Binary Trigger Signals for Deep Reinforcement Learning in Equity Markets

Juan Manuel Montoya

PhD Candidate — University of Salzburg

Outline

- Introduction
- Financial Basics
- Experimental Setup
- Results
- Interpretation

Section: Introduction

Motivation

Reinforcement Learning for Financial Markets

- Why are financial markets a difficult environment for RL?

Key challenges:

- Non-stationarity
- Partial observability
- Delayed rewards
- Noisy signals
- Regime changes

Section: Financial Basics

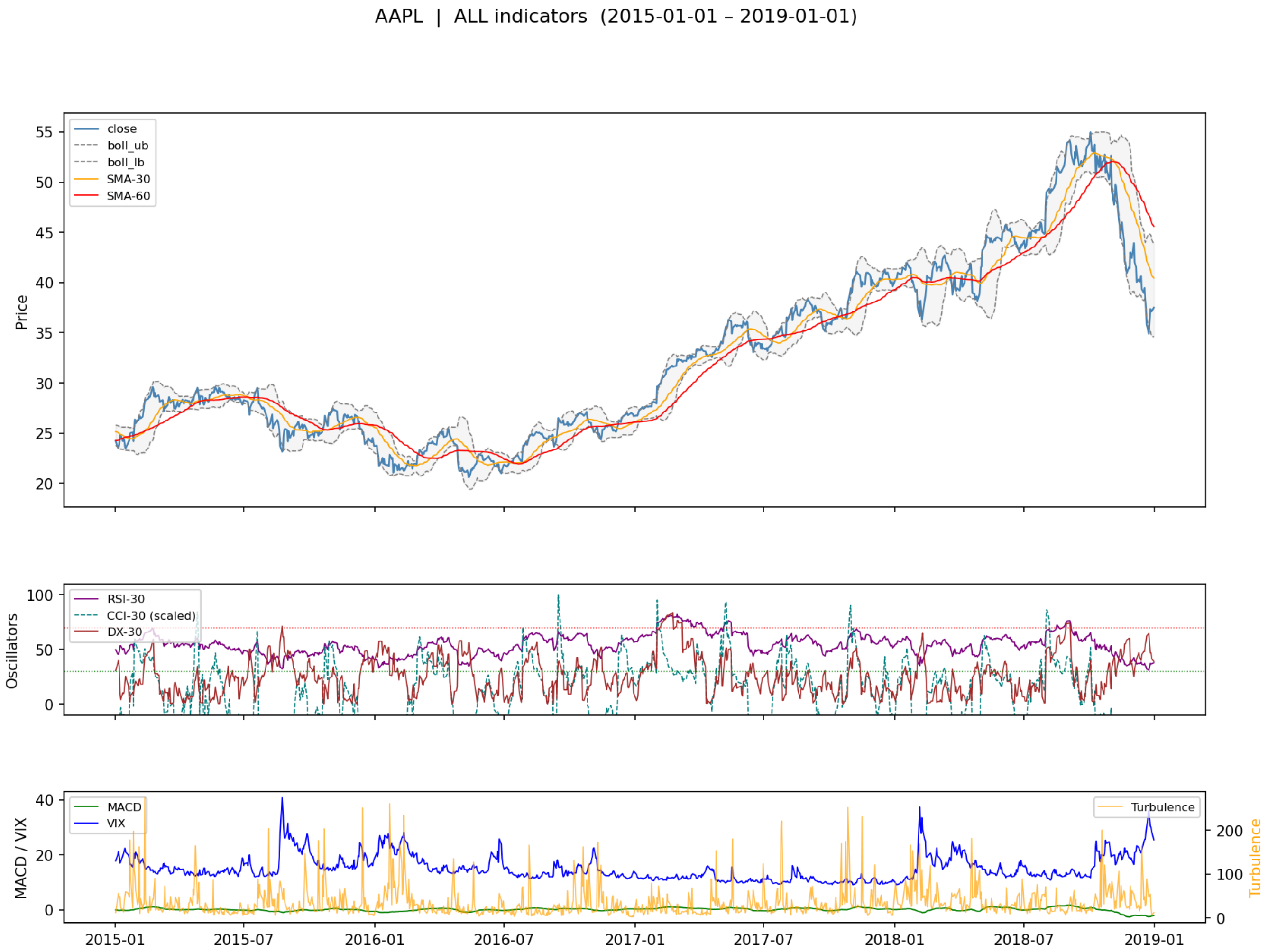
Financial Indicators and Trading Decisions

- Indicators extract patterns from price and volume
- Used to generate **buy** and **sell** signals

Two categories:

- Continuous indicators (RSI, MACD, Bollinger, CCI, DX)
- Event-based indicators (triggers)

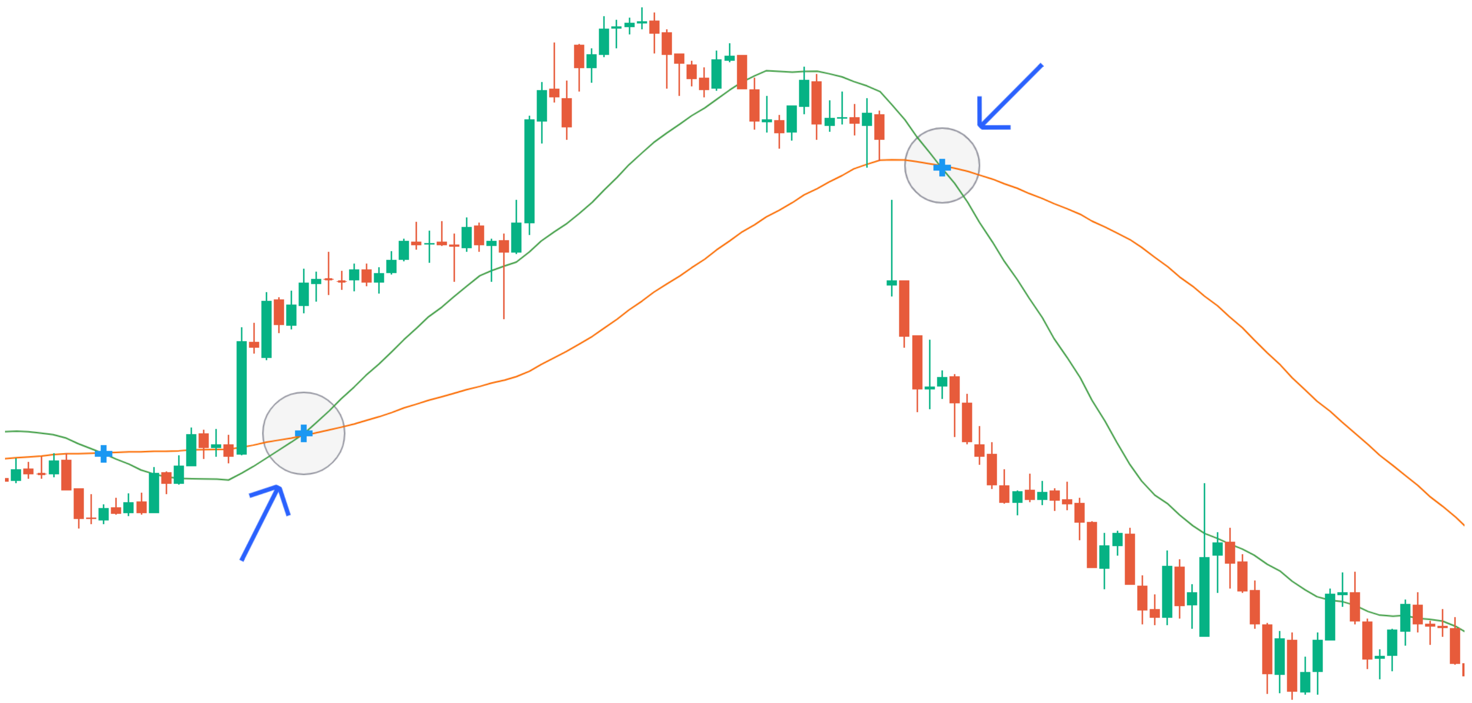
All-in-one: Continuous indicators



Moving Average Crossover with Trigger

- Crossover logic:
  - Buy when fast MA 20 crosses above slow MA 40
  - Sell when fast MA 20 crosses below slow MA 40

Moving Average Crossover with Trigger



Research Questions

RQ	Question
H1	*(Paradigm shift)* Can a single-indicator trigger outperform the all-in-one approach? How? Why?

The **strongest novel claim** is the **event-driven representation** discovery:

- RL algorithm pairs best with event-driven in Equity Markets
- Against general assumption of *all-in-one* approach
- Even-driven well know in CT, Econometrics, and RL

Section: Experimental Setup

Basic Experimental Setup

Data

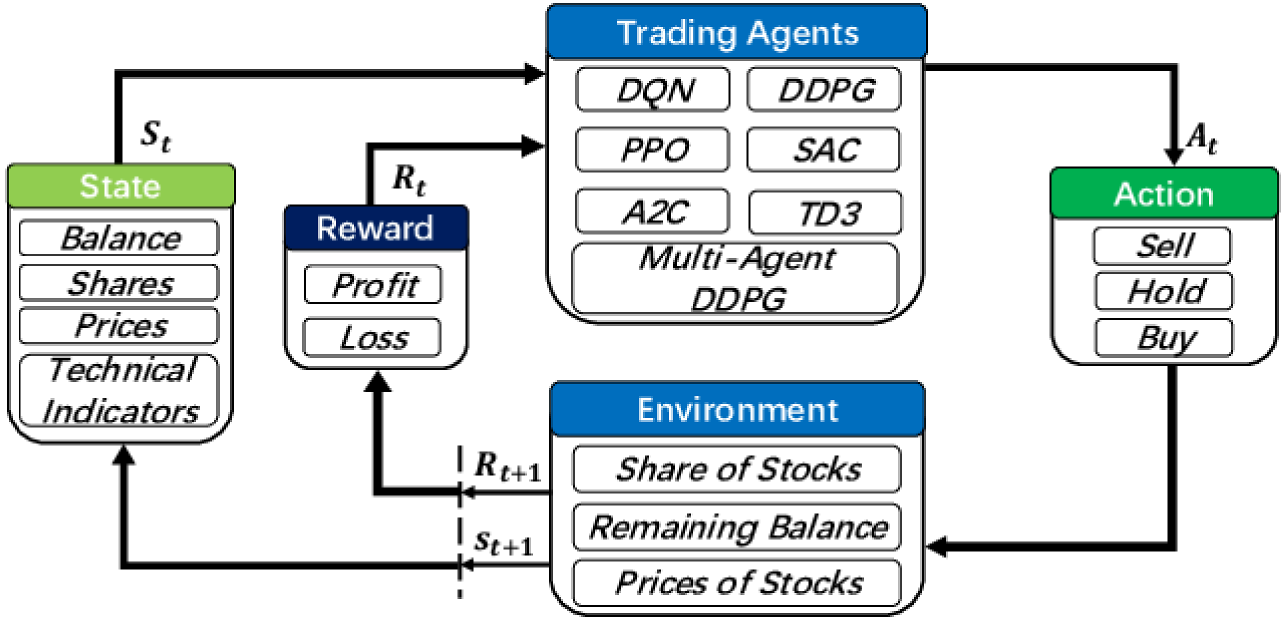
- Dow Jones 30 stocks with daily candles

Window	Period	Regime	Training
W1	2020 Mar-Dec	COVID crash + recovery	2005 Jan – Feb 2020
W2	2022 Jan-Dec	Inflation / bear market	2005 Jan – Nov 2021
W3	2025 Jan-Dec	Recent bull + Trump tariffs	2005 Jan – Nov 2024

Goal: Validate the algorithm-indicator interaction across multiple market regimes

Environment

- Action space: 30 stocks with [-100, 100]
- State space: indicators + prices + **all 30 stocks**
- Reward: portfolio (1M) return
- Alg: A2C, PPO, SAC with 5 rand. seeds



igure 1: Overview of automated trading in FinRL, using

My Experimental Setup

Trigger-Based State Representation

- Binary crossover events
- Breakout detection

Subsampling

- Training only on trigger days

My Experimental Setup

Why is this important?

- Explainability
- Noise reduction
- Novel representation
- Learning efficiency
- Better results

All-in-one state with 6 Stocks

Ticker	Day	Close	Volume	MA20	MA40	RSI	MACD
AMZN	0	191.23	42.1M	188.45	185.67	0.52	+1.45
AAPL	0	229.67	50.2M	227.34	224.56	0.48	+0.87
MSFT	0	425.34	18.9M	422.12	418.90	0.45	-0.92
GOOGL	0	138.67	26.3M	136.45	134.23	0.55	+1.23
INTC	0	44.56	36.7M	43.78	42.34	0.42	-0.67
NVDA	0	870.12	11.8M	865.34	860.12	0.63	+2.12
AMZN	1	194.56	45.2M	189.78	186.23	0.58	+2.34
AAPL	1	232.89	52.1M	229.45	226.12	0.62	+1.87
MSFT	1	428.34	19.7M	423.89	420.12	0.51	-0.45
GOOGL	1	141.56	28.4M	138.92	135.67	0.64	+2.12
INTC	1	45.89	38.9M	44.56	42.89	0.48	-1.23
NVDA	1	876.89	12.3M	869.78	862.45	0.71	+3.45

Moving Average with Trigger

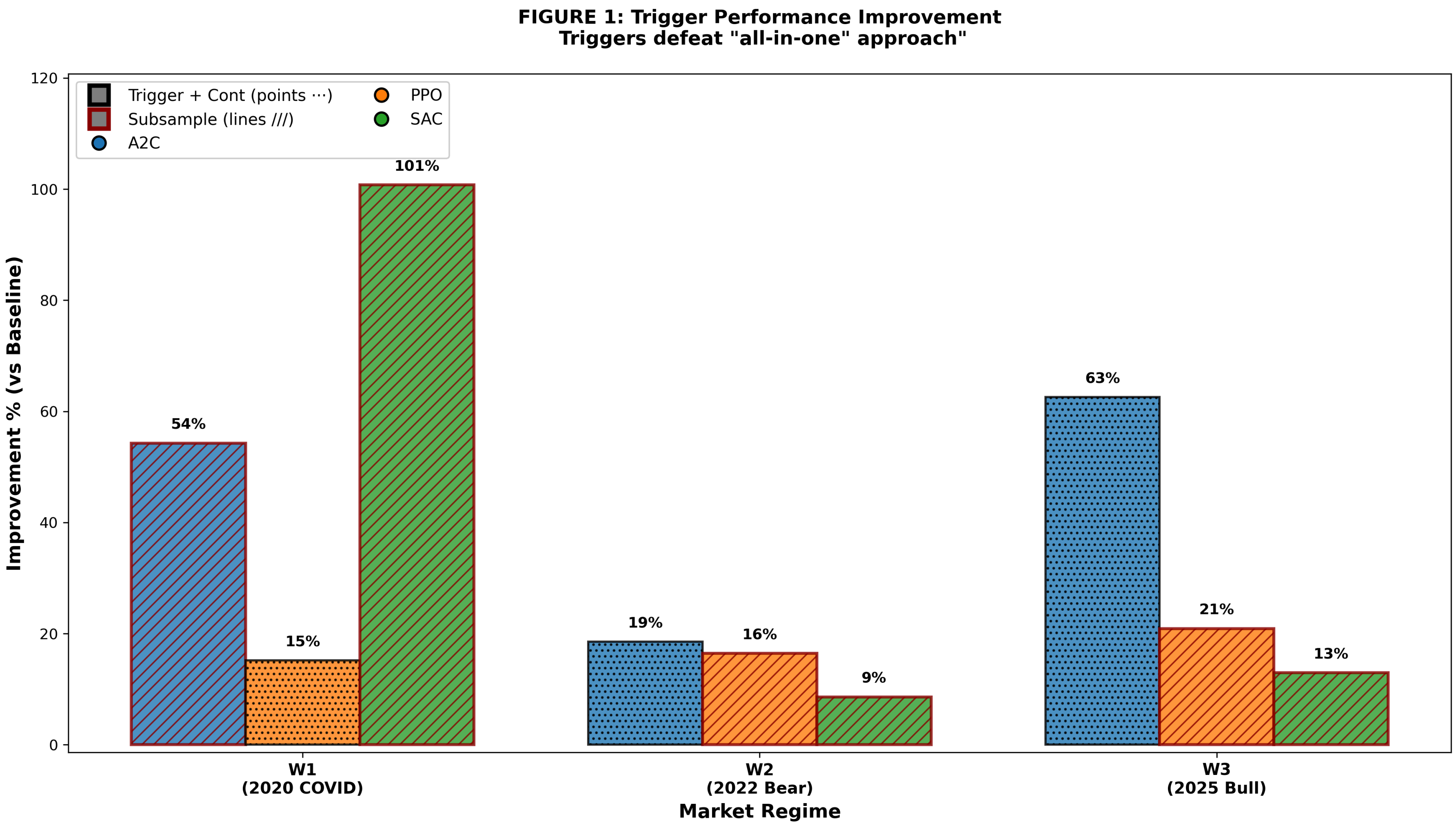
Ticker	Day	Close	Volume	MA20	MA40	MVX
AMZN	0	191.23	42.1M	188.45	185.67	0
AAPL	0	229.67	50.2M	227.34	224.56	1
MSFT	0	425.34	18.9M	422.12	418.90	0
GOOGL	0	138.67	26.3M	136.45	134.23	1
INTC	0	44.56	36.7M	43.78	42.34	1
NVDA	0	870.12	11.8M	865.34	860.12	1
AMZN	1	194.56	45.2M	189.78	186.23	0
AAPL	1	232.89	52.1M	229.45	226.12	0
MSFT	1	428.34	19.7M	423.89	420.12	0
GOOGL	1	141.56	28.4M	138.92	135.67	0
INTC	1	45.89	38.9M	44.56	42.89	0
NVDA	1	876.89	12.3M	869.78	862.45	0

Moving Average with Trigger and Subsample

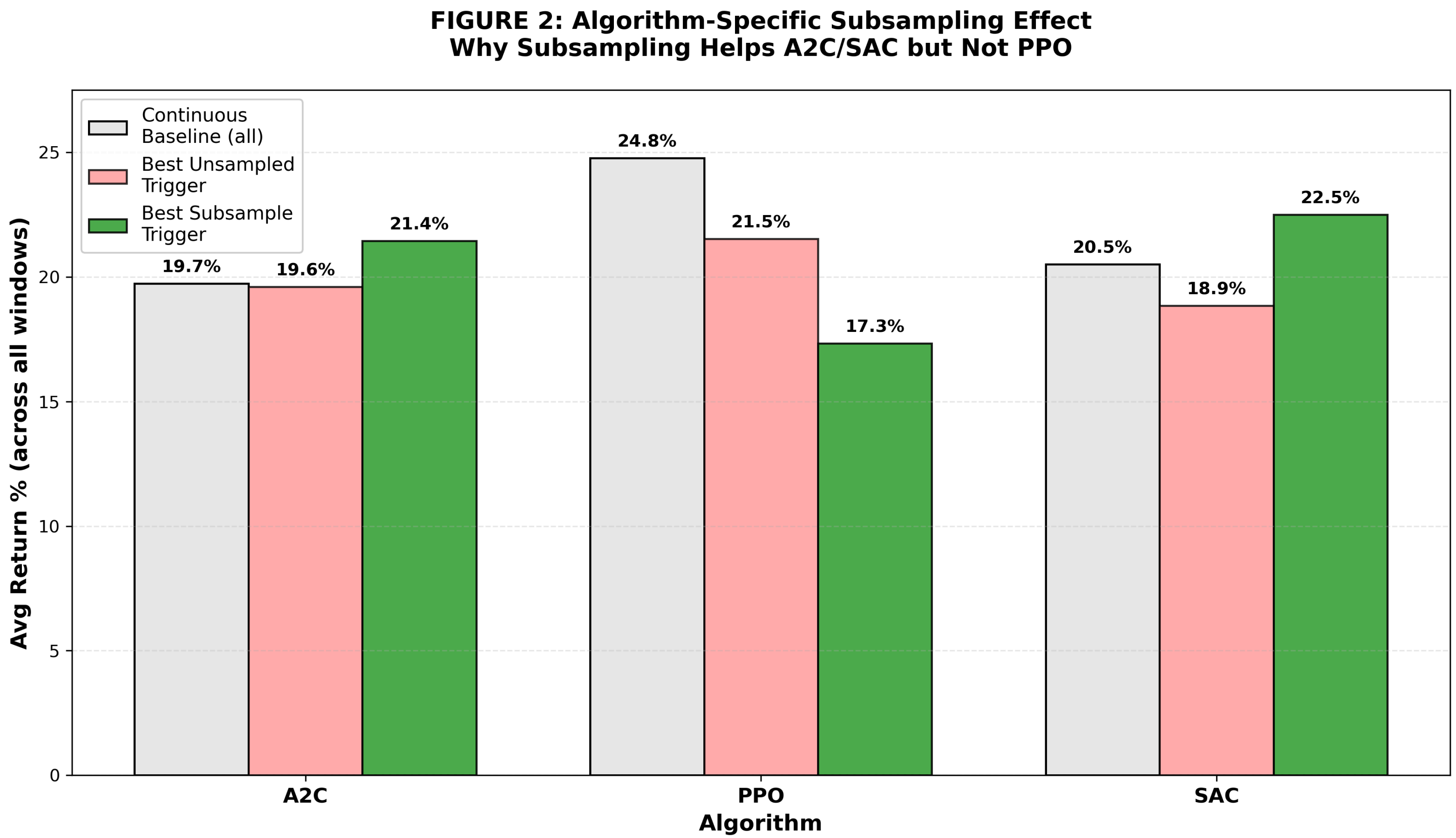
Ticker	Day	Close	Volume	MA20	MA40	MVX
AMZN	O	191.23	42.1M	188.45	185.67	0
AAPL	O	229.67	50.2M	227.34	224.56	1
MSFT	O	425.34	18.9M	422.12	418.90	0
GOOGL	O	138.67	26.3M	136.45	134.23	1
INTC	O	44.56	36.7M	43.78	42.34	1
NVDA	O	870.12	11.8M	865.34	860.12	1
AMZN	P	194.56	45.2M	189.78	186.23	0
AAPL	P	232.89	52.1M	229.45	226.12	0
MSFT	P	428.34	19.7M	423.89	420.12	1
GOOGL	P	141.56	28.4M	138.92	135.67	0
INTC	P	45.89	38.9M	44.56	42.89	0
NVDA	P	876.89	12.3M	869.78	862.45	0

Section: Results

Stable Results though all Market Regimes

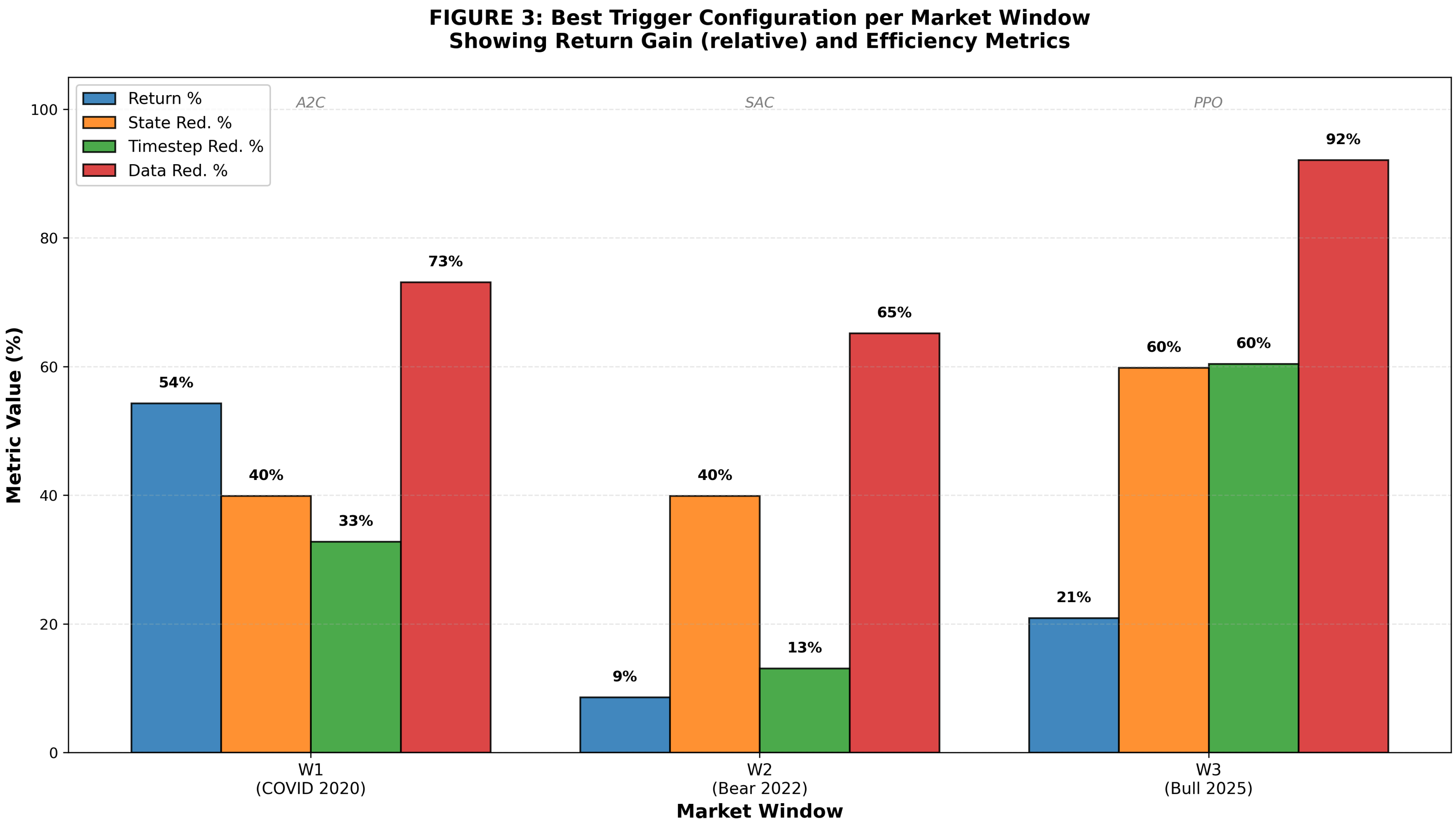


Algorithm Differences



- Buy and Hold Baseline: around 11%!

Subsampling Effect

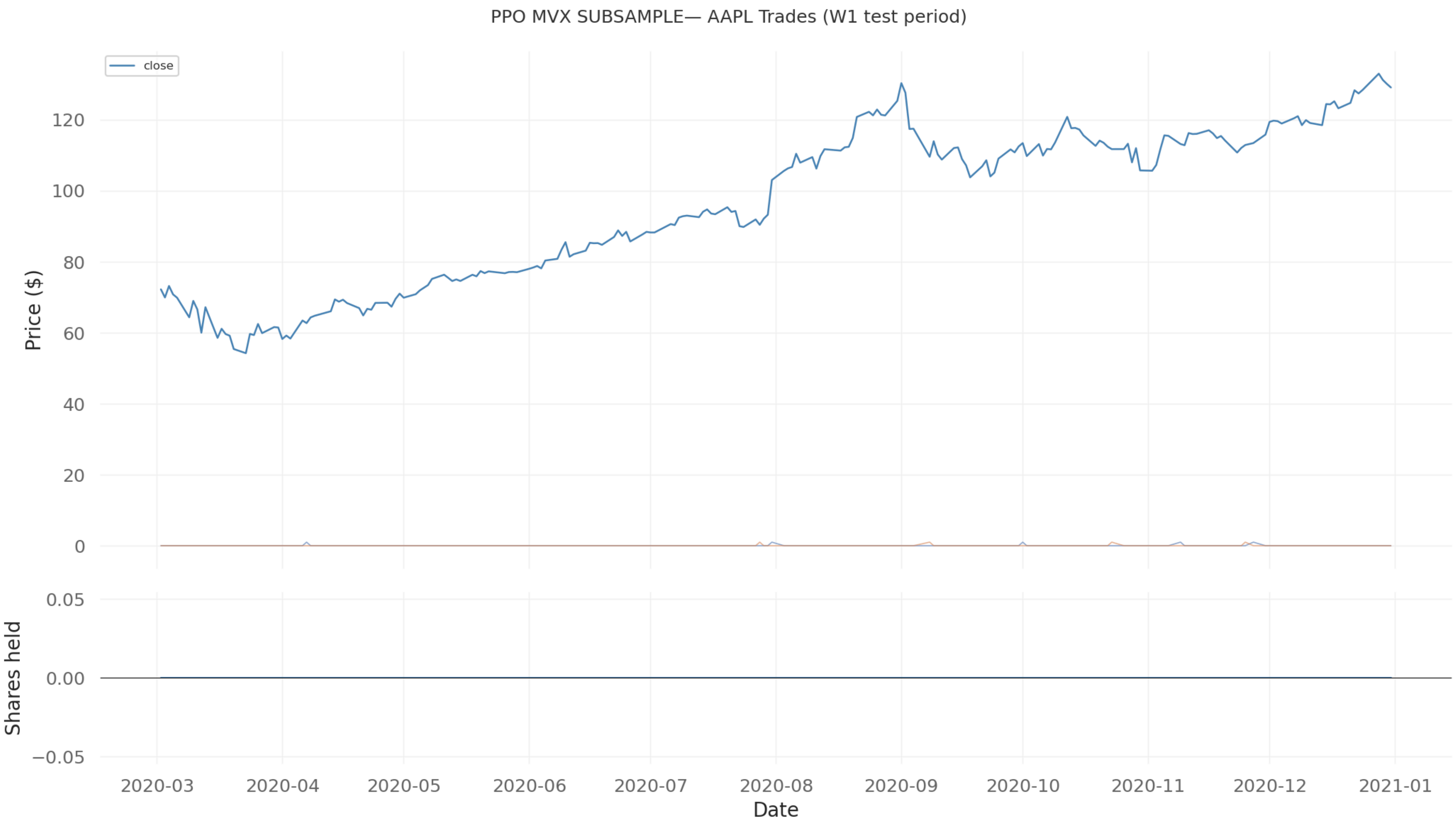


Hypothesis Pre-Results Summary

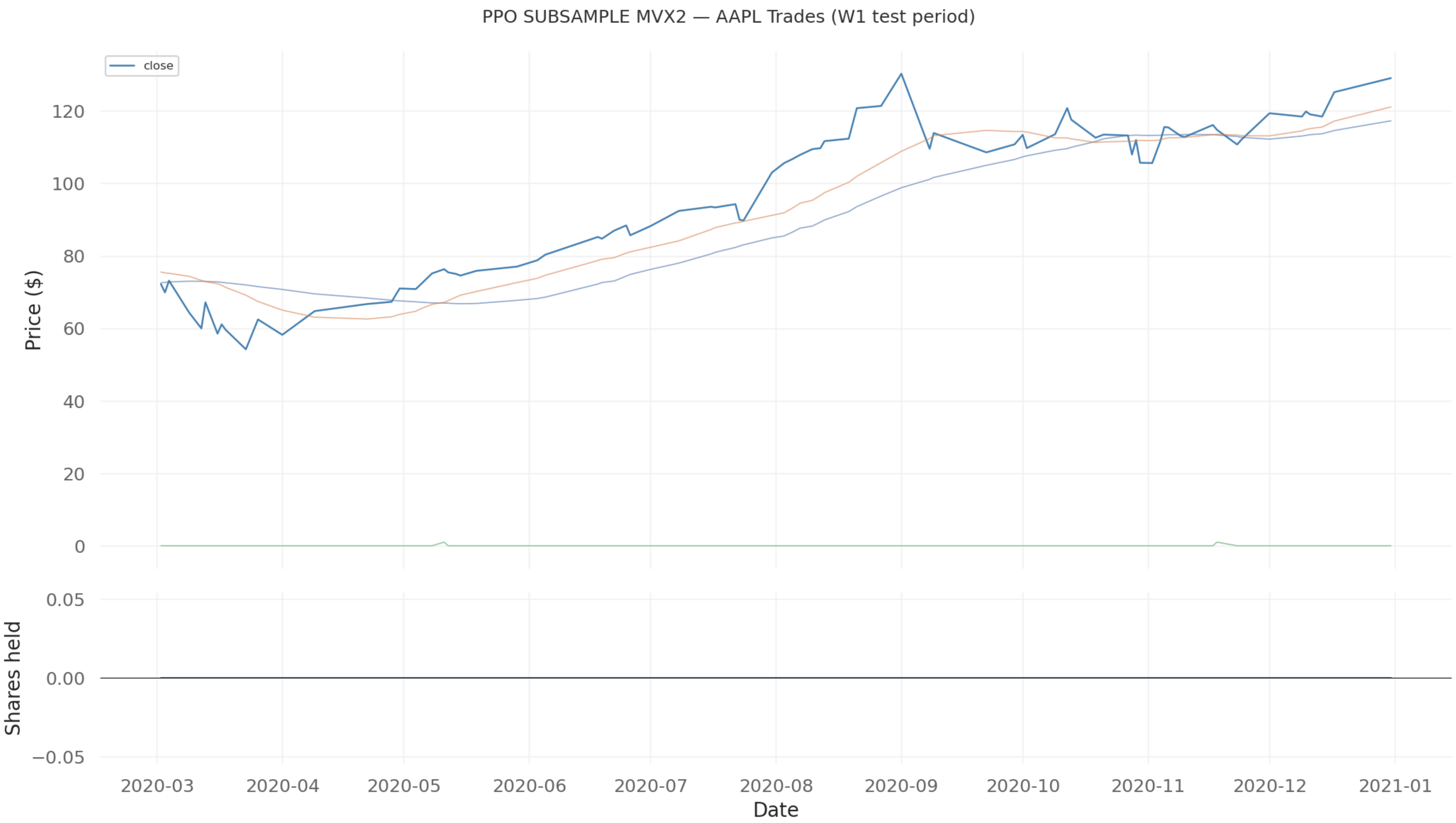
- H1a: Trigger based improves results against all-in-one and baselines
- H1b: How? Using Triggers and Subsampling
- H1c: Why? -> Good question...

Core finding: RL algorithm work better with event-driven representation in Equity Markets.

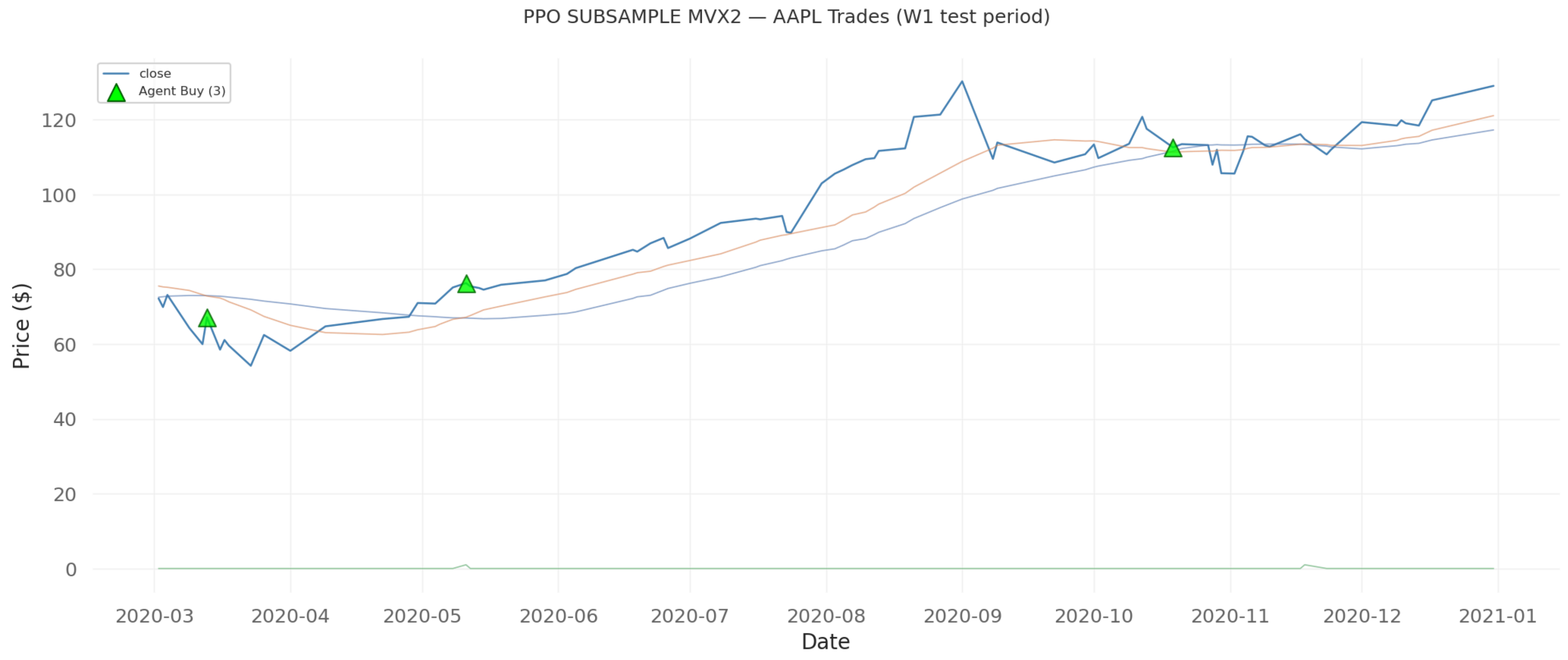
Interpretability: Agent Apple Stocks



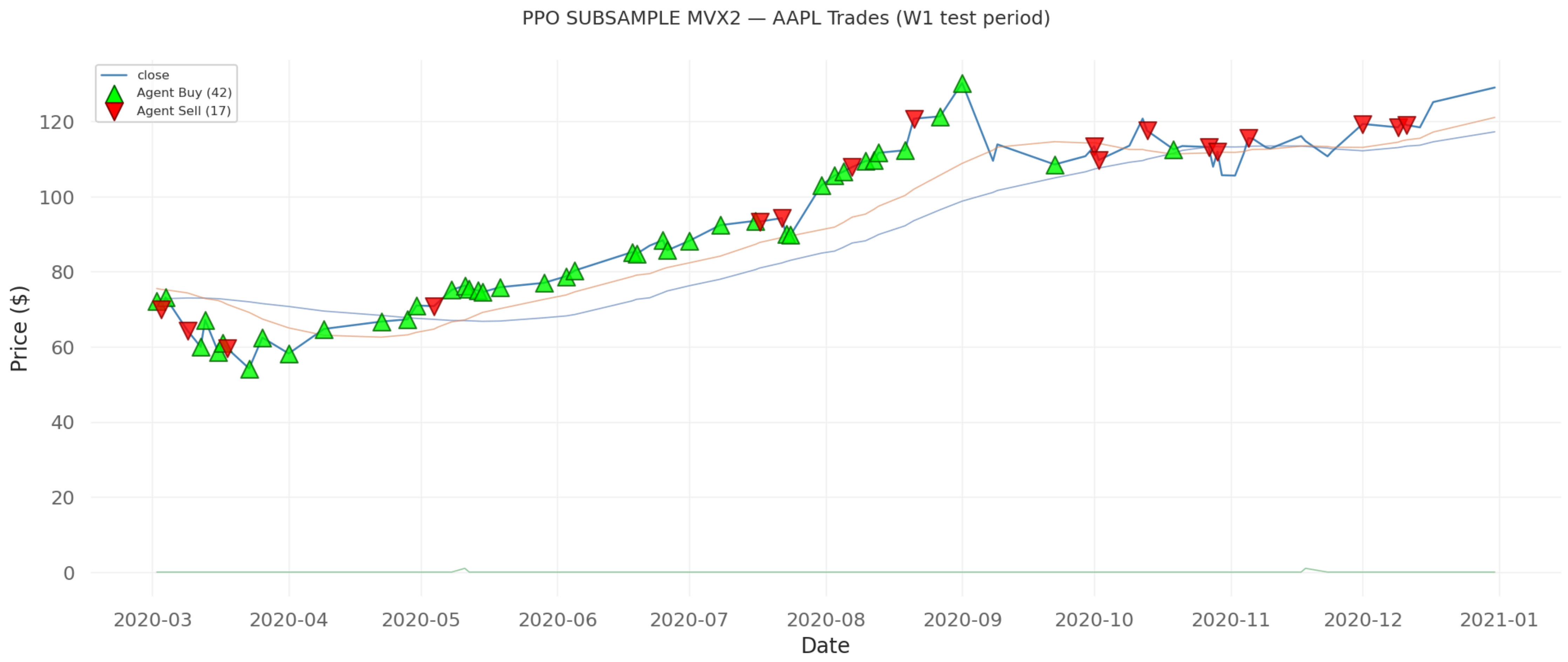
Interpretability: Agent Apple Stocks



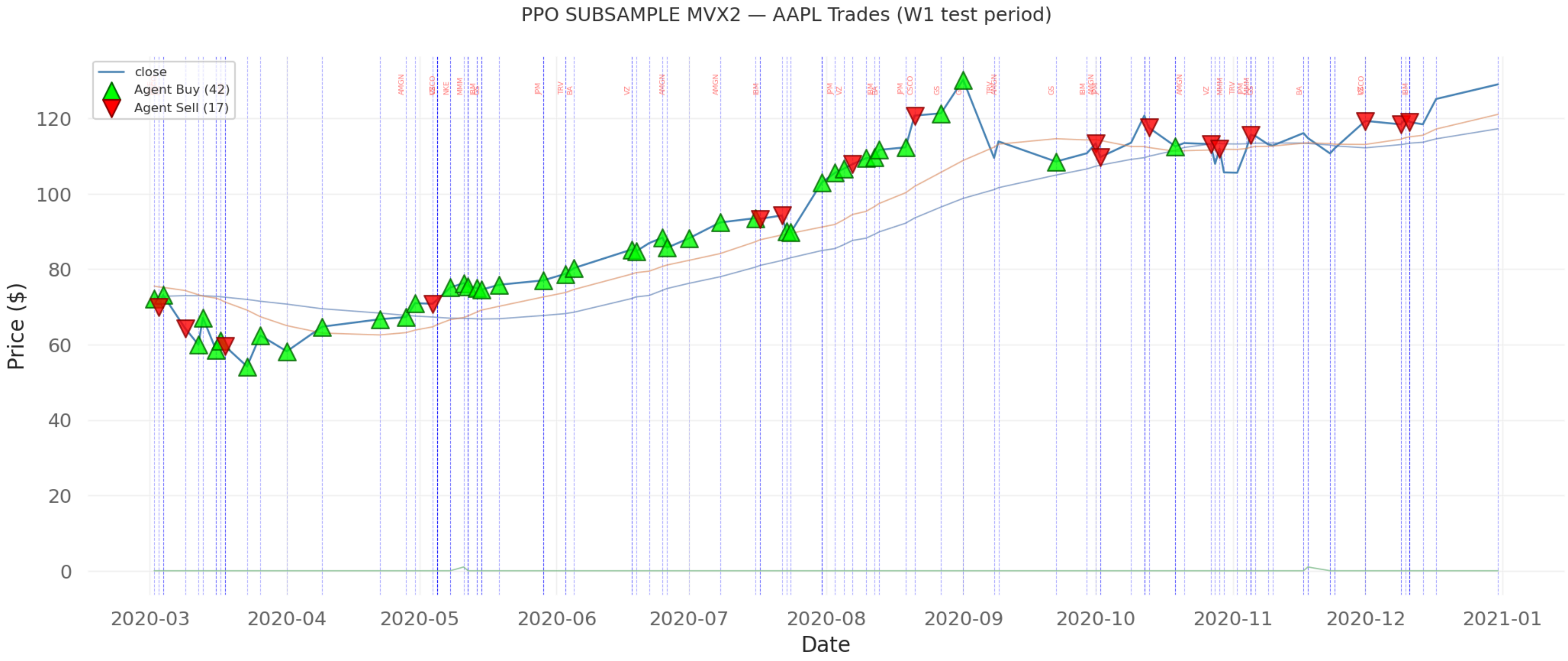
Interpretability: Agent Apple Stocks



Interpretability: Agent Apple Stocks



Interpretability: Agent Apple Stocks



Interpretability: Events really mathers?

1. Event vs random
   - proves events carry information
2. Per-ticker shuffle
   - proves temporal structure matters
3. Cross-asset shuffle
   - proves synchronization matters
4. Feature ablation for single stocks
   - proves which assets matter

ALgorithms: PPO vs SAC & A2C?

1. PPO relies heavily on GAE (Generalized Advantage Estimation)
    - Subsampling introduces jumps
2. Clipping + Distribution Shift (PPO instability)
    - Jumps caused instability in policy updates
3. On-policy vs Off-policy Data Efficiency
    - But A2C is on policy!

Questions?