However this work against our findings
  - List tokens that contain information
  - Sample matching based on volume

Liquidity is negatively related with the *PrivateInfoShare*

- LPs are sensitive to adverse selection

AMMs can provide price discovery when they are the only market

- Support low cost trading

Can AMMs lead price discovery?

Informed traders face a venue selection choice

- Profit maximisation: Implicit cost and explicit costs
- More liquid AMM (LOB) have lower implicit costs
- Informed traders are sensitive to gas fees on Ethereum (Capponi, Jia, and Yu, 2023)
- Information is revealed to block builder in AMM
- Ambiguous effect of volatility

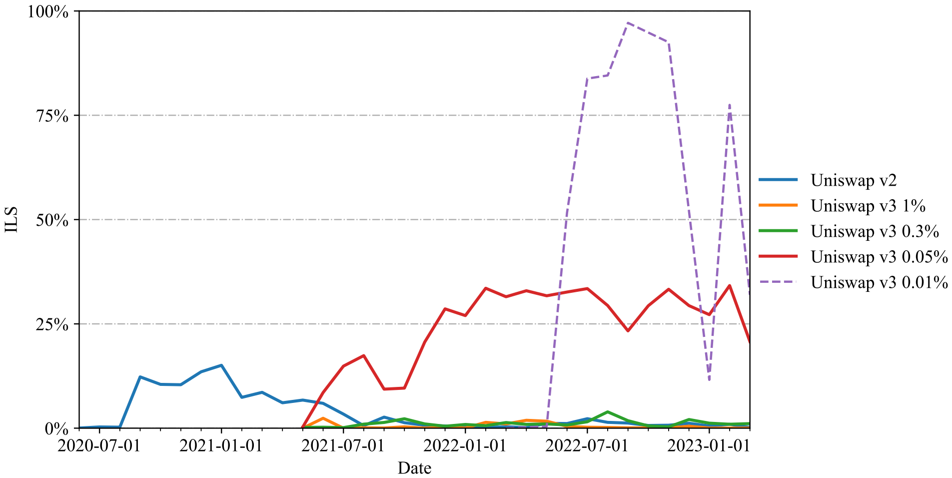
Measuring price leadership

Comparing trading pairs that trade on both the AMM and the LOB

Extend Yan-Zivot-Putnins Information Leadership Share (*ILS*) to *n* markets (Putniņš, 2013)

- Significant difference in volatility between markets
  - ◇ Information Share (IS) and Component Share (CS) both measure a relative avoidance of noise
  - ◇ IS is also reported for completeness
- VECM with 300 lags ($\sim$1 hour)
  - ◇ Mid-quote prices for all Uniswap markets and Binance
  - ◇ Ethereum blocktime (12 seconds*)

Information Leadership share



What AMM determines price leadership?

Average ILS for AMM pool is 8.7%

- Uniswap v3 0.05% has an ILS of 24.1%
- Uniswap v3 0.01% has an ILS of 61.1% (FUNToken)

Statistically test AMM price leadership

- $PDS_{i,t} = \alpha_{i,t} + Liquidity_{i,t} + Volatility_{i,t} + Gas_t + Controls_{i,t} + \epsilon_{i,t}$

Insight on the informed trader venue selection choice

- Driven by costs: Implicit and explicit

Determinants of price discovery

Dependent variable:	*ILS*	*ILS*	*IS*	*IS*
QuotedSpread	−0.085***	−0.070***	−0.085***	−0.070***
	(−9.54)	(−9.04)	(−9.54)	(−9.04)
EffectiveSpread	−0.006	0.000	−0.006	0.000
	(−1.52)	(0.03)	(−1.52)	(0.03)
$Depth_{[1.5\%]}$	−0.005***	−0.003*	−0.005***	−0.003*
	(−3.49)	(−1.84)	(−3.49)	(−1.84)
$DepthVariance_{[1.5\%]}$	−0.017***	−0.001	−0.017***	−0.001
	(−4.82)	(−0.43)	(−4.82)	(−0.43)
Volume	0.015***	0.020***	0.015***	0.020***
	(5.35)	(6.86)	(5.35)	(6.86)
RealizedVolatility	0.015***	−0.011**	0.015***	−0.011**
	(3.95)	(−2.12)	(3.95)	(−2.12)
Gas	−0.018***		−0.018***	
	(−5.32)		(−5.32)	
Pair Effects	N	Y	N	Y
Month Effects	N	Y	N	Y
Observations	1,577	1,577	1,577	1,577
Adjusted R^2	28.2%	47.1%	28.2%	47.1%

Significance: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Including the limit order book metrics

Dependent variable:	*ILS*	*ILS*	*IS*	*IS*
QuotedSpread	−0.079***	−0.068***	−0.079***	−0.068***
	(−9.45)	(−8.93)	(−9.45)	(−8.93)
$Depth_{[1.5\%]}$	−0.004***	−0.003*	−0.004***	−0.003*
	(−2.96)	(−1.71)	(−2.96)	(−1.71)
Volume	0.025***	0.022***	0.025***	0.022***
	(8.67)	(7.36)	(8.67)	(7.36)
LOBQuotedSpread	0.016***	0.004	0.016***	0.004
	(5.18)	(1.43)	(5.18)	(1.43)
LOBEffectiveSpread	0.005	0.022***	0.005	0.022***
	(0.76)	(3.03)	(0.76)	(3.03)
$LOBDepth_{[1.5\%]}$	−0.026***	−0.021***	−0.026***	−0.021***
	(−3.90)	(−2.79)	(−3.90)	(−2.79)
LOBVolume	0.008*	−0.011	0.008*	−0.011
	(1.66)	(−1.63)	(1.66)	(−1.63)
Controls	Y	Y	Y	Y
Pair Effects	N	Y	N	Y
Month Effects	N	Y	N	Y
Observations	1,577	1,577	1,577	1,577
Adjusted R^2	34.4%	48.6%	34.4%	48.6%

Significance: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

AMM price leadership

- AMM price leadership increasing with AMM liquidity
- AMM price leadership decreasing with gas fees
- AMM price leadership decreasing with LOB liquidity
- Venue selection choice for informed traders impacted by both implicit and explicit costs

Conclusion

Paper examines if AMMs can provide price discovery

- Find that AMMs when they are the only market to trade an asset can provide price discovery
- Show that LPs in AMMs are sensitive to adverse selection
- Support AMMs as a use case for low cost trading infrastructure

Appendix

Descriptive Statistics

Panel A. Limit Order Book Summary Statistics	Mean	Standard Deviation	p10	Median	p90
Volume ($)	4,532	20,109	3	24	4,018
Trades	3,966.2	19,541.8	16.1	85.8	4,803.5
Mean Trade ($)	480	486	113	294	1,223
Median Trade ($)	147	161	30	90	370
Realized Volatility (%)	1.46	2.02	0.24	0.85	3.41
Half Quoted Spread (%)	0.21	0.17	0.04	0.18	0.40
Effective Spread (%)	0.25	0.30	0.03	0.18	0.51
$\text{Depth}_{[1.5\%]}$ ($)	1.81	6.13	0.01	0.07	3.30
$\text{Depth Variance}_{[1.5\%]}$ (%)	45.34	24.52	23.77	40.04	71.33

Descriptive Statistics

Panel B. Automated Market Maker Summary Statistics					
	Mean	Standard Deviation	p10	Median	p90
Volume ($)	298	1,596	1	8	455
Trades	9.5	31.5	0.4	1.7	15.0
Mean Trade ($)	15,395	35,404	1,348	5,101	29,865
Median Trade ($)	7,167	19,007	393	2,266	12,649
Realized Volatility (%)	17.43	211.33	0.22	1.53	11.43
Half Quoted Spread (%)	0.33	0.20	0.30	0.30	0.30
Effective Spread (%)	4.57	17.67	0.44	1.50	4.30
$\text{Depth}_{[1.5\%]}$ ($)	2.47	14.13	0.00	0.02	1.38
$\text{Depth Variance}_{[1.5\%]}$ (%)	36.98	54.22	6.74	18.54	82.68

Variance decomposition estimation

Two key inputs are needed to construct the components of variance which we can get by estimating a reduced form structural VAR model: the variance of the innovations in each variable, $\sigma^2_{\epsilon_{rm}}$, $\sigma^2_{\epsilon_x}$, $\sigma^2_{\epsilon_r}$ and the long-run cumulative responses to these shocks θ_{r_m}, θ_x, θ_r.

$$r_{m,t} = a_0^* + \sum_{l=1}^{12} a_{1,l}^* r_{m,t-l} + \sum_{l=1}^{12} a_{2,l}^* x_{t-l} + \sum_{l=1}^{12} a_{3,l}^* r_{t-l} + e_{r_m,t}$$

$$x_t = b_0^* + \sum_{l=1}^{12} b_{1,l}^* r_{m,t-l} + \sum_{l=1}^{12} b_{2,l}^* x_{t-l} + \sum_{l=1}^{12} b_{3,l}^* r_{t-l} + e_{x,t}$$

$$r_t = c_0^* + \sum_{l=1}^{12} c_{1,l}^* r_{m,t-l} + \sum_{l=1}^{12} c_{2,l}^* x_{t-l} + \sum_{l=1}^{12} c_{3,l}^* r_{t-l} + e_{r,t}$$

Variance decomposition estimation

The reduced form residuals can be written as linear models of the structural-model residuals:

$$
\begin{aligned}
e_{r_m,t} &= \epsilon_{r_m,t} \\
e_{x,t} &= \epsilon_{x,t} + b_{1,0}\epsilon_{r_m,t} = b_{1,0}e_{r_m,t} + \epsilon_{x,t} \\
e_{r,t} &= \epsilon_{r,t} + (c_{1,0} + c_{2,0}b_{1,0})\epsilon_{r_m,t} + c_{2,0}\epsilon_{x,t} = c_{1,0}e_{r_m,t} + c_{2,0}e_{x,t} + \epsilon_{r,t}
\end{aligned}
$$

Although the structural-model residuals are uncorrelated contemporaneously by design, the Equation shows that the reduced-form residuals exhibit contemporaneous correlation. This correlation allows for inference about the structural-model residuals. We use linear regressions to estimate the parameters $b_{1,0}$, $c_{1,0}$ and $c_{2,0}$

Variance decomposition estimation

Using the estimated parameters $b_{1,0}$, $c_{1,0}$ and $c_{2,0}$, along with the estimated variances of the reduced-form residuals $\sigma^2_{e_{rm}}$, $\sigma^2_{e_x}$, and $\sigma^2_{e_r}$ we derive estimates for the variances of the structural model shocks by rearranging the variance expression of the Equation.

$$
\begin{aligned}
\sigma_{\epsilon_{rm}} &= \sigma_{e_{rm}} \\
\sigma_{\epsilon_x} &= \sigma_{e_x} - b^2_{1,0}\sigma^2_{e_{rm}} \\
\sigma_{\epsilon_r} &= \sigma_{e_r} - (c^2_{1,0} + 2c_{1,0}c_{2,0}b_{1,0})\sigma^2_{e_{rm}} - c^2_{2,0}\sigma^2_{e_x}
\end{aligned}
$$

Variance decomposition estimation

To estimate the long-run cumulative impulse response functions of the structural model, we compute the equivalent reduced-form shocks and feed them through the reduced-form model. Specifically:

1. A structural shock to market returns $[\epsilon_{r_m,t}, \epsilon_{x,t}\epsilon_{r,t}]' = [1, 0, 0]'$ has a reduced-form equivalent $[e_{r_m,t}, e_{x,t}e_{r,t}]' = [1, b_{1,0}, (c_{1,0} + c_{2,0}b_{1,0})]'$
2. A structural shock to market returns $[\epsilon_{r_m,t}, \epsilon_{x,t}\epsilon_{r,t}]' = [0, 1, 0]'$ has a reduced-form equivalent $[e_{r_m,t}, e_{x,t}e_{r,t}]' = [0, 1, c_{2,0}]'$
3. A structural shock to market returns $[\epsilon_{r_m,t}, \epsilon_{x,t}\epsilon_{r,t}]' = [0, 0, 1]'$ has a reduced-form equivalent $[e_{r_m,t}, e_{x,t}e_{r,t}]' = [0, 0, 1]'$

The cumulative return response to each of these shocks, evaluated at $t = 36$ (the point where the responses stabilize), provides estimates for θ_{r_m}, θ_x, and θ_r respectively.

Variance decomposition estimation

Taking the variance of the innovations in the efficient price we get $\sigma_w^2 = \theta_{r_m}^2 \sigma_{\epsilon_{r_m}}^2 + \theta_x^2 \sigma_{\epsilon_x}^2 + \theta_r^2 \sigma_{\epsilon_r}^2$. The errors in the structural model are contemporaneously uncorrelated by construction and therefore the covariance terms are all zero. The contribution to the variation in the efficient price from each of the information components is $\theta_{r_m}^2 \sigma_{\epsilon_{r_m}}^2$ (market-wide information), $\theta_x^2 \sigma_{\epsilon_x}^2$ (private firm-specific information), and $\theta_r^2 \sigma_{\epsilon_r}^2$ (public firm-specific information). The estimated components of variance are therefore

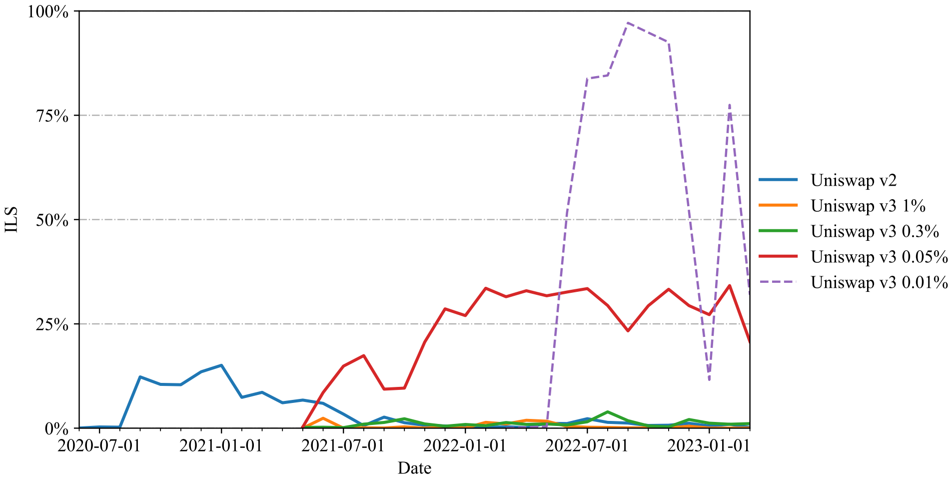
$$
\begin{aligned}
MarketInfo &= \theta_{r_m}^2 \sigma_{\varepsilon_{r_m}}^2 \\
PrivateInfo &= \theta_x^2 \sigma_{\varepsilon_x}^2 \\
PrivateInfo &= \theta_r^2 \sigma_{\varepsilon_r}^2 \\
Noise &= \sigma_s^2
\end{aligned}
$$

Variance decomposition estimation

Normalizing these variance components to sum to 100% gives variance shares:

$$MarketInfoShare = \theta_{r_m}^2 \sigma_{\varepsilon_{r_m}}^2 / (\sigma_w^2 + \sigma_s^2)$$
$$PrivateInfoShare = \theta_x^2 \sigma_{\varepsilon_x}^2 / (\sigma_w^2 + \sigma_s^2)$$
$$PrivateInfoShare = \theta_r^2 \sigma_{\varepsilon_r}^2 / (\sigma_w^2 + \sigma_s^2)$$
$$NoiseShare = \sigma_s^2 / (\sigma_w^2 + \sigma_s^2)$$

Information Share



Price discovery estimation

For each pair-month, we estimate a reduced form VECM of the log price series ($p_{1,t}$ to $p_{n,t}$) with 300 lags (prices are sampled based on the Ethereum block time where trading is continuous in the AMM and the LOB).

$$\Delta p_t = aZ_{t-1} + \sum_{i=1}^{300} b_i \Delta p_{t-i} + \epsilon_t \tag{1}$$

where Δp_t is the $n \times 1$ midquote return vector, α is the $n \times (n-1)$ matrix of error correction coefficients, Z_{t-1} is the $n \times 1$ co-integrating vector, b_i is the $n \times n$ coefficient matrix for lag i and ϵ_t is the $n \times 1$ vector of residuals.

Price discovery estimation

From the reduced form VECM estimates in 1 we derive the corresponding infinite lag VMA representation in structural form assuming recursive contemporaneous causality running from the first through to the last price series.

$$
\begin{aligned}
\Delta p_{1,t} &= \sum_{l=0}^{\infty} A_{1,l}\varepsilon_{1,t-1} + \sum_{l=0}^{\infty} A_{2,l}\varepsilon_{2,t-1} + \cdots + \sum_{l=0}^{\infty} A_{n,l}\varepsilon_{n,t-1} \\
\Delta p_{2,t} &= \sum_{l=0}^{\infty} B_{1,l}\varepsilon_{1,t-1} + \sum_{l=0}^{\infty} B_{2,l}\varepsilon_{2,t-1} + \cdots + \sum_{l=0}^{\infty} B_{n,l}\varepsilon_{n,t-1} \\
\vdots & \qquad\qquad\qquad \vdots \qquad\qquad\qquad \vdots \\
\Delta p_{n,t} &= \sum_{l=0}^{\infty} N_{1,l}\varepsilon_{1,t-1} + \sum_{l=0}^{\infty} N_{2,l}\varepsilon_{2,t-1} + \cdots + \sum_{l=0}^{\infty} N_{n,l}\varepsilon_{n,t-1}
\end{aligned}
$$

Price discovery estimation

We obtain the structural VMA coefficients by computing the orthogonalized impulse response functions and the (contemporaneously uncorrelated) structural VMA errors ($\varepsilon_{1,t}$ to $\varepsilon_{n,t}$) by mapping their relation to the reduced form errors. Innovations in the permanent component (the efficient price, m_t) are given by

$$\Delta m_t = \theta_{\varepsilon 1}\varepsilon_{1,t} + \theta_{\varepsilon 2}\varepsilon_{2,t} + \cdots + \theta_{\varepsilon n}\varepsilon_{n,t}$$

The variance of the innovations in the efficient price is therefore:

$$\begin{aligned} Var(\Delta m_t) &= Var(\theta_{\varepsilon 1}\varepsilon_{1,t} + \theta_{\varepsilon 2}\varepsilon_{2,t} + \cdots + \theta_{\varepsilon n}\varepsilon_{n,t}) \\ &= \theta_{\varepsilon 1}^2 Var(\varepsilon_{1,t}) + \theta_{\varepsilon 2}^2 Var(\varepsilon_{2,t}) + \cdots + \theta_{\varepsilon n}^2 Var(\varepsilon_{n,t}) \end{aligned}$$

Price discovery estimation

Information shares (IS) are obtained as each price's contribution to the variance of the efficient price innovations

$$IS_n = \frac{\theta_{\varepsilon n}^2 Var(\varepsilon_{n,t})}{Var(\Delta m_t)}$$

Component shares (CS) are obtained by normalizing the permanent price impacts of each price series in the reduced form model.

$$CS_n = \frac{\theta_{\epsilon n}}{\sum_{i=1}^{n} \theta_{\epsilon i}}$$

Price discovery estimation

Finally, we calculate the information leadership share (ILS). In the two-price case, market's propensity to reflect new information (how much market *i*s price responds to an innovation in the efficient price) can be obtained from the ratio $\beta_i = \frac{IS_i}{CS_i}$, which when normalized gives the information leadership share

$$ILS_n = \frac{\beta_n^2}{\sum_{i=1}^{n} \beta_i^2}$$

Relative determinants of price discovery

Dependent variable:	*ILS*	*ILS*	*IS*	*IS*
QuotedSpreadRatio	−0.046***	−0.040***	−0.041***	−0.038***
	(−7.34)	(−6.94)	(−7.91)	(−7.96)
EffectiveSpreadRatio	0.022***	0.011*	0.014**	0.006
	(3.39)	(1.87)	(2.54)	(1.11)
$Depth_{[1.5\%]}Ratio$	−0.003	−0.006**	−0.001	−0.005***
	(−1.46)	(−2.50)	(−0.90)	(−2.72)
VolumeRatio	0.016***	0.026***	0.017***	0.029***
	(6.04)	(7.27)	(7.62)	(9.48)
RealizedVolatility	−0.022***	−0.019***	−0.011**	−0.007
	(−3.80)	(−3.04)	(−2.30)	(−1.18)
Gas	−0.010**		−0.012***	
	(−2.45)		(−3.24)	
Controls	Y	Y	Y	Y
Pair Effects	N	Y	N	Y
Month Effects	N	Y	N	Y
Observations	1,577	1,577	1,577	1,577
Adjusted R^2	18.2%	39.8%	24.3%	45.3%

Significance: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$