Natural Centralization in Decentralized Finance

Pablo D. Azar Adrian G. Casillas Maryam Farboodi

December 9, 2025

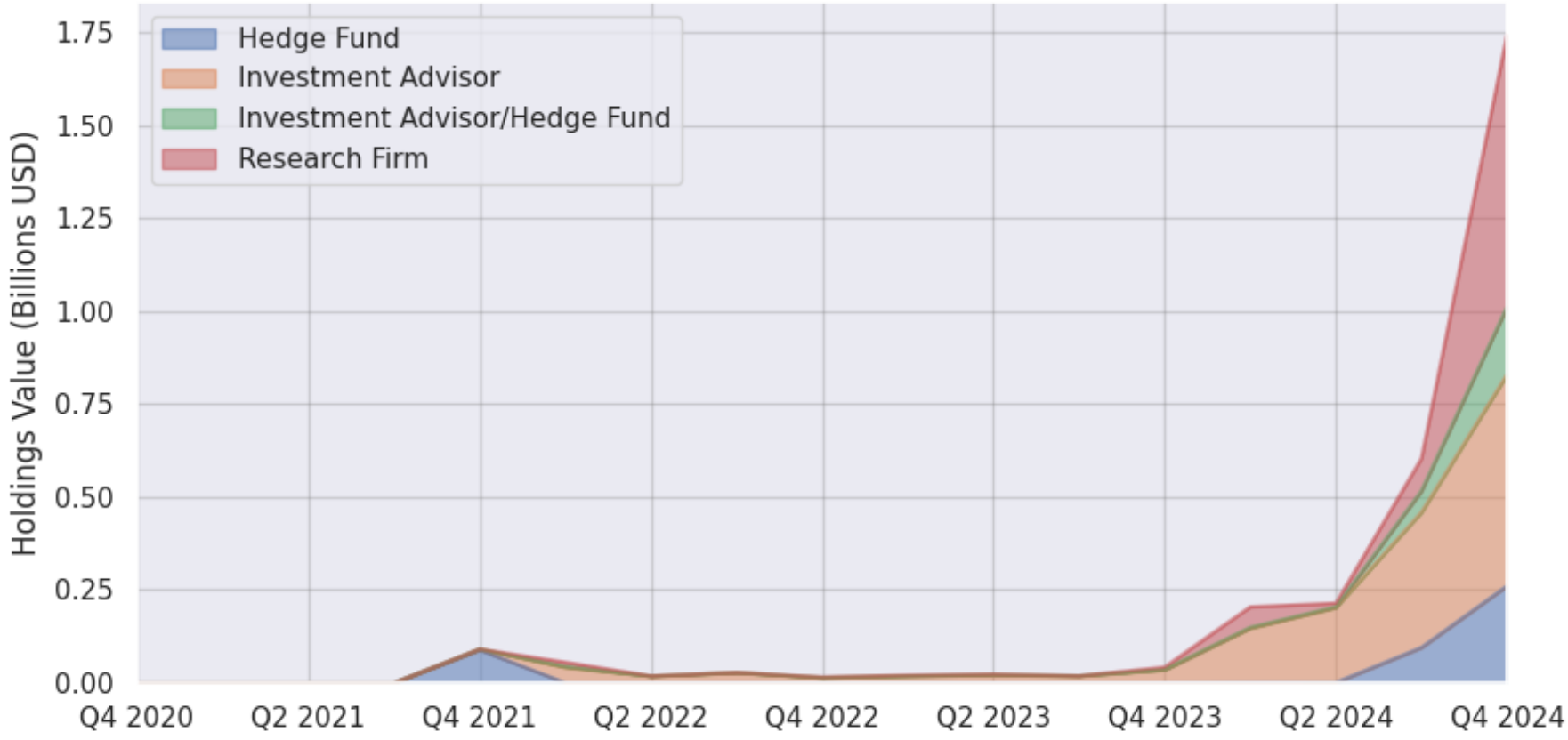
Natural Centralization in Decentralized Markets

- Canonical Economic Models: free entry yields perfect competition.
- Yet real-world markets have significant intermediaries, rents, and concentration.
- We use DeFi to show how frictions and intermediaries arise endogenously—even in markets with free entry.

Market Structure Questions

- How are profit shares determined along intermediation chains?
- What drives concentration and market power across layers?
- DeFi data enables causal tests of information-driven market power.

Interconnections with Traditional Finance



Major Developments: Crypto–TradFi Integration

- United States: Spot BTC/ETH ETFs; GENIUS Act; custody/tokenization build-outs.
- Europe: MiCA taking effect with EU-wide licensing; DLT Pilot Regime enabling tokenized securities; regulated crypto ETPs on major venues.
- Rest of World: UK FMI/Digital Securities sandboxes; Hong Kong spot crypto ETFs; Singapore stablecoin framework/licensing; UAE, Brazil, Japan tokenization initiatives.

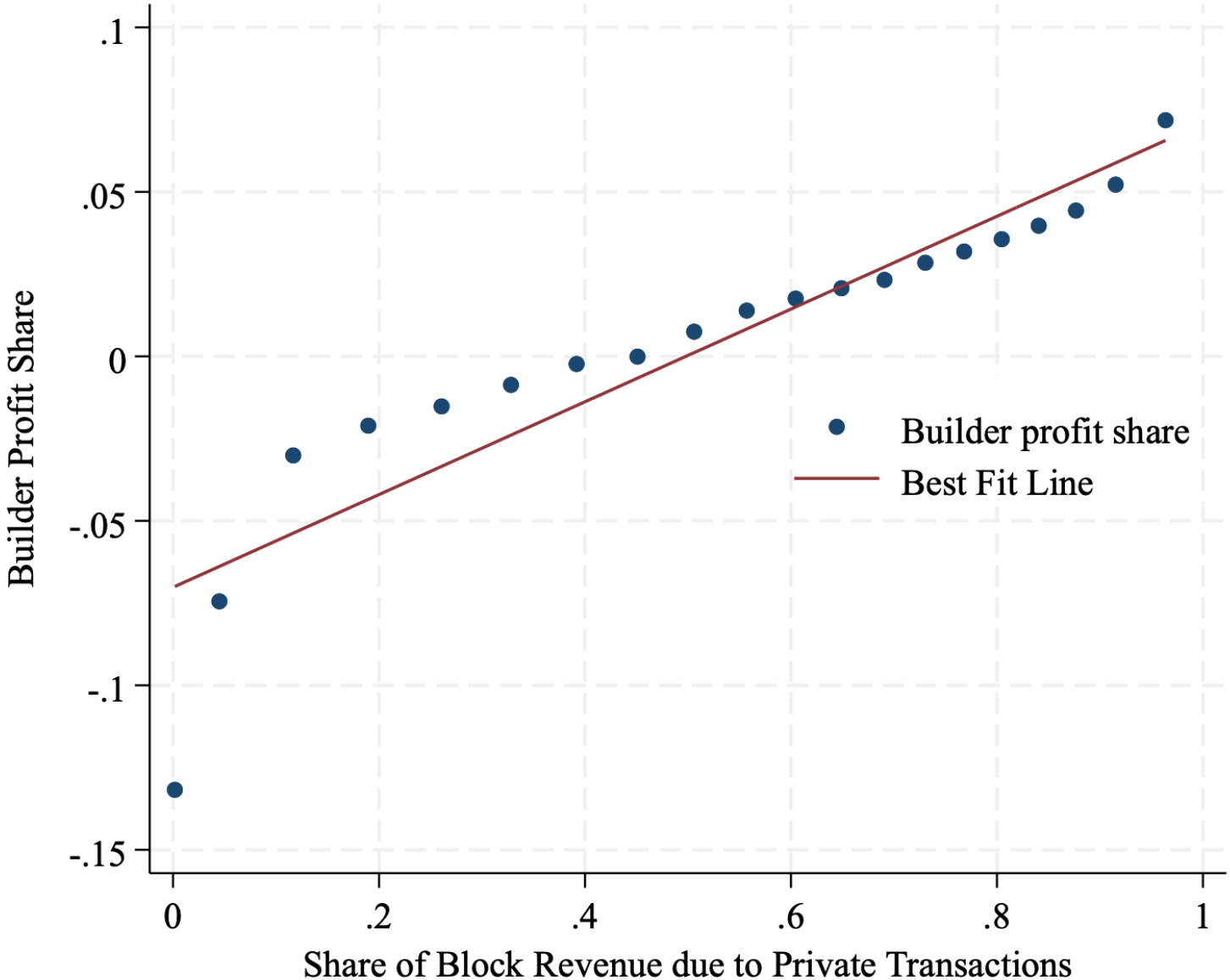
Risks to Financial Stability

- Increasing Interconnections with Traditional Finance
- Cryptoization
- Resurrection of Shadow Banking
- Illicit Financial Flows

This Paper: DeFi Intermediation

- The DeFi intermediation chain: arbitrageurs, builders, proposers, ETH holders.
- Private vs. public information and their impact on profit sharing and concentration.
- A new limit to arbitrage: privacy needs vs. blockchain transparency.
- New bargaining model with free entry and endogenous bargaining power.

Preview of Empirical Results



Literature Review

- *Theory of financial intermediation:* Leland and Pyle (1977), Diamond and Dybvig (1983), Diamond (1984), Allen and Gale (1997), Boot and Thakor (1997), Diamond and Rajan (2001)
- *information in financial intermediation:* Shleifer and Vishny (1997), Glode and Opp (2006)
- *Empirical documentation of intermediation rents in traditional finance:* Green et al. (2007), Di Maggio et al. (2017), Hollifield et al. (2017), Li and Schürhoff (2019), Farboodi et al. (2019)

Literature Review 2

- *Blockchain technology in economics and finance:*
  - *PoW:* Raskin and Yermack (2018), Abadi and Brunnermeier (2018), Biais et al. (2019), Cong et al. (2021a,b), Lehar and Parlour (2023)
  - *PoS:* Choi et. al. (2023), Saleh (2021), Socu and Saleh (2021), Fanti et al. (2019)
- *Cryptocurrency markets:* Makarov and Schoar (2020, 2021), Cong et al. (2021c), Kim (2022), Huberman et al. (2021), Ferreira et al. (2023), Capponi et al. (2023), Gupta et al. (2023), Heimbach et al. (2024), Liu et al. (2023), Gandal and Halaburda (2016), Bakos et. al. (2021), Halaburda(2025)

The Ethereum Blockchain

- Ethereum is the largest blockchain that allows for DeFi protocols
- Two main types of transactions: simple token transfers, and smart contract interactions
- Smart contracts are blockchain-stored programs which execute pre-defined functions when tokens are sent to them
  - Swapping tokens, Providing Liquidity, Depositing Collateral, Lending Tokens, etc...

Risk-Free Arbitrage

- The large number of competing DeFi protocols creates violations of the law of one price
  - e.g. different prices across exchanges, or fire sales from collateral liquidation
- Ethereum allows for atomic transaction bundles: either all transactions in a bundle execute or none do
- Atomic transaction bundles guarantee risk-free arbitrages

New Limit to Arbitrage: The Need for Privacy

- Even when an arbitrageur identifies a discrepancy, they need to have their transaction executed
- If the arbitrageur does not act quickly, the opportunity will disappear
- However, the arbitrageur cannot broadcast their transaction to the network, or it will be stolen
- Reputable DeFi intermediaries arise due to this need for privacy

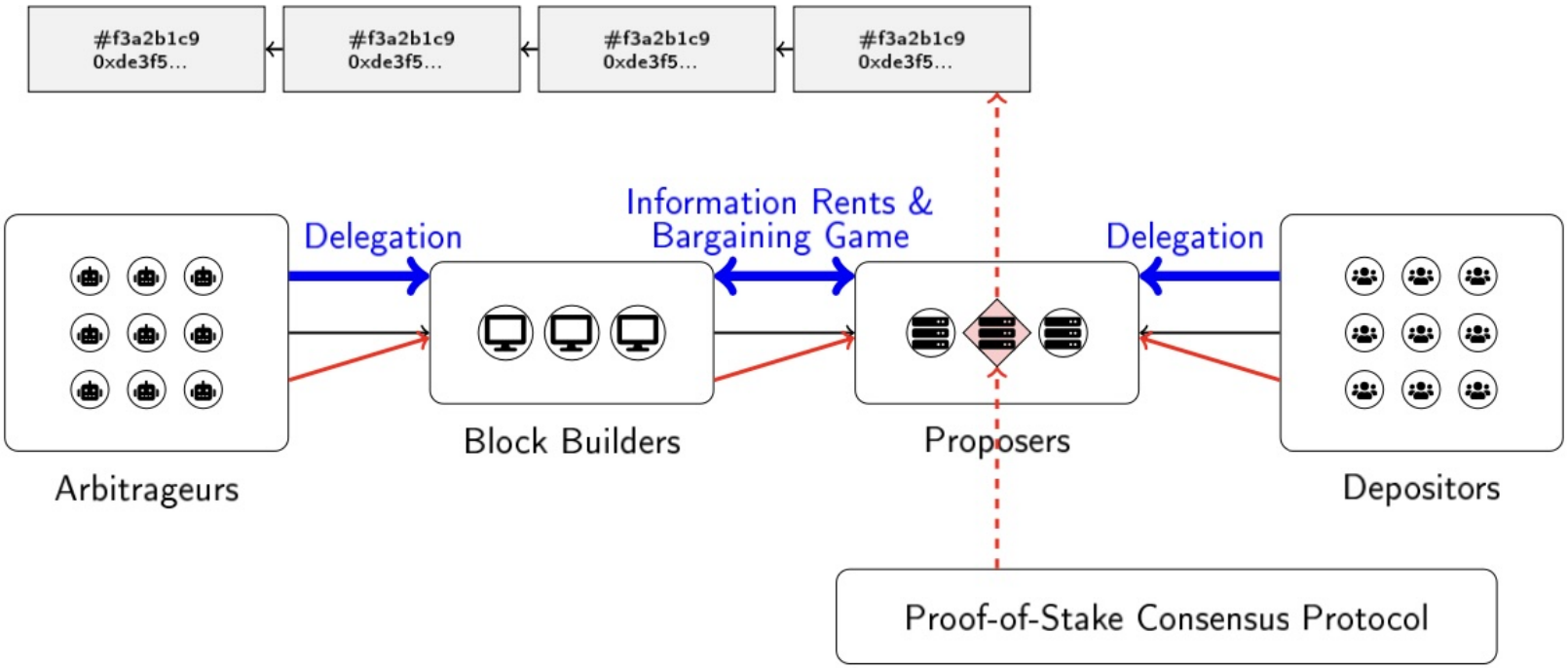
Emergence of the DeFi Intermediation Chain

- **Arbitrageurs** find profitable trades
- **Block Builders** aggregate transactions into blocks and bid for a slot on the chain
- **Block Proposers** choose one of the proposed builder blocks and appends it to the blockchain
- **ETH Holders** delegate their ETH stake to proposers

Market Concentration Across the Chain

- **Arbitrageurs** Crypto Hedge Funds such as Wintermute and SCP
- **Block Builders** 2 builders create more than 50% of the blocks
- **Block Proposers** 4 Proposers have over 50% of the stake (Coinbase, Binance, Kraken, Lido)
- **ETH Holders** contribute to concentration by pooling their shares

The DeFi Intermediation Chain



▸ market concentration

Why Information Rents?

- At each time t:
  - Selected proposer sees nothing except the set of bids by block builders
  - Each block builder sees every transaction in their block
    - They see a larger pool of transactions and choose among them
- All builders have an information advantage over the proposer:
  - Information advantage might be shared among all builders (public)
  - It might be unique to a particular builder (private)
- **Does private information the builder has over other builders affect share of profit builder gets when bargaining with proposer?**

Private versus Public Information

Definition

Public Transaction: *A transaction in the block added at time t is public if it is broadcast to the network before time t and is not a payment to the builder of this block.*

Private Transaction: *Every other transaction!*

- Public transactions are non-exclusive among builders

Data

- Period: September 15, 2022 (switch to proof-of-stake) - September 30, 2024
- MEV blocks in Ethereum: 3,903,112 blocks (73%) ▸ MEV
  - Builder different from proposer and use MEV-boost
- Dune Analytics:
  - Block-level data
  - Identity of the builder and proposer
  - Total revenue and split between the builder and proposer
- **Mempool** Guru project (Secure Decentralized Systems Lab):
  - Tracks whether transactions were broadcast to the network before being appended to the blockchain
  - Enables identification of public vs. private transactions

Variables

- B_t: block appended to blockchain at time t

Rev_t	Total revenue of block builder and proposer from block t
$\Pi_{B,t}$	Revenue of block builder
$\Pi_{P,t}$	Revenue of block proposer
$\theta_{B,t}$	$\frac{\Pi_{B,t}}{Rev_t}$, Revenue share of block builder
$\theta_{P,t}$	$\frac{\Pi_{P,t}}{Rev_t}$, Revenue share of proposer
$\log Public_t$	Total value of public transactions in block t
$\log Private_t$	Total value of private transactions in block t

Block-Level Summary Statistics

	Mean	Std. Dev.	Min	5th	Median	95th	99th	Max	Skewness	Kurtosis
Rev_t	0.14	1.49	0.00	0.02	0.05	0.35	1.28	802.79	227.79	82779.54
$\Pi_{B,t}$	0.02	0.66	-0.30	-0.00	0.00	0.03	0.17	791.74	615.44	613508.02
$\Pi_{P,t}$	0.12	1.22	0.00	0.02	0.05	0.32	1.09	691.96	259.88	108182.04
$\theta_{B,t}$	0.05	0.12	-0.10	-0.04	0.01	0.29	0.57	1.00	3.27	16.25
$\theta_{P,t}$	0.95	0.12	0.00	0.71	0.99	1.04	1.08	1.10	-3.27	16.25
log $Private_t$	0.08	0.17	0.00	0.00	0.03	0.27	0.78	6.69	8.82	128.58
log $Public_t$	0.02	0.05	0.00	0.01	0.02	0.06	0.14	5.20	27.87	1461.36
Hack Dummy	0.07	0.25	0.00	0.00	0.00	1.00	1.00	1.00	3.43	12.74
Crisis Dummy	0.01	0.11	0.00	0.00	0.00	0.00	1.00	1.00	8.56	74.21
Observations	3903112									

- Total revenue, profits, and profit shares very skewed
- Builder profit even more skewed

▸ Normal/Pareto

Information-Driven Market Power

- **Do builders with superior information compared to other builders get a higher share of profit from the proposers?**
- Plain-vanilla reduced-form evidence

$$\theta_{B,t} = \alpha + \beta \ln \mathit{Private}_t + \epsilon_t$$

Potential Endogeneity Concerns

1. **Omitted variable bias:**
   1. Relationships between builders and proposers may lead builders to treat some proposers differently:
      - Solution: Include fixed effects
   2. Other block characteristics can affect $\theta_{B,t}$: any other characteristic impacts the share of builder only through total block revenue
      - Solution: Include value of public transactions $Public_t$

$$\theta_{B,t} = \beta \ln Private_t + \gamma \ln Public_t + \psi_{i(t)} + \eta_{j(t)} + \phi_{i(t),j(t)} + \epsilon_t$$

2. **Simultaneity:**
   - Builders simultaneously decide:
     - Which transactions to insert into the block
     - Payment to the proposer
   - Equivalent to deciding:
     - How much private information to capitalize in this block
     - How much of that value to share with the proposer

Instrument

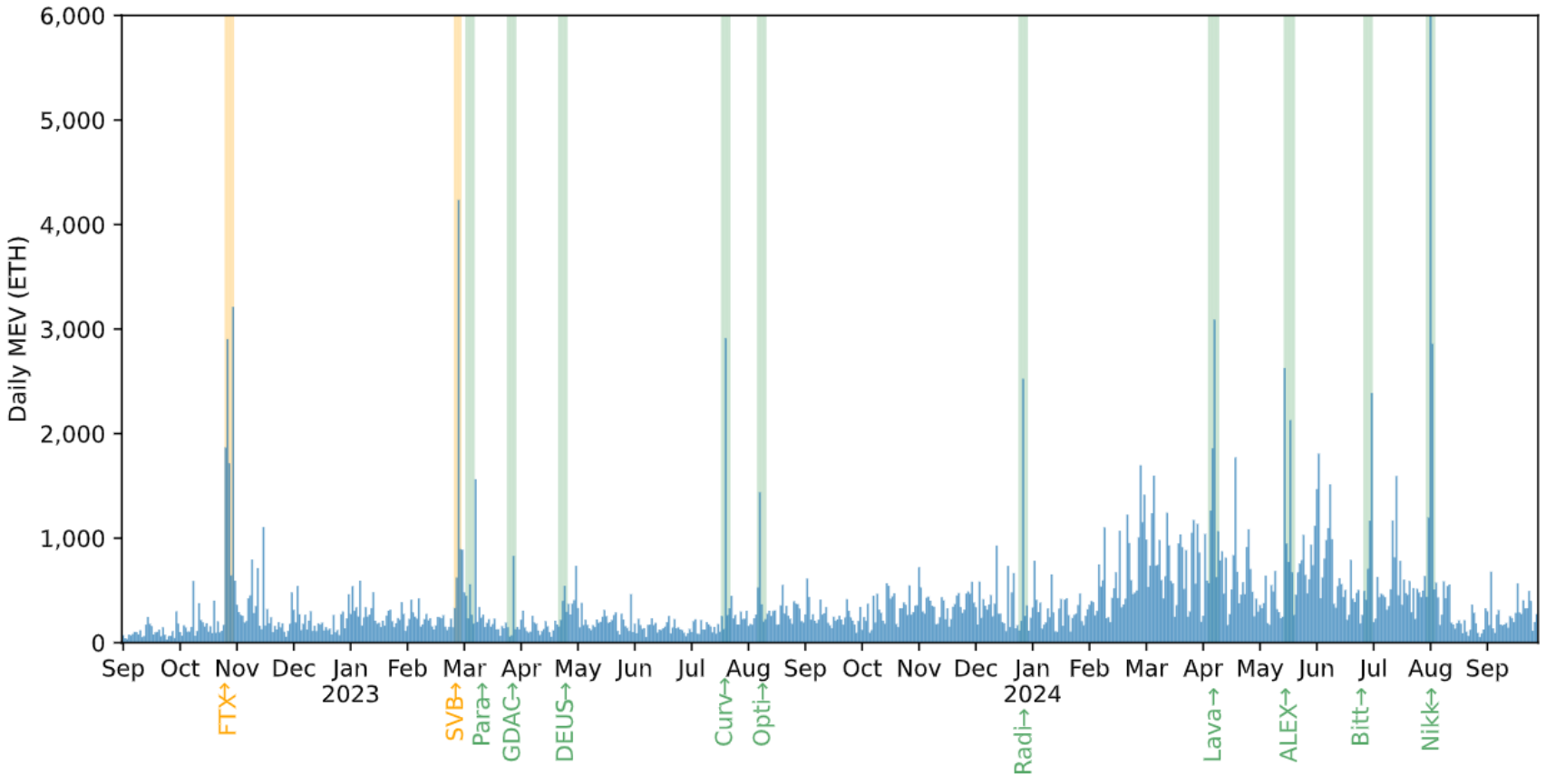
- Instrument for total value of private and public transactions
- Two instruments, both dummies:
  1. $\textbf{Hacked}_\textbf{t} = \textbf{1}$ if the block is appended on a day where there is a crypto protocol hack: 50 hacks
  2. $\textbf{Crisis}_\textbf{t} = \textbf{1}$ if the block is appended during either the FTX or SVB crises: 9 days

$$\ln Private_t = \hat{\beta}_1 Hacked_t + \hat{\gamma}_1 Crisis_t + \hat{\psi_{1,i(t)}} + \hat{\eta_{1,j(t)}} + \hat{\phi_{1,i(t),j(t)}} + \hat{\epsilon_{1,t}};$$

$$\ln Public_t = \hat{\beta}_2 Hacked_t + \hat{\gamma}_2 Crisis_t + \hat{\psi_{2,i(t)}} + \hat{\eta_{2,j(t)}} + \hat{\phi_{2,i(t),j(t)}} + \hat{\epsilon_{2,t}};$$

- **Exclusion restriction:** Hacks and crises affect the share of builders only through the value of private and public transactions available to include in the block

Instruments and Total Daily Revenue



- Crises and hacks are first-order institutional features of DeFi
- Decentralization makes rolling back bugs/thefts highly unlikely

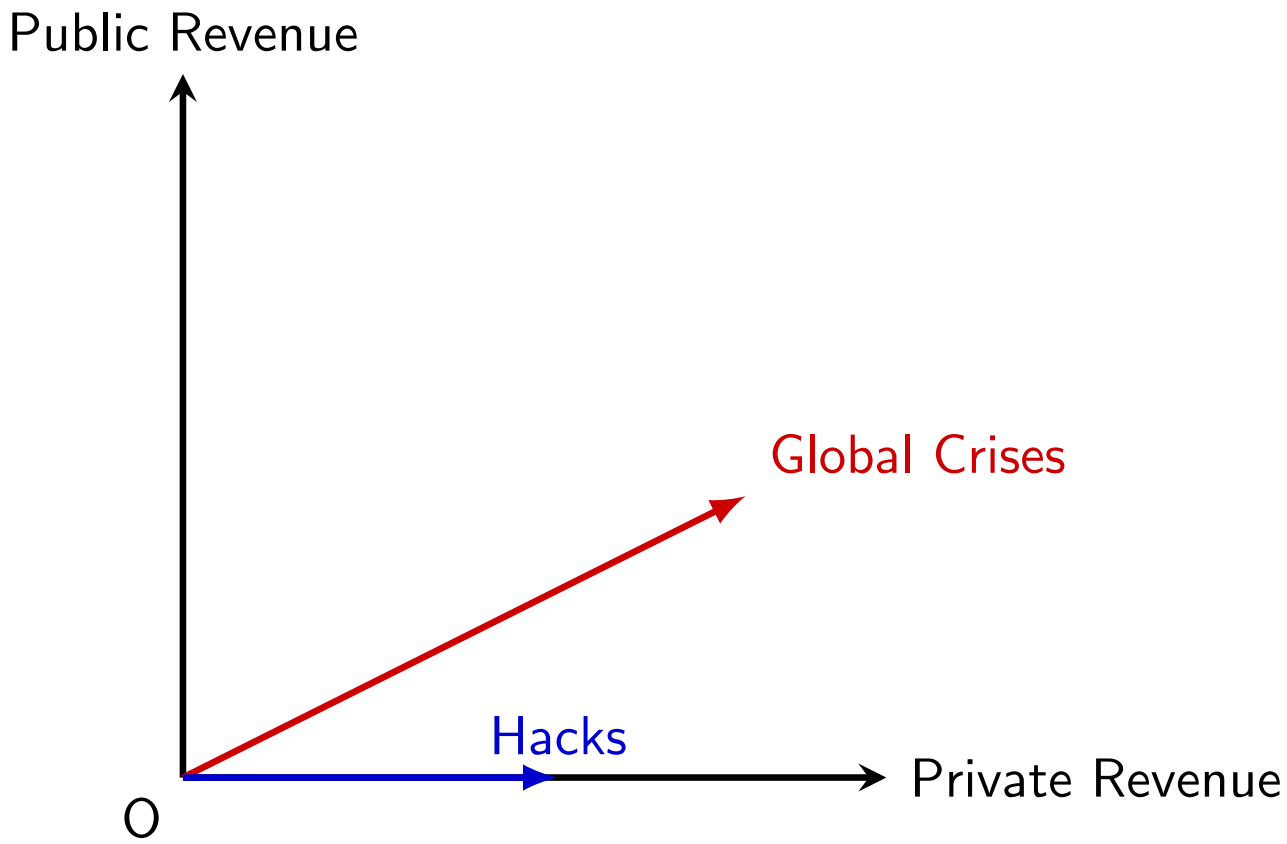
First Stage

	(1) log $Private_t$	(2) log $Private_t$	(3) log $Public_t$	(4) log $Public_t$
Hack Dummy	0.0107***	0.0101***	-0.0012***	-0.0000
	(0.0007)	(0.0008)	(0.0002)	(0.0002)
Crisis Dummy	0.1187***	0.1228***	0.0218***	0.0136***
	(0.0095)	(0.0098)	(0.0014)	(0.0014)
Constant	0.0732***		0.0240***	
	(0.0016)		(0.0007)	
Observations	3903112	3446649	3903112	3446649

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

First Stage: Spanning Instruments



Note: Vector lengths not drawn to scale

Second Stage

	(1) OLS No FE	(2) OLS FE	(3) IV No FE	(4) IV FE
log $Private_t$	0.153***	0.141***	0.401***	0.476***
	(0.0127)	(0.0156)	(0.0582)	(0.105)
log $Public_t$	-0.206***	-0.106***	-2.932***	-3.924***
	(0.0147)	(0.0141)	(0.334)	(0.967)
Constant	0.0404***		0.0877***	
	(0.00292)		(0.00598)	
N	3903112	3446649	3903112	3446649
F Statistic			461.35	66.29
Robust F Statistic			80.507	15.013

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Empirical Finding

1. Higher value of private transactions in a given block:
   - Causally increases the profit share of the block builder
2. Higher value of public transactions, controlling for the value of private transactions:
   - Causally decreases the profit share of the block builder

Model Overview

- Bargaining between proposers and builders
- Three transaction types:
  1. Regular (low value, frequent)
  2. Public arbitrage (high value, all builders know)
  3. Private arbitrage (high value, single builder knows)

Full Model

Builder's Bargaining Power

- Private information $\rightarrow$ credible threat to wait
- Builder can:
  - Include transaction now at proposed terms
  - Wait for better terms in future period
- Public information $\rightarrow$ no bargaining power
  - Competition drives builder profits to zero

Key Result

Proposition (Information-Driven Market Power)

Private transactions:

- *Increase builder's profit share*

Public transactions:

- *Decrease builder's profit share*

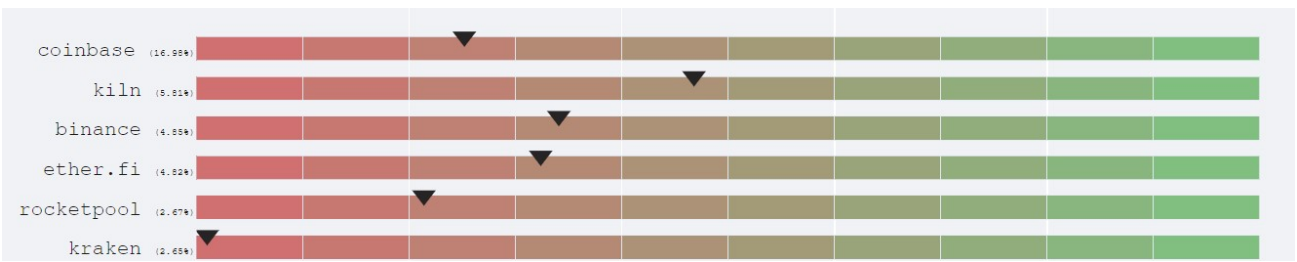
Conclusion

- Private information is a powerful driver of market power in DeFi
- Despite decentralization, significant concentration occurs due to information asymmetry
- Traditional financial institutions may become key intermediaries
- Novel empirical contribution:
  - First causal identification of private information's impact on market power
  - Using cyber attacks and market crises as instruments
- A repeated bargaining model that illustrates how information asymmetry affects profit sharing between intermediaries

Origins of DeFi Intermediation

- *emergence of block builders:* need for privacy by arbitrageurs
  - being frontrun or having the arbitrage opportunity stolen
  - commodification of AI-driven frontrunning bots: Robinson and Konstantopoulos (2020)
- *separation of builders & proposers:*
  - OFAC-sanctioned transactions
  - builder sees transactions, proposer *only* the bid

back



Market Concentration

Number of Builders	Number of Proposers
109	189,126

Builder Name	% Share of Total Builder Revenue	% Share of Blocks
Titan Builder	37.59	17.81
beaverbuild.org	31.74	32.95
rsync-builder.xyz	7.19	16.12

Proposer Name	% Share of Total Proposer Revenue	% Share of Blocks
lido	29.55	27.66
coinbase	10.48	10.41
binance	4.41	4.52
kraken	4.04	4

back

MEV: Maximum Extractable Value

- revenue from the ordering of transactions in a block, in excess of revenue from the value of transactions alone
- transaction fees paid to the block builders + direct payments sent by arbitrageurs to builders (to include transaction in the block)
- how to recognize an MEV block
  - last transaction of the block is issued from builder to proposer

back

Public & Private Trans.: Alternative Defn

- private $\equiv$ non-atomic
  - not on Mempool
  - involve a direct transfer
  - appear to be a one-way swap (public leg of an arbitrage)
  - involve a CEX traded token
- public $\equiv$ atomic
  - the rest!
- differences
  1. shorter sample: September 15, 2022 - January 31, 2024
  2. $(Private_t, Rev_t)$ instead of $(Private_t, Public_t)$

back

Environment

- Bargaining game between *proposer* and *block builders*
- Ignore the rest of the intermediation chain

Environment

- Bargaining game between *proposer* and *block builders*
- Ignore the rest of the intermediation chain
- Infinitely repeated game $t = 0, 1, \ldots$
- M block builders, $m \in \{1, ..., M\}$
- N proposers, $n \in \{1, ..., N\}$, each with (constant) stake w_n
- PoS consensus mechanism
  - Chooses proposer n with probability $\psi_n = \frac{w_n}{\sum_{i=1}^{N} w_i}$
  - n adds block B_t to blockchain

Transactions at Time t

- Blocks made up of three type of transactions
  1. *Regular transactions:* low payoff ($\underline{R}$), happen every period
  2. *Arbitrage transactions:* high payoff ($\bar{R}$), but happen rarely
     - 2-1 *Public arbitrage:* builders w/ symmetric info
     - 2-2 *Private arbitrage:* builders w/ asymmetric info
- Most periods: normal times with no arbitrage opportunities
- Arbitrage transaction
  - Happens with small Poisson rate $\rho \ll 1$ per period
  - Public/private with i.i.d. probability $\pi_u/1 - p_u$
  - Private arbitrage turns public at Poisson rate $\rho_d \gg \rho$ per period

Block Revenue

- Depends on whether it includes an arbitrage transaction or not block of builder m at time t:

$$\widehat{R}_{m,t} \in \{\underline{\mathrm{R}}, \bar{R}\}$$

- **Regular period** t**:** no arbitrage transaction $\widehat{R}_{m,t} = \underline{\mathrm{R}}, \forall\ m$
- **Abnormal period** t**:** arbitrage transaction
  - Public arbitrage transaction
    - Is broadcast on the network $\Rightarrow$ included in every block built at time t
    - $\widehat{R}_{m,t} = \bar{R}, \forall\ m$
  - Private arbitrage transaction
    - Only known by a single builder $b^* \Rightarrow$ only included in b^* block

$$\widehat{R}_{m,t} = \begin{cases} \bar{R} & m = b^* \\ \underline{\mathrm{R}} & \forall\ m \neq b^* \end{cases}$$

Bargaining Game

$$\begin{cases} B_t & \text{Block added at time } t \\ R_t & \text{Revenue of } B_t \\ p_t & \text{Proposer randomly selected by PoS at time } t \\ b_t & \text{Block builder chosen by } p_t \text{ at time } t \end{cases}$$

- Proposer p_t and block builder b_t play an alternating-offer bargaining game a la Rubinstein (1982)
- *Trade* and divide the *net* block revenue with
  - **Exogenous** asymmetric bargaining powers $(1-\delta, \delta)$
  - **Endogenous** outside options $\Upsilon_{P,t}$ and $\Upsilon_{B,t}$

Bargaining Game

$$
\begin{aligned}
S_t &= R_t - \Upsilon_{P,t} - \Upsilon_{B,t} \\
\Pi_{B,t} &= \delta S_t + \Upsilon_{B,t} \\
\Pi_{P,t} &= (1-\delta) S_t + \Upsilon_{P,t}
\end{aligned}
$$

- Simplification: $\delta \to 0$

$$
\begin{aligned}
\theta_{B,t} &= \frac{\Upsilon_{B,t}}{R_t} \\
\theta_{P,t} &= \frac{R_t - \Upsilon_{B,t}}{R_t} = 1 - \frac{\Upsilon_{B,t}}{R_t}
\end{aligned}
$$

Result

Proposition (**Information-Driven Market Power**)

More private transactions in a given block:

- *increases the profit share of the block builder,*
- *decreases that of the proposer.*

Higher value of public transactions, controlling for the value of private transactions, has the opposite impact:

- *decreases the profit share of the block builder,*
- *increases that of the proposer.*

Intuition

- A block builder's outside option governed by the information he has but other nlock builders do not
  - his information advantage over other builders
- **Public vs private period**
  - Public period: regular or abnormal (arbitrage)
    - All transactions public
  - Private period: abnormal
    - One of the blocks offered to the proposer has a private high value arbitrage transaction
- Block builder outside option varies with period being public or private

Public Period

- all blocks are perfect substitutes for each other
- having an arbitrage opportunity does **not** give the builder **comparative advantage**
  - every other builder also has it!
- each builder's outside option is zero
- **no private information relative to other builders**
  - **high value of public transactions only increases proposer share**

Private Period

- If private block not chosen, its value remains unexploited.
- It stays in the pool for tomorrow.
  - p_t might or might not be the proposer tomorrow.
  - The private arbitrage might turn public at rate $\rho_d \gg \rho$.
- Expected amount that p_t can get from this block X.

$$
\begin{aligned}
X \leq\ & \underline{\mathrm{R}} \\
& + \psi_n \left[\rho_d \bar{R} + (1-\rho_d) X\right] \\
& + (1-\psi_n)\psi_n \left[\rho_d \times 0 + (1-\rho_d)\left[\rho_d \bar{R} + (1-\rho_d) X\right]\right] \\
& + (1-\psi_n)^2 \psi_n \left[\rho_d \times 0 + (1-\rho_d)\rho_d \times 0 + (1-\rho_d)^2 \left[\rho_d \bar{R} + (1-\rho_d) X\right]\right] + \cdots
\end{aligned}
$$

Private Period: Builder Outside Option

- Long-time proposer payoff:

$$X \leq \bar{X} = \frac{1-(1-\psi_n)(1-\rho_d)}{\rho_d}\underline{R} + \psi_n \bar{R}$$

  - **Increasing in ψ_n:** p_n more likely to be the proposer soon $\Rightarrow$ add this block.
  - **Decreasing in ρ_d:** private arbitrage more likely to turn public before proposer p_t is chosen again $\Rightarrow$ exploited by some other proposer.
    $\rho_d \downarrow \Rightarrow$ private arbitrage more likely to turns public before proposer p_t is chosen again $\Rightarrow$ exploited by some other proposer

$\Rightarrow$ $\frac{1}{\psi_n}$ and ρ_d: **effective discount rate** for proposer p_t

$\Rightarrow$ $\underline{R} \ll \bar{R} \Rightarrow \bar{X} < \bar{R}$

$$\Upsilon_{B,t}^{\text{private}} \geq \bar{R} - \bar{X} > 0$$

Where Theory and Empirics Coincide!

Proposition (**Information-Driven Market Power**)

More private transactions in a given block:

- *increases the profit share of the block builder,*
- *decreases that of the proposer.*

Higher value of public transactions, controlling for the value of private transactions, has the opposite impact:

- *decreases the profit share of the block builder,*
- *increases that of the proposer.*

Comparison: Normal and Pareto Dists

- Normal distribution

$$\text{skewness} = 0$$

$$\text{excess kurtosis} = 0$$

- Pareto distribution

$$\text{Skewness} = \frac{2(1+\alpha)}{\alpha-3}\sqrt{\frac{\alpha-2}{\alpha}} \Rightarrow \alpha = 3.01 \rightarrow 464.49$$

$$\text{excess Kurtosis} = \frac{6(\alpha^3+\alpha^2-6\alpha-2)}{\alpha(\alpha-3)(\alpha-4)} \Rightarrow \alpha = 4.001 \rightarrow 80,973.78$$

back