Statistical Arbitrage

Prediction Market Spatial Arbitrage

Spring 2026

Intro

Statistical Arbitrage Team


Ian May


Davis Lynn


Jackson Sher


Quang Nguyen


Miguel Gutierrez

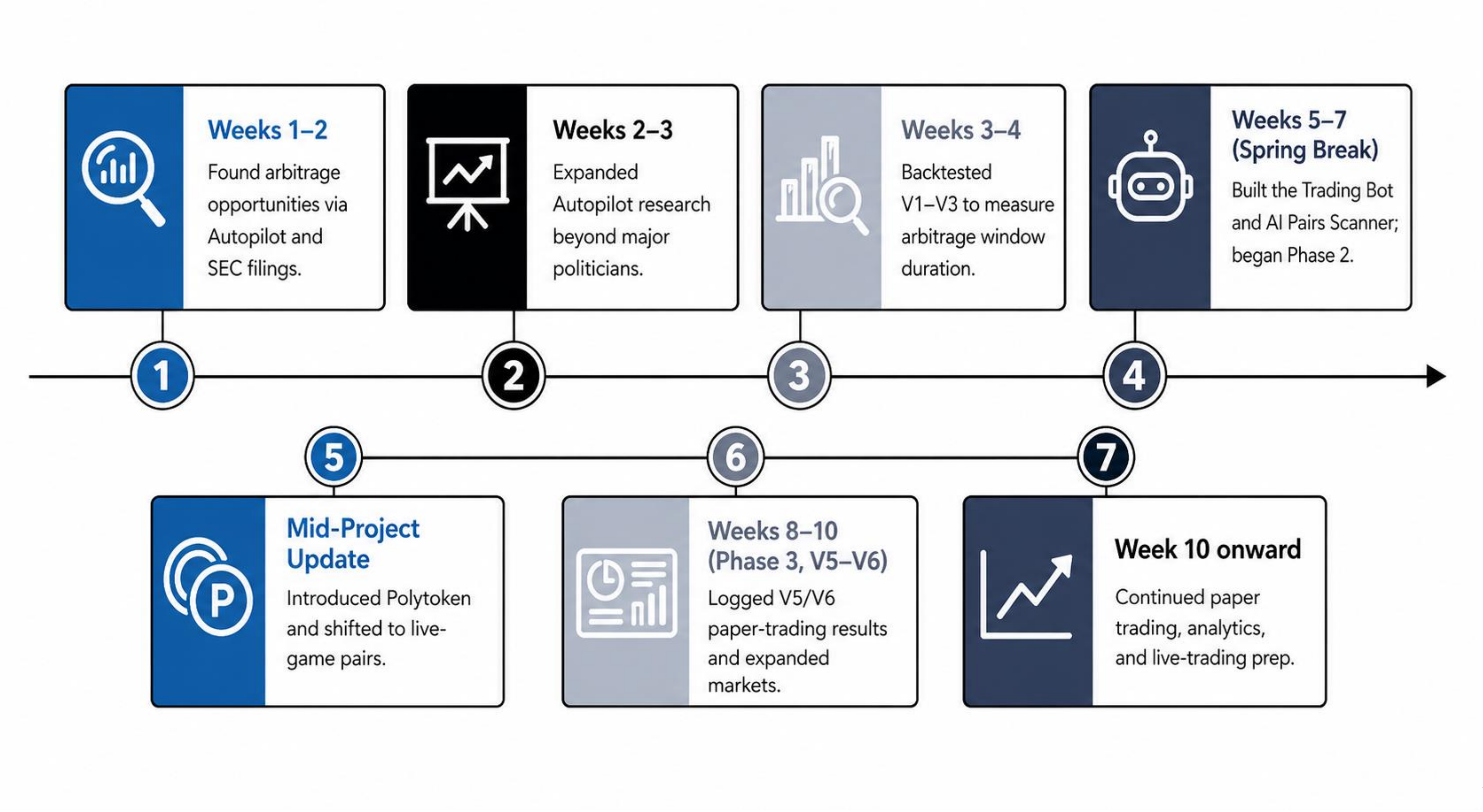

Hayden Kreikemeier

Statistical Arbitrage

Agenda

- **Market Microstructures**
  - What is a Prediction Market
  - Limit Order Book
  - Arbitrage Opportunity
  - Binary vs 3-Outcome Markets
- **Prediction Market Model**
  - Backtests
  - Flowchart
  - Pairs Generator
  - Execution Engine
- **Results & Conclusions**
  - Paper Trading Results
  - Risks & Limitations
  - Next Steps
  - What This Taught us About Markets
- **Q&A**

Timeline


Weeks 1–2
Found arbitrage opportunities via Autopilot and SEC filings.
Weeks 2–3
Expanded Autopilot research beyond major politicians.
Weeks 3–4
Backtested V1–V3 to measure arbitrage window duration.
Weeks 5–7 (Spring Break)
Built the Trading Bot and AI Pairs Scanner; began Phase 2.
1
2
3
4
5
6
7
Mid-Project Update
Introduced Polytoken and shifted to live-game pairs.
Weeks 8–10 (Phase 3, V5–V6)
Logged V5/V6 paper-trading results and expanded markets.
Week 10 onward
Continued paper trading, analytics, and live-trading prep.
P

What is a Prediction Market?

Trade contracts on real-world events

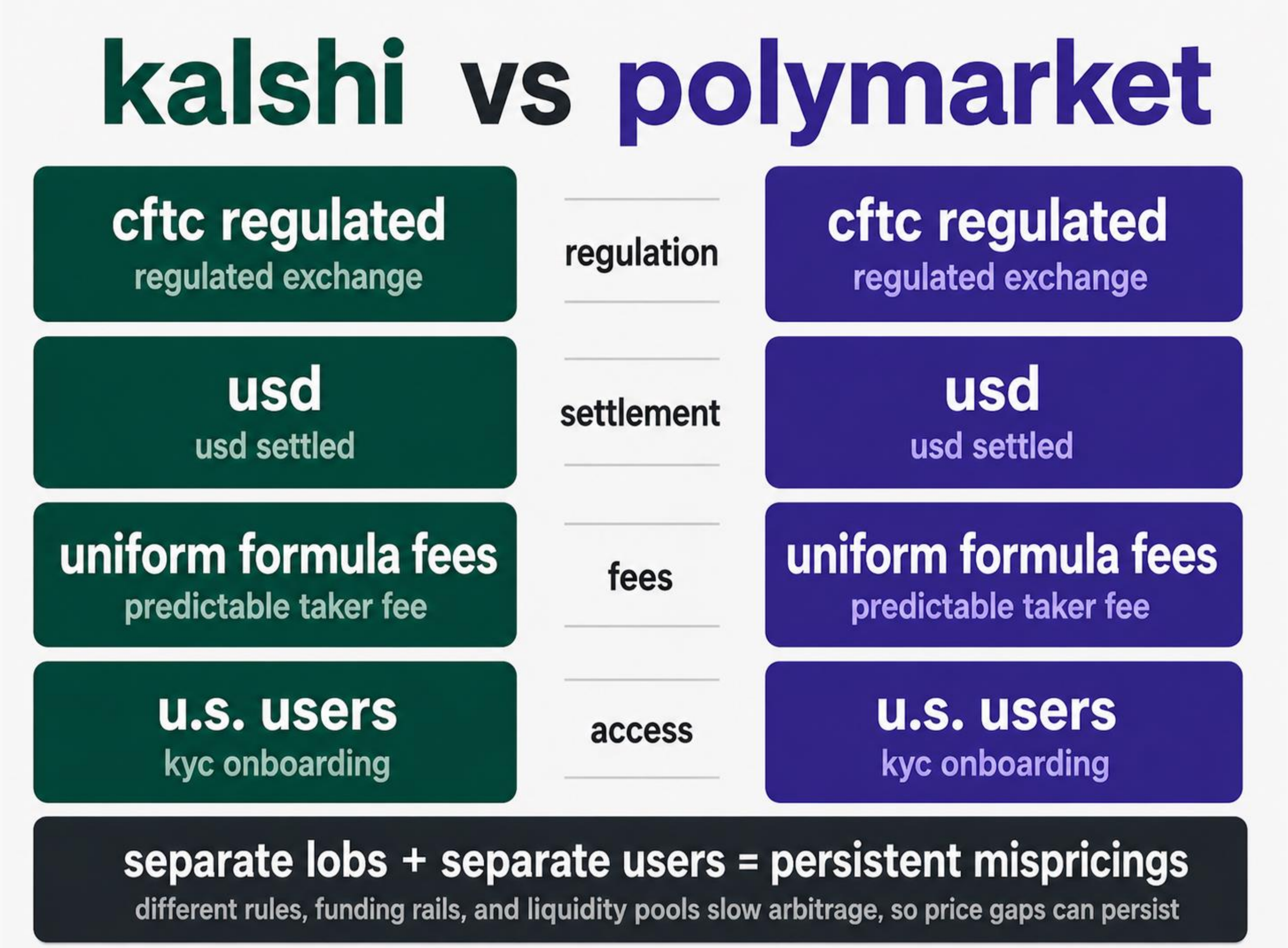
- Will the Stars or Wild win?
- Will the Fed cut rates?

The price IS the probability

- A YES contract pays $1.00 if the event happens, $0 if it does not happen
- A contract trading at 65¢ implies a 65% chance of happening

Two major exchanges in industry

- Kalshi & Polymarket



Limit Order Book

Same market-structure as equities

- Bids, asks, spread, and visible depth
- Two venues, two separate order books

We always lift the offer

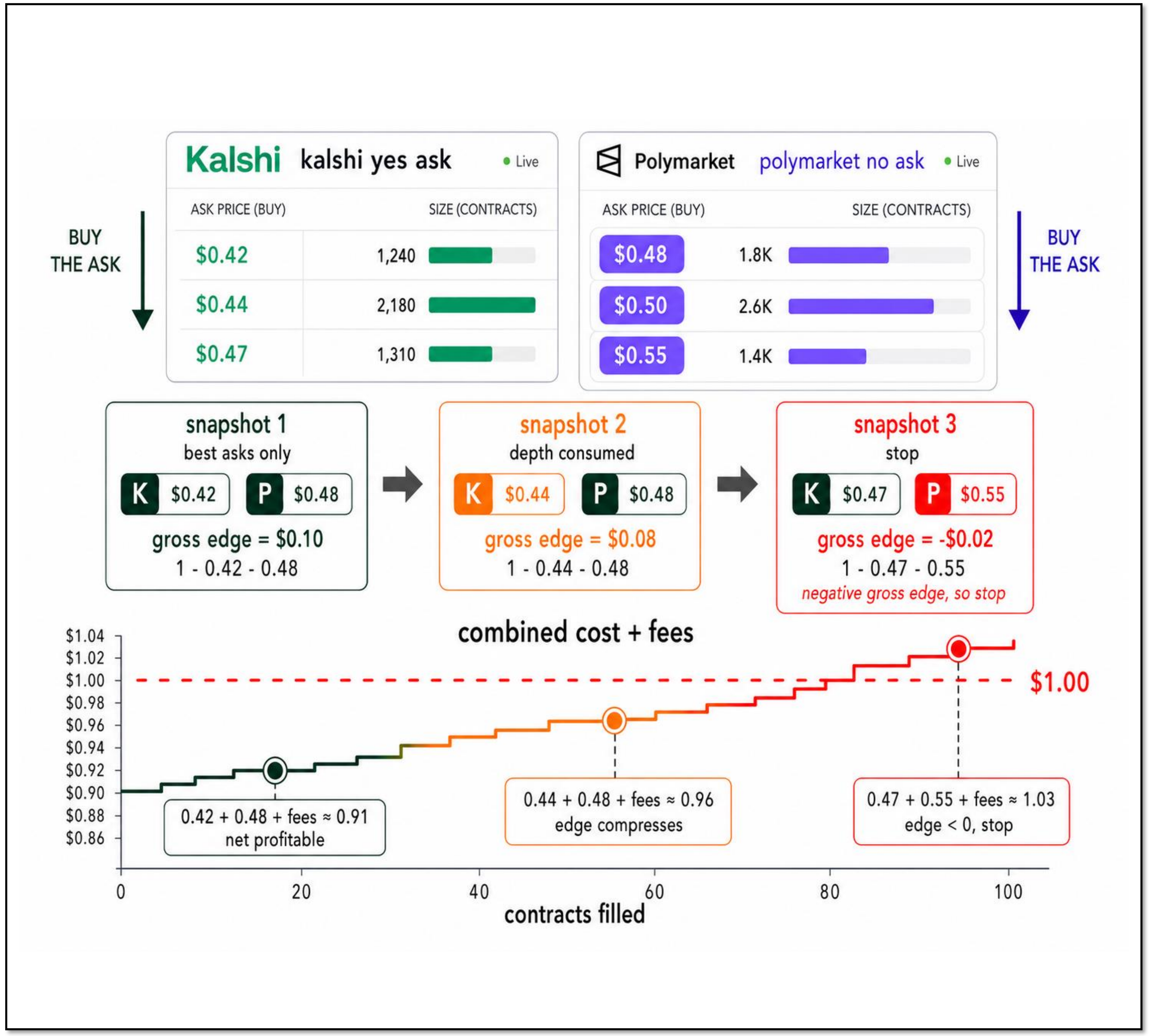
- Quoted best price is only starting point

Same market-structure as equities

- Size at **level 1** (best offer) is *limited*

Same market-structure as equities

- Recompute blended cost after each step
- Stop when next contract takes cost to $1

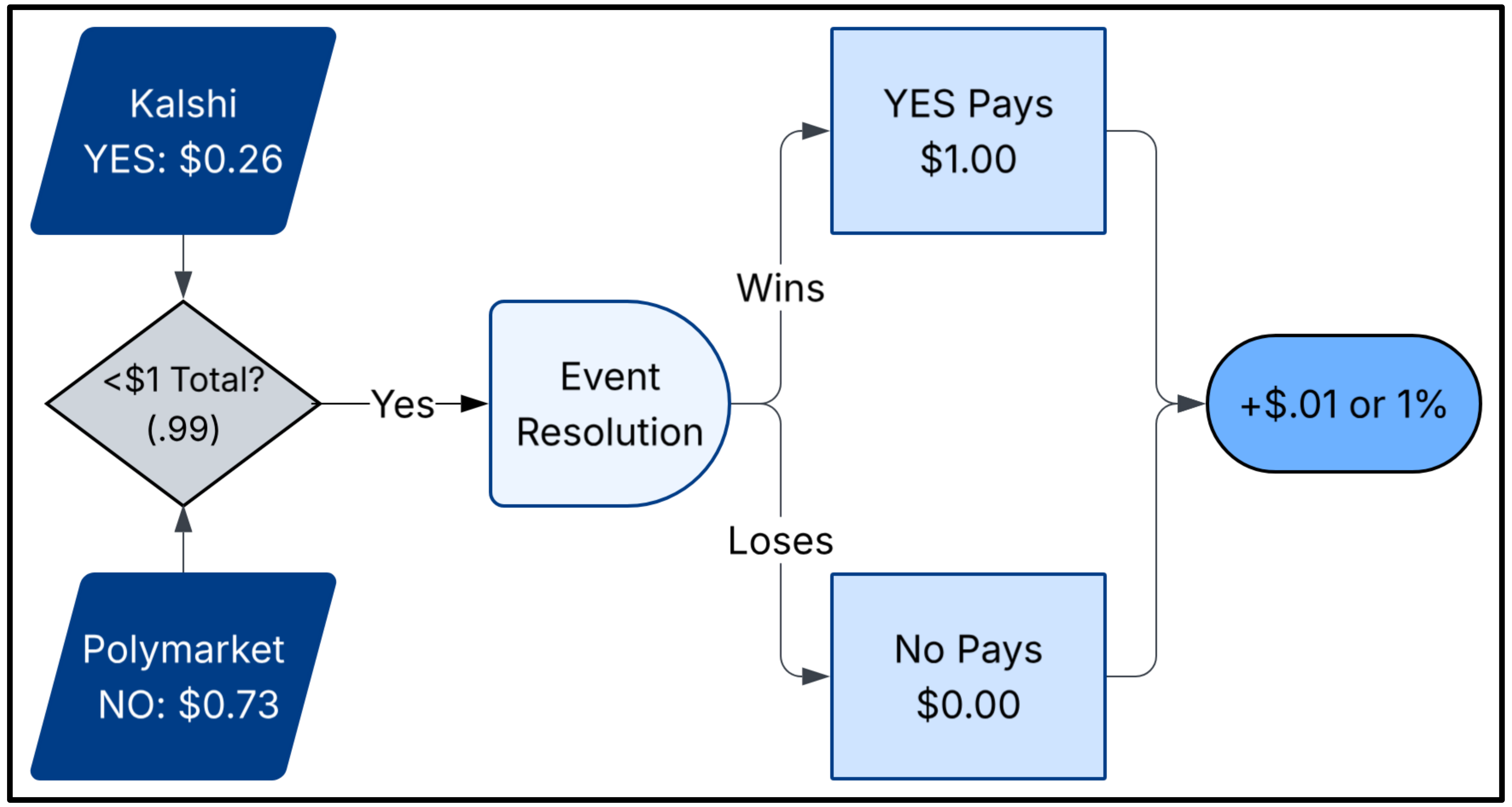


Arbitrage Opportunity

Core Arb Condition

- Arbitrage exists when the Yes on platform A + No on Platform B + fees <$1.00
  - The winning contract is guaranteed to pay $1

Example trade – Will the Spurs win the Western Conference Championship?



Why do arbs happen?

- Separate LOBs
- Asynchronous noise
- Large bets on one platform

Fees are probability-dependent

- Kalshi uses 7% x P(1-P)
- Polymarket uses θ x P(1-P) with category specific θ
- Sports fees are 3%
- Fees are lowest on extreme price, highest at 50¢

Traders@SMU

Binary vs 3-Outcome Markets

#	KALSHI	POLYMARKET
1	Buy T1 YES	Buy T2_YES
2	Buy T1 NO	Buy T1_YES
3	Buy T2 YES	Buy T1_YES
4	Buy T2 NO	Buy T2_YES

Group A 6 structures · binary arb, one outcome at a time · target $1

#	KALSHI	POLYMARKET
1	Buy T1 YES	Buy T1 NO
2	Buy T1 NO	Buy T1 YES
3	Buy T2 YES	Buy T2 NO
4	Buy T2 NO	Buy T2 YES
5	Buy Tie YES	Buy Tie NO
6	Buy Tie NO	Buy Tie YES

Group B · 6 structures · 3-leg YES · target $1

Buy YES on all 3 outcomes — exactly one pays $1. Arb if total cost < $1.

#	KALSHI	POLYMARKET
7	Buy Tie YES	T1 YES + T2 YES
8	T1 YES + T2 YES	Buy Tie YES
9	Buy T2 YES	T1 YES + Tie YES
10	T1 YES + Tie YES	Buy T2 YES
11	Buy T1 YES	T2 YES + Tie YES
12	T2 YES + Tie YES	Buy T1 YES

Group C · 6 structures · 3-leg NO · target $2

Buy NO on all 3 outcomes — the two losing NOs each pay $1. Arb if total cost < $2.

#	KALSHI	POLYMARKET
13	Buy Tie NO	T1 NO + T2 NO
14	T1 NO + T2 NO	Buy Tie NO
15	Buy T2 NO	T1 NO + Tie NO
16	T1 NO + Tie NO	Buy T2 NO
17	Buy T1 NO	T2 NO + Tie NO
18	T2 NO + Tie NO	Buy T1 NO

Groups B and C only exist in 3-outcome markets. Group A structures also apply here, but B and C are entirely new arb types that don't exist in binary matches.

Validating the Model through Backtesting

The Data

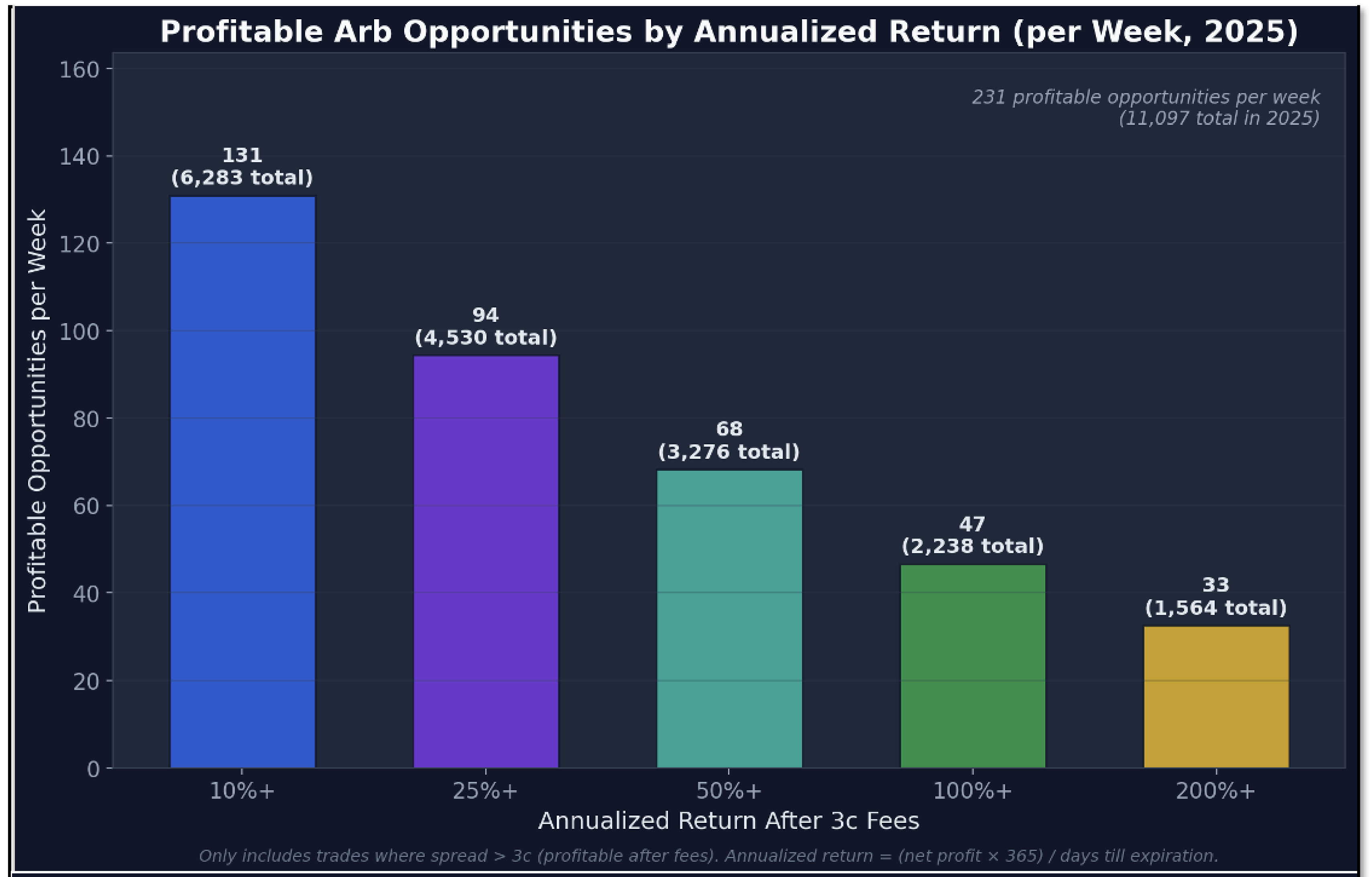
- 477M trades from 2024 to 2026
- 820 market pairs matched by title

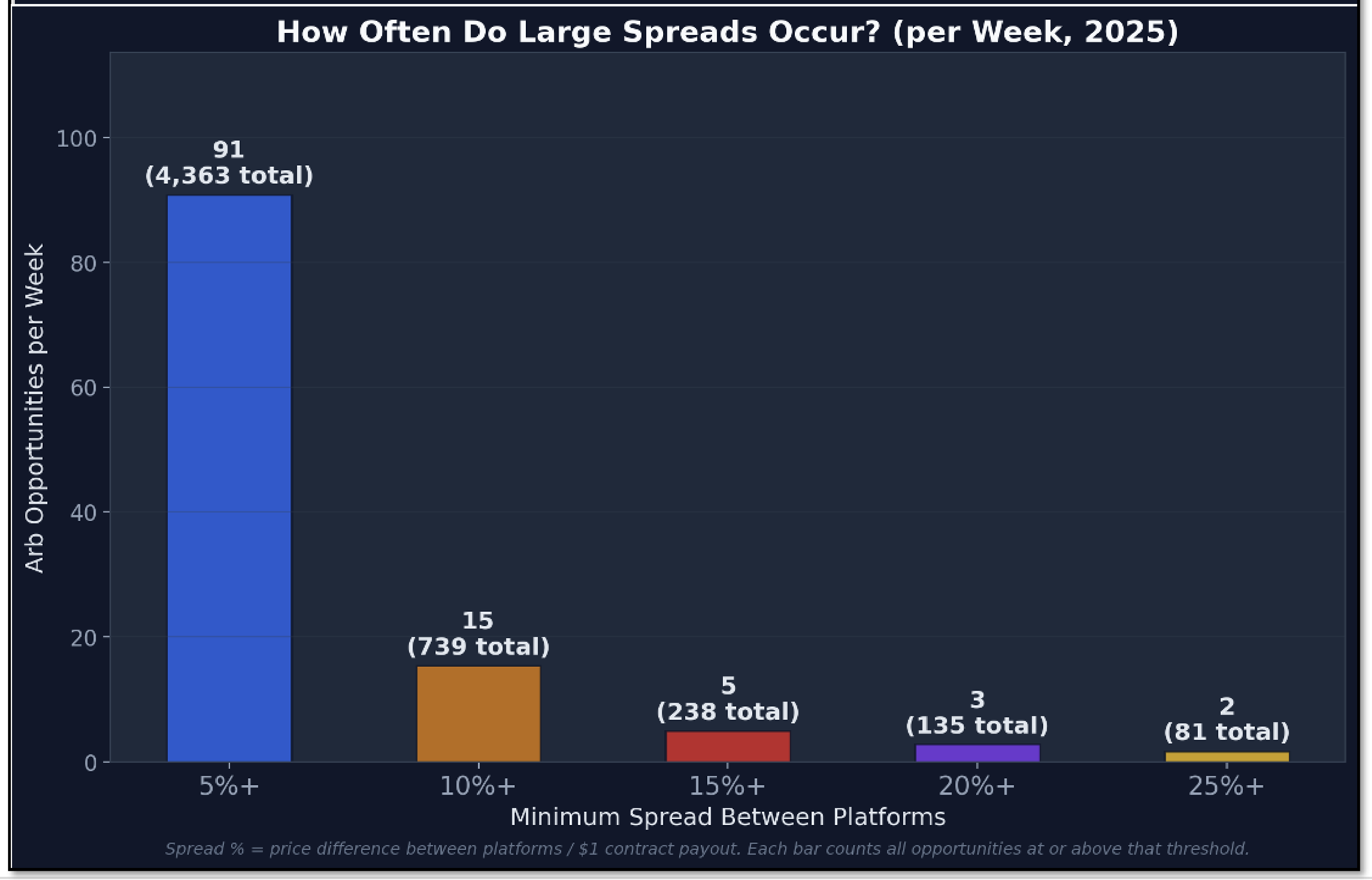
How it Works

- Trades aggregated into 30-seec buckets to create price snapshots across both platforms
- Arb-window flagged when price gap is >= 2 cents (fees)
- Each arb tagged with direction: buy YES on Kalshi and NO on Polymarket, or vice versa

Other Features

- Liquidity filter: Both sides must have at least 100 contracts worth of depth
- ARR graph and large spread graph





Model Transition

Model V5 (Full Market and Convergence)

Pros:

- Broader array of markets
- Less efficient markets (more arbitrage)

Cons:

- Capital locked up for weeks to months
- Bid/Ask spread causes rare convergence
- Low ARR on non-converging pairs
- More inconsistent pairs
- Higher total fees: exit fees

Model V6 (Sports and Contract Expiry)

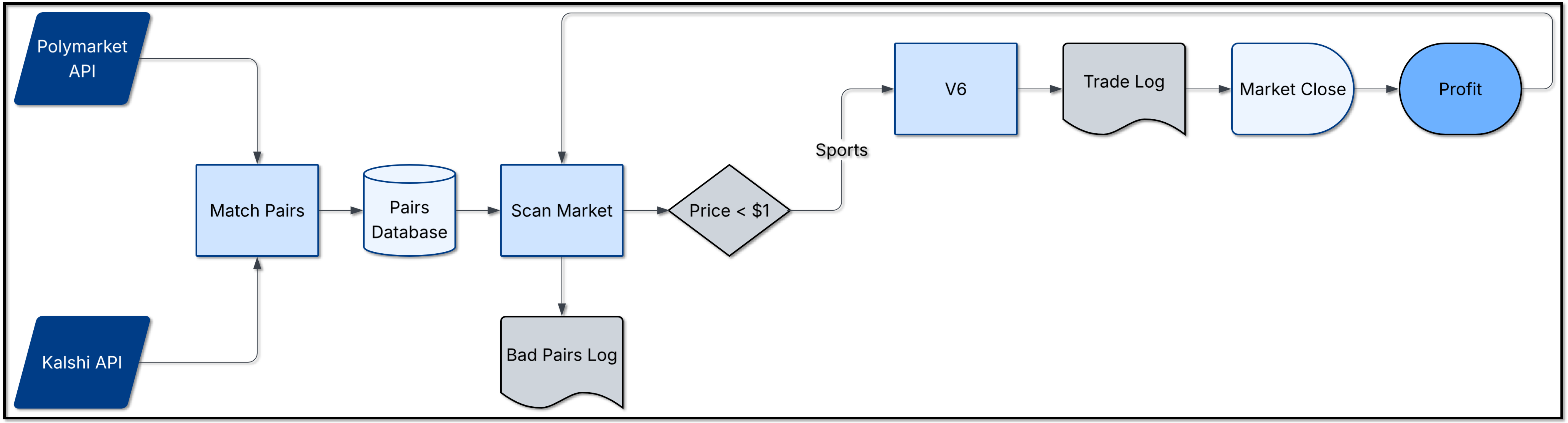
Pros:

- Daily capital reallocation
- Definitive results on contracts
- Faster fetch and calculation time
- Easier pairing, greater standardization
- More liquid markets, faster moving
- No exit fees

Cons:

- Incomplete arbitrage capture
- Faster moving markets

Flowchart


Polymarket API
Kalshi API
Match Pairs
Pairs Database
Scan Market
Bad Pairs Log
Price < $1
Sports
V6
Trade Log
Market Close
Profit

Pairs Generator

Algorithm Overview

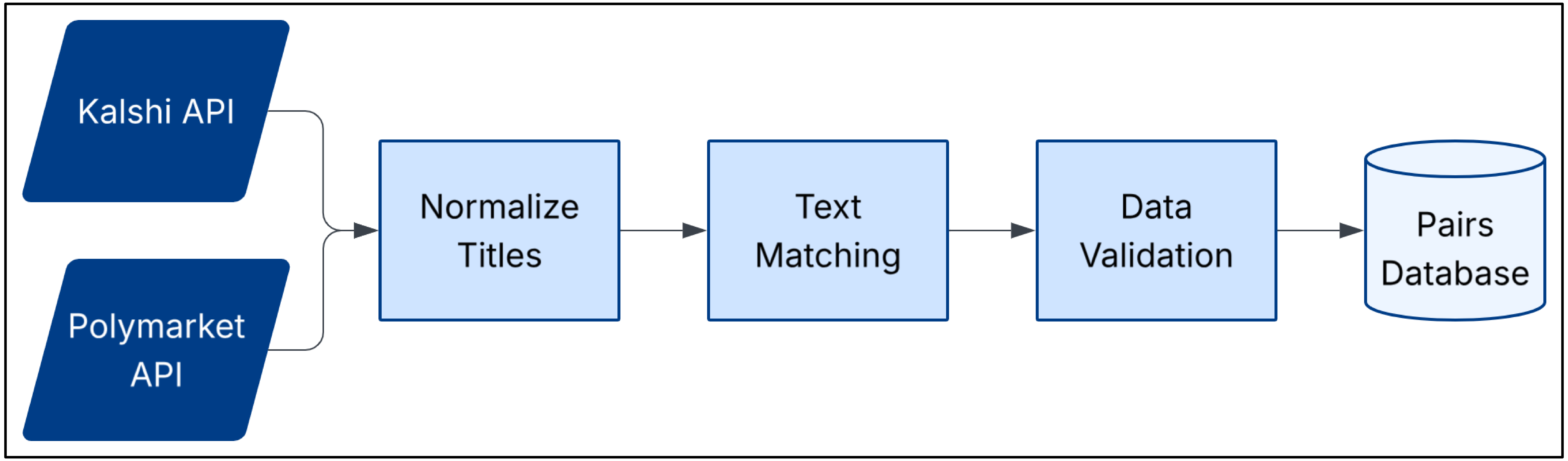
- Data pipeline for inputs for trading model
- Ingests, Normalizes, Matches, and Validates Pairs

Accurate and Configurable

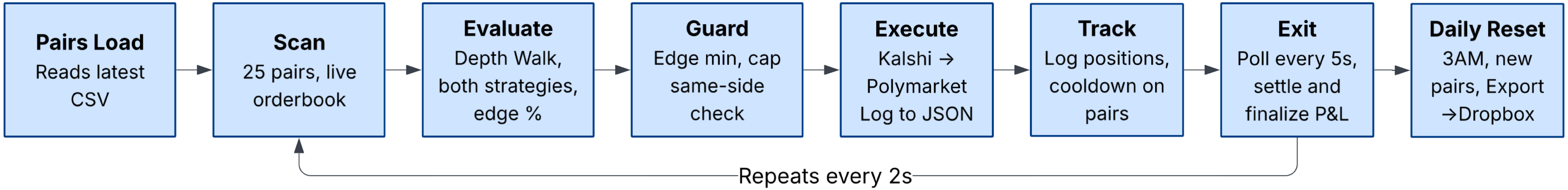
- 98.2% accuracy
- Configurable by category, predominantly sports.

Big Data

- ~650,000 Kalshi, ~50,000 Polymarket Markets
- Ingests and processes ~2.5gb per execution



Automated Trading Cycle



Paper Trading vs Funded Trading

- Reentry on open positions
- Bad execution – hold -> -3% vs -10%
- Speed Limitation

Guard Filters

- Min edge % = 0.1%
- Max edge % = 20%
- Min contracts available = 100 contracts
- Max contracts bought per trade = 1000 contracts

Paper Trading Results

Paper Trading Results 4/16-4/28 (Week Result)

```
==================================================================================
========
              OFFICIAL PERFORMANCE SUMMARY: SYSTEMATIC MULTI-EXCHANGE ARBITRAGE
==================================================================================
========
  Reporting Period:          April 16, 2026 - April 28, 2026
  Total Realized Profit:     $6,855.93
  Cumulative Throughput:     $162,726.12 (Total Deployed Volume)
  Maximum Capital Exposure:  $26,712.90 (Single-Day Peak)
  System-Wide ARR:           720.6%
  Aggregate Strategy ROI:    25.67% (on max exposure)
----------------------------------------------------------------------------------
```

Breakdown of Information

- **Daily Performance:** Average daily gain of **$571.33**.
- **Capital Redeployed (ARR):** Games resolve in the same day, capital can be redeployed
  - Limited arbitrage in the markets, can only invest so much (about $30,000).
- **Scalability: 446 unique positions** executed, generating **$6,856** in the estimated profit.

Arbitrage Window Timing 4/16-4/27

STRATEGY EFFICIENCY: EDGE % VS. ARB PERSISTENCE (TOTAL PROFIT)

Edge Bucket	<15s	15s-1m	1m-5m	5m-30m	30m+	Total
<1%	$97	$14	$12	$18	$79	$220
1-2%	$139	$17	$23	$81	$277	$538
2-5%	$284	$25	$127	$185	$1,004	$1,625
5-10%	$426	$0	$107	$327	$1,164	$2,024
10%+	--	$11	$196	$1,100	$1,143	$2,449

- **Profit Concentration:** 34% of gains driven by high-conviction **10%+ arbs**.
- **Capture Stability:** 60% of total profit derived from windows open **>5 minutes**.
- **Premium Edge:** Consistent **7.68% weighted average edge** maintained across all market sectors.

Sport by Sport breakdown 4/16-4/27

MARKET SECTOR PERFORMANCE SUMMARY

Tennis	Positions: 107	Net Profit: $2,902.26	ROI: 5.4%
MLB	Positions: 99	Net Profit: $1,328.35	ROI: 3.6%
Niche/Other	Positions: 125	Net Profit: $1,213.03	ROI: 3.2%
Cricket	Positions: 7	Net Profit: $402.29	ROI: 7.2%
Soccer	Positions: 8	Net Profit: $349.05	ROI: 5.3%
NBA	Positions: 20	Net Profit: $312.51	ROI: 3.6%
KBO	Positions: 56	Net Profit: $134.91	ROI: 2.8%
Esports	Positions: 19	Net Profit: $113.93	ROI: 2.3%
NHL	Positions: 5	Net Profit: $99.59	ROI: 3.2%

- **Tennis Alpha: $2,912 profit | 5.2% ROI** (Highest net contributor)
- **MLB Liquidity: $37,947 volume | 3.4% ROI** (Primary capital engine)
- **Cricket Yield: $402 profit | 6.0% ROI** (Most efficient sector)
- **NBA Velocity: $291 profit | 4.3% ROI** (High-turnover liquidity)
- **KBO Baseball** top performer in windows open **>30 minutes**.

Paper Trading Results

13-Day Total Results

Date	Positions	Contracts	Capital Deployed	Est. Profit	Total Edge
04-16	39	18,805	$17,755.33	$723.20	2.89%
04-17	37	20,459	$19,377.62	$699.14	2.64%
04-18	39	11,123	$10,497.75	$390.98	3.72%
04-19	27	10,085	$9,594.13	$310.71	3.24%
04-20	59	29,095	$26,712.90	$1,829.47	6.85%
04-21	57	20,232	$19,192.81	$641.42	3.34%
04-23	33	8,418	$7,990.72	$299.36	2.39%
04-24	31	12,970	$12,238.48	$488.28	2.89%
04-25	33	9,188	$8,609.63	$382.29	2.66%
04-26	23	9,479	$8,964.68	$351.13	2.66%
04-27	22	7,209	$6,952.88	$123.10	1.73%
04-28	46	15,779	$14,839.19	$616.85	3.20%
TOTAL	446	172,842	$162,726.12	$6,855.93	3.18%

The objective of the Paper Trading summary is to break down the trading for the specific day and analyze where our profits are coming from.

Paper Trading Summary consists of:

- Each Trade
- Edge
- Profit
- Window Open
- ARR

```
==========================================================================================
  PAPER TRADING SUMMARY
==========================================================================================
  Period            04-16 Breakdown (Excluding Soccer)
  Positions         39
  Contracts         18,805
  Capital deployed  $17,755.33
  Est. profit       $723.20
  Total edge        2.91%

  Note: each session uses max(total_profit) — Soccer trades filtered out.
==========================================================================================

------------------------------------------------------------------------------------------
  TRADE LOG
------------------------------------------------------------------------------------------
Market                                       Ctrs     Edge     Profit  Window                ARR
------------------------------------------------------------------------------------------
mumbai indians punjab kings                 1,000    8.76%  $   80.54  02:22pm ( 2h 18m)    436%
pirates nationals                           1,000    8.52%  $   78.49  05:42pm ( 2h 25m)    433%
rockies astros                              1,000    8.03%  $   74.35  01:50am ( 1h 26m)    372%
bnk fearx drx                               1,000    7.52%  $   69.98  06:25am ( 1h 34m)   3890%
white sox rays                              1,000    5.25%  $   49.89  06:55pm ( 2h 28m)    269%
ellis kukushkin                               656    6.64%  $   40.88  08:11am (14m 38s)    316%
rocha sakamoto                                936    4.21%  $   37.81  03:23pm ( 1h 18m)    208%
rybakina shnaider                           1,000    3.70%  $   35.68  04:12pm ( 1m 22s)    184%
ducks predators                             1,000    3.61%  $   34.82  01:09am (     0s)   1383%
angels yankees                              1,000    3.39%  $   32.83  05:46pm ( 1h 54m)    170%
crawford kennedy                            1,000    3.18%  $   30.83  06:33pm ( 2h 17m)    162%
ruiz forejtek                               1,000    2.60%  $   25.38  02:28pm (42m 39s)    129%
tigers royals                               1,000    2.48%  $   24.19  07:42pm ( 2h 19m)    126%
salkova maria                                 734    3.18%  $   22.65  10:25am (     0s)    153%
tabur martin                                1,000    1.52%  $   14.98  08:28pm (46m 17s)     90%
jets sharks                                   404    3.71%  $   14.44  12:36am ( 34m 0s)   1388%
safiullin assche                            1,000    1.05%  $   10.35  04:36pm (57m 50s)     60%
gadamauri vithoontien                         248    3.10%  $    7.45  08:10am (49m 16s)    147%
athletics rangers                             310    2.18%  $    6.62  08:35pm (40m 50s)    111%
shin jones                                    211    3.24%  $    6.61  08:10am (49m 31s)    154%
zink krueger                                  490    0.65%  $    3.15  04:50pm (    57s)     32%
shapovalov marozsan                           245    1.30%  $    3.15  03:18pm (     0s)     64%
fernandez sonmez                              560    0.55%  $    3.07  07:09pm ( 2h 21m)     27%
maia chwalinska                               129    2.44%  $    3.07  01:36pm (24m 10s)    119%
padres mariners                               177    1.53%  $    2.67  02:12am (     0s)     70%
baptiste shymanovich                          176    1.09%  $    1.90  12:31pm (  1m 6s)     53%
dignitas disguised                            125    1.53%  $    1.88  08:04pm (     0s)    478%
kt wiz nc dinos                                65    1.87%  $    1.20  09:43am ( 1h 52m)     90%
bondar oliynykova                              70    1.41%  $    0.97  07:38pm (     0s)     71%
osuigwe kraus                                  64    1.34%  $    0.85  10:52am (  1h 3m)     75%
shopify rebellion team liquid                  80    1.05%  $    0.83  07:28pm ( 1m 59s)    322%
shanghai sharks xinjiang flying tigers         50    1.14%  $    0.56  07:41pm (    12s)     50%
doosan bears ssg landers                       10    4.61%  $    0.44  11:18am (     0s)    223%
```

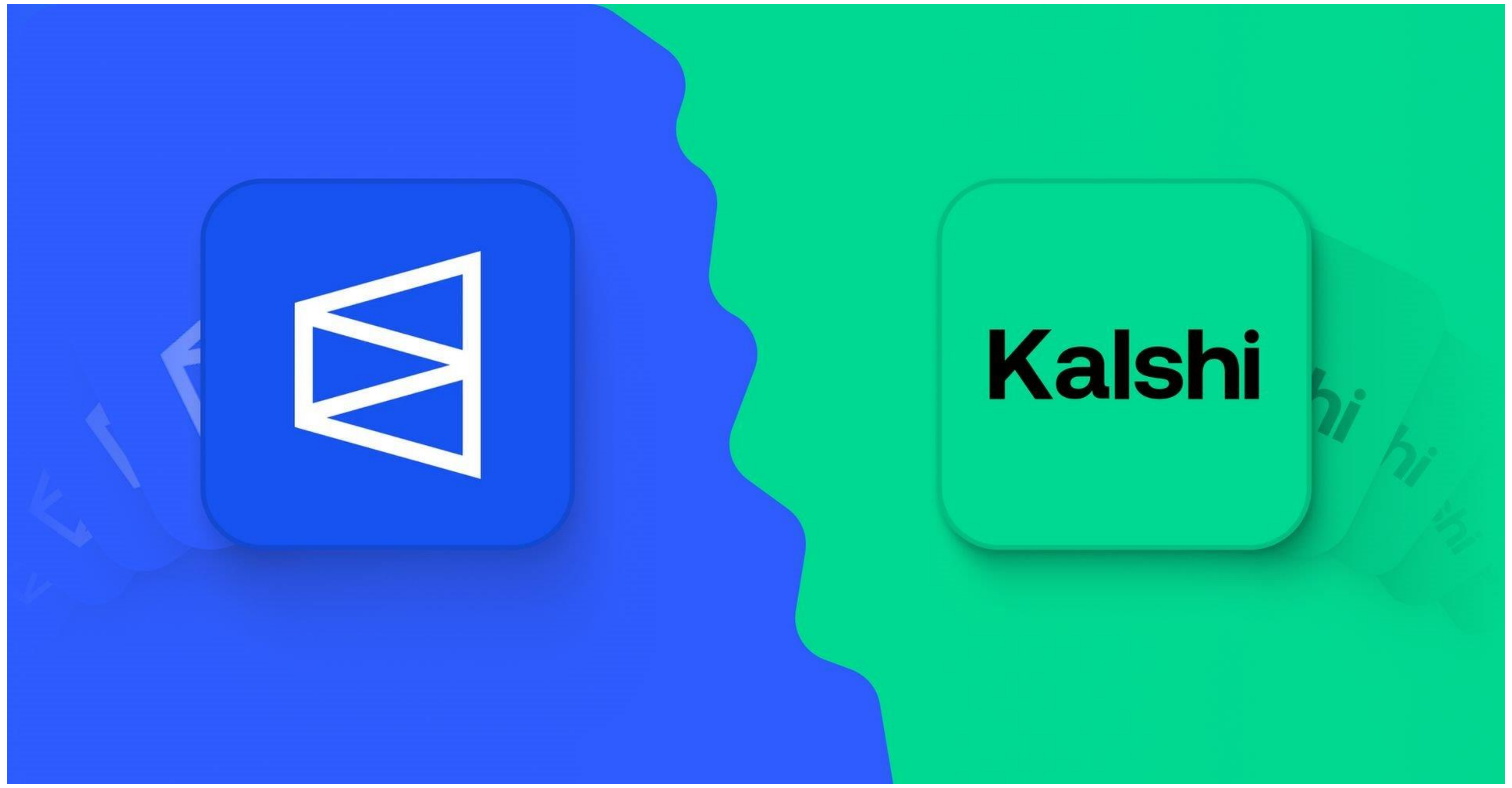
Risks & Limitations

	CATEGORY	FAILURE MODE	RISK	MITIGATION
1	Latency	Window closure	Faster bots take the trade first.	Optimize speed; trade niche markets.
2	Execution	Leg failure	One side fills; the other doesn't.	Fill-or-Kill orders only.
3	Environment	Paper vs. live	Backtests ignore fees and slippage.	Apply a 20–30% haircut.

Next Steps

PHASE 01	PHASE 02	PHASE 03
Short-Term Next 4 Weeks	**Medium-Term** Summer	**Long-Term** Next Semester
01 Go Live Deploy $500–$1k real capital to validate execution against paper-trade results.	**01 Latency** Replace REST polling with WebSocket feeds for sub-second quote updates.	**01 ML Entry Timing** Predict when edges appear, not just where they exist.
02 Expand Pair Universe Scale from ~200 -> 1,000+ verified pairs across politics, crypto, tech.	**02 Alternative Venues** Integrate PredictIt, Manifold, Robinhood, Coinbase to widen arb opportunity surface.	**02 Club Paper Book** Onboard new Traders@SMU members to contribute and test strategies.
03 3-Way Arbitrage Support games with ties (soccer, chess). Opens ~30% more markets.		**03 Cross-Asset** Crypto perpetuals vs. prediction-market BTC/ETH contracts.
Validate ↗	**Accelerate ↗**	**Scale ↗**

What this Taught us about Markets



Arbitrage is real in Inefficient Markets

- Same event, two platforms, different prices. We saw 3–8% edges hold for hours, not seconds.

Fees eat naïve arbitrage

- A 2% raw edge becomes negative after fees

The hard part is knowing it's the same trade

- Matching "Barca FC vs Madrid FC?" on Polymarket to Kalshi's equivalent ticker is a language problem, not a finance one. AI made 40k × 40k comparisons tractable.

Speed isn't always the moat

- HFT shops skip these markets because size is too small. Our edge comes from patience.

Thank you! Questions?

Ian May

Davis Lynn

Jackson Sher

Quang Nguyen

Miguel Gutirrez

Hayden Kreikemeier

Traders@**SMU**