

Bocconi Students Investment Club

Polymarket Microstructure Analysis Project

Behavioral Wallet Classification, Kyle's λ Estimation, Avellaneda–
Stoikov Market Making, and Directional Strategy Backtesting on
the 2024 US Presidential Election Markets

Authors

Bocconi Students Investment Club

Pieter Christiaan Rotteveel

Martin Patrikov

Rayi Makori

2026

Polymarket Microstructure Analysis Project

Overview

The goal of this project was to understand the Polymarket microstructure and the effect of different types of wallets on the price. This research utilizes public 2024 US Election data from Polymarket to analyze market microstructure and the efficacy of algorithmic trading strategies within decentralized prediction markets. The study first develops a behavioral taxonomy by clustering wallet addresses into interpretable participant types such as informed, retail, or arbitrageur, and then quantifies information asymmetry by estimating type-specific Kyle's λ to measure price impact. These empirical findings inform the design and back testing of an Avellaneda–Stoikov-style market-making framework specifically adapted to manage the unique inventory risks associated with political event contracts. Finally, the project evaluates the performance of mechanical edge strategies, assessing whether systematic mean-reversion or programmatic wallet-copying can generate consistent alpha in high-volatility, binary-outcome environments.

Executive Summary

This paper analyzes the microstructure of Polymarket's 2024 US election markets using trade-level data. We show that behavioural wallet classification resolves biased price-impact estimates, that order flow has persistent positive impact (Kyle's $\lambda > 0$), and that passive market making is structurally profitable under realistic frictions. In contrast, directional strategies such as mean-reversion and wallet-copying fail after transaction costs. These results suggest that prediction markets in this setting are dominated by retail flow, rewarding liquidity provision over information-based trading.

Prediction Markets and the Polymarket Protocol

A prediction market is a specialized financial exchange designed to aggregate information and forecast the likelihood of future events through the collective wisdom of its participants. Unlike traditional opinion polling, which relies on stated preferences, prediction markets require participants to back their beliefs with financial capital, creating a "skin in the game" incentive structure that filters out noise and personal bias. This mechanism operates on the economic principle that market prices reflect the most accurate, real-time probability of an event's occurrence as traders compete to profit from their private information or superior analysis. By incentivizing accuracy, these markets often outperform expert pundits and traditional data collection methods in predicting complex outcomes.

The functional mechanics of these markets typically involve the trade of binary shares that represent a specific outcome, such as the result of a political election or a policy decision. These shares are priced on a scale from zero to one dollar, where the current trading price directly correlates to the market's perceived probability of the event. For instance, a share trading at \$0.60 implies a 60% probability of success. Upon the resolution of the event, the shares associated with the winning outcome are redeemed for exactly \$1.00, while all other shares expire worthless.

All the views expressed are opinions of Bocconi Students Investment Club members and can in no way be associated with Bocconi University. All the financial recommendations offered are for educational purposes only. Bocconi Students Investment Club declines any responsibility for eventual losses you may incur implementing all or part of the ideas contained in this website. The Bocconi Students Investment Club is not authorised to give investment advice. Information, opinions, and estimates contained in this report reflect a judgment at its original date of publication by Bocconi Students Investment Club and are subject to change without notice. The price, value of and income from any of the securities or financial instruments mentioned in this report can fall as well as rise. Bocconi Students Investment Club does not receive compensation and has no business relationship with any mentioned company.

This structure distinguishes the platform from a traditional sportsbook, as participants trade against one another in a peer-to-peer fashion rather than against a centralized bookmaker who sets fixed odds and takes the opposite side of every bet.

Within the modern landscape of prediction markets, two primary models have emerged as industry leaders, represented by Kalshi and Polymarket. Kalshi operates as a highly regulated exchange within the United States, overseen by the Commodity Futures Trading Commission (CFTC). It functions within the traditional financial system, utilizing US dollars and maintaining a centralized custody model to facilitate hedging against macroeconomic shifts and political events. While Kalshi provides a regulated avenue for domestic participants, its centralized nature and restricted access create a different data environment compared to its decentralized counterparts, as transaction history and participant behavior are often shielded from public analysis.

In contrast, this research focuses on Polymarket, the world's largest decentralized prediction market built on the Polygon blockchain. Polymarket utilizes a hybrid-decentralized model that combines an off-chain Central Limit Order Book (CLOB) for high-speed matching with on-chain settlement via smart contracts and the USDC stablecoin. This decentralized architecture provides an unprecedented level of transparency, as every trade, wallet interaction, and order flow event is recorded on a public ledger. Such transparency allows for the granular analysis of market microstructures. This is why we opted to do our project on Polymarket data.

This paper makes three main contributions. First, we show that trader classification based on ex-post performance introduces significant bias in price-impact estimation, and that behavioural classification resolves this issue. Second, we provide evidence that order flow in prediction markets has persistent positive price impact, particularly in high-information regimes. Third, we demonstrate that liquidity provision strategies outperform directional strategies under realistic execution frictions, suggesting that these markets are structurally favourable to market makers.

Data

The project utilizes trade-level data sourced from the SII-WANGZJ/Polymarket_data repository on HuggingFace [6], encompassing a total of 38 distinct prediction markets. This dataset includes the national "Presidential Election Winner 2024" market, which represents 17 candidates and approximately \$3.7bn in volume, as well as 21 swing-state markets covering Pennsylvania, Georgia, Wisconsin, Michigan, Arizona, Nevada, and North Carolina. The primary analysis focuses on the Trump and Harris markets due to their high liquidity and volume. This focus introduces a selection bias toward high-liquidity markets, but ensures that microstructure effects such as price impact and inventory dynamics are identifiable. Lower-probability markets, such as those for Obama or RFK, were maintained in the dataset but excluded from the headline analysis because their proximity to zero-price boundaries introduces significant boundary effects that are incompatible with the Kyle and Avellaneda-Stoikov frameworks. The data processing pipeline evolved from initial repository setup and daily sample sanity checks to the generation of hourly mid-price series and sophisticated wallet-level features. These metrics, such as trade frequency, maker ratios, inventory roughness, and realized P&L per volume, allow for a granular assessment of market participant behavior and inventory risk.

All the views expressed are opinions of Bocconi Students Investment Club members and can in no way be associated with Bocconi University. All the financial recommendations offered are for educational purposes only. Bocconi Students Investment Club declines any responsibility for eventual losses you may incur implementing all or part of the ideas contained in this website. The Bocconi Students Investment Club is not authorised to give investment advice. Information, opinions, and estimates contained in this report reflect a judgment at its original date of publication by Bocconi Students Investment Club and are subject to change without notice. The price, value of and income from any of the securities or financial instruments mentioned in this report can fall as well as rise. Bocconi Students Investment Club does not receive compensation and has no business relationship with any mentioned company.

Trader Segmentation and Behavioural Classification

Clustering is an unsupervised learning technique that partitions a set of observations into groups such that members of the same group are more similar to each other than to members of other groups, where similarity is defined by a distance metric in feature space. The most widely used algorithm, k-means, formalizes this as an optimization problem: given n observations $\{x_1, \dots, x_n\} \subset \mathbb{R}^d$ and a chosen number of clusters K the algorithm seeks centroids minimizing the within-cluster sum of squares, $\sum_{k=1}^K \sum_{x_i \in C_k} \|x_i - \mu_k\|^2$. This is solved iteratively by Lloyd's algorithm [4], which alternates between assigning each point to its nearest centroid and recomputing each centroid as the mean of its assigned points; the procedure converges to a local minimum and is typically restarted from multiple random seeds. Because k-means is sensitive to feature scale and to the heavy tails common in financial data, the inputs are standardized to zero mean and unit variance, and right-skewed variables (trade counts, sizes, holding periods) are first compressed with a $\log(1+x)$ transform. Cluster quality is assessed with the silhouette coefficient $s(i) = (b(i) - a(i)) / \max(a(i), b(i))$ where $a(i)$ is the mean intra-cluster distance and $b(i)$ the mean nearest-other-cluster distance, taking values in $[-1, 1]$. Rule-based classification is a complementary approach in which group membership is determined by hand-coded thresholds on interpretable features rather than geometric proximity in a scaled feature space; it sacrifices flexibility for transparency and is appropriate when the goal is downstream causal inference rather than pure description.

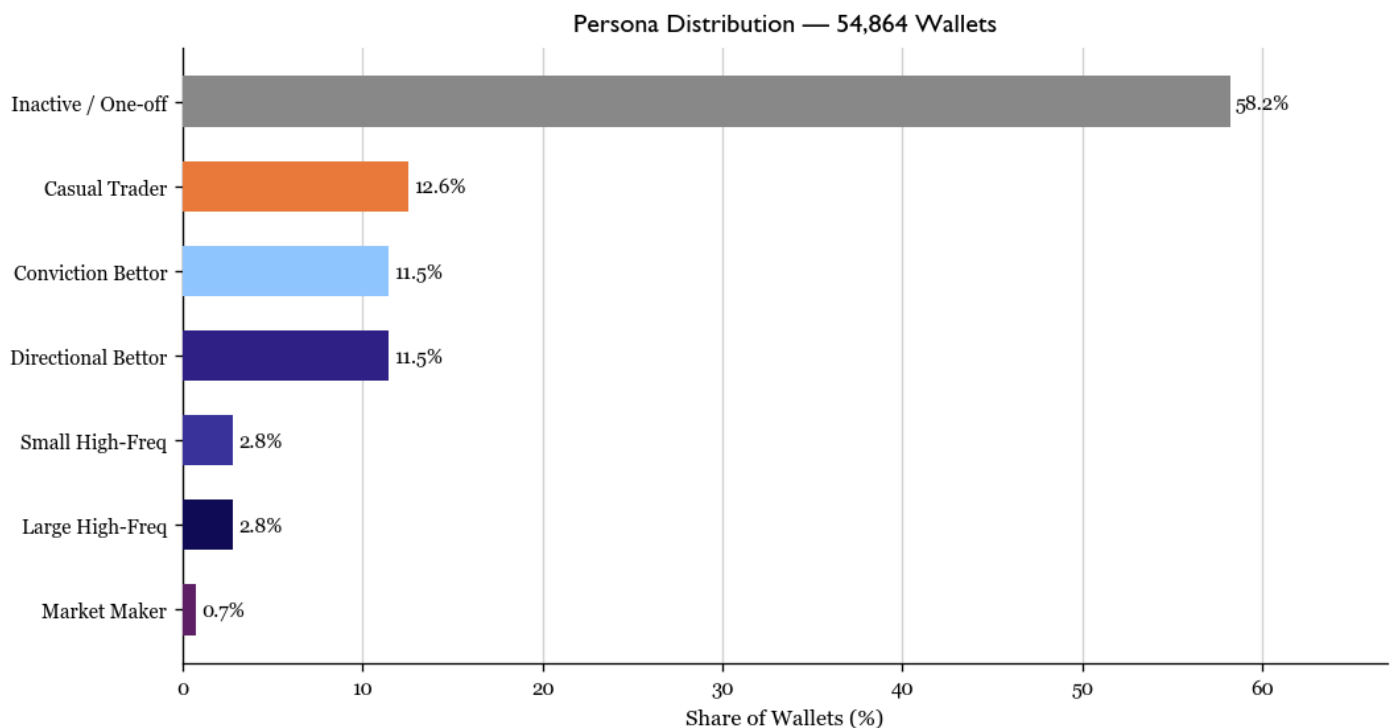
We applied both approaches in sequence. An initial k-means run with $K = 3$ on standardised, log-transformed wallet features (trade frequency, average trade size, holding period, inventory roughness, realised PnL per dollar volume, number of markets) produced three coherent groups that we labelled Market Maker, Directional Trader, and Retail Trader, with a silhouette of around 0.3. We then moved to a rule-based classifier that split wallets into five personas (Informed Bettor, Market Maker, Lucky Bettor, Negative PnL, and Inactive / One-off) using interpretable thresholds on frequency, maker ratio, buy/sell balance, and realized PnL. This five-persona scheme fed directly into our first attempt at decomposing Kyle's λ , a measure of price impact whose theoretical foundations and full specification we develop in the next section.

A key challenge in this step is avoiding look-ahead bias. Any classification based on realised PnL over the full sample risks incorporating future information into trader labels, which can distort downstream price-impact estimates. In v1(`week3\_kyles\_lambda.ipynb`) we estimated a separate λ_k per persona and found that the Informed Bettor coefficient was significantly negative on three of our five top markets, the wrong sign relative to theory, which predicts that informed flow moves price in its own direction. In v2(`week3\_kyles\_lambda\_v2.ipynb`) we added winsorisation, a `signed\_log` flow transform, HAC standard errors, and market fixed effects; R^2 on the Trump market collapsed from 0.026 to 0.008 and on Harris from 0.089 to 0.009, which we traced to a units bug in the `signed\_log` transform. In v3(`week3\_kyles\_lambda\_v3.ipynb`) we applied three isolated fixes: dropping the buggy transform, switching Δp to log-odds, and, most importantly, rebuilding personas using only pre-Aug-31 trades to remove ex-post PnL contamination. The wrong-signed Informed Bettor result disappeared once labels were de-contaminated; it had been a labelling artifact rather than a property of the market.

The canonical replacement pipeline(`week3\_updated\_`) removes PnL from classification entirely in favour of seven behavioural personas, adds a lag-1 regressor, restricts pooling to the two large markets (Trump and Harris),

All the views expressed are opinions of Bocconi Students Investment Club members and can in no way be associated with Bocconi University. All the financial recommendations offered are for educational purposes only. Bocconi Students Investment Club declines any responsibility for eventual losses you may incur implementing all or part of the ideas contained in this website. The Bocconi Students Investment Club is not authorised to give investment advice. Information, opinions, and estimates contained in this report reflect a judgment at its original date of publication by Bocconi Students Investment Club and are subject to change without notice. The price, value of and income from any of the securities or financial instruments mentioned in this report can fall as well as rise. Bocconi Students Investment Club does not receive compensation and has no business relationship with any mentioned company.

and raises the minimum-trades filter from five to ten. It produces $R^2 = 0.134$ for Trump, 0.096 for Harris, and 0.051 pooled, and on the large markets no persona carries a significantly wrong-signed λ , confirming that the earlier anomaly was generated by the classification step rather than by the price-impact regression itself. The absolute level of R^2 here is expected in high-frequency settings; low explanatory power is the norm in the microstructure literature, where per-bin price changes are dominated by news arrivals, inventory rebalancing, and noise that order flow alone cannot explain. The economic significance lies in coefficient stability and sign consistency rather than fit: all persona-level coefficients are correctly signed, statistically significant, and stable across specifications, and those properties are precisely what the updated pipeline delivers.



Source: Polymarket trades_election_2024 dataset, Sep–Nov 2024

Price Impact Estimation: Kyle's λ

In Kyle's (1985) [1] seminal model of informed trading, a single risk-neutral insider with private information about an asset's terminal value submits a market order alongside an exogenous flow of liquidity ("noise") trades, and a competitive market maker observes only the aggregate net order flow before setting a price. Because the market maker cannot tell informed orders from noise, she rationally adjusts the price linearly in the total flow she observes; the slope of that adjustment is what the literature calls Kyle's λ .

Mechanically, λ is the inverse of market depth: a low λ means a market can absorb large directional orders with little price movement, while a high λ means the market maker treats incoming flow as highly informative and shifts quotes aggressively. The empirical baseline regression derived from this setup is $\Delta P_t = \alpha + \lambda \cdot \text{OFI}_t + \varepsilon_t$ where ΔP_t is the price change over bin t and OFI_t is the signed (buyer-initiated minus seller-initiated) order flow

All the views expressed are opinions of Bocconi Students Investment Club members and can in no way be associated with Bocconi University. All the financial recommendations offered are for educational purposes only. Bocconi Students Investment Club declines any responsibility for eventual losses you may incur implementing all or part of the ideas contained in this website. The Bocconi Students Investment Club is not authorised to give investment advice. Information, opinions, and estimates contained in this report reflect a judgment at its original date of publication by Bocconi Students Investment Club and are subject to change without notice. The price, value of and income from any of the securities or financial instruments mentioned in this report can fall as well as rise. Bocconi Students Investment Club does not receive compensation and has no business relationship with any mentioned company.

imbalance over the same bin. The slope $\hat{\lambda}$ is interpreted as the basis-point price move per unit of net flow and is by construction expected to be positive: net buying pressure ought to push prices up, and net selling pressure down. Because the textbook regression collapses every trader into a single OFI series, it cannot tell us which traders are doing the price moving.

Three design choices deserve explicit justification. First, the dependent variable is the change in log-odds rather than the change in raw price. Because Polymarket contracts are bounded between zero and one, raw price differences are asymmetric near the boundaries, and the residuals of an OLS regression on raw prices would be heteroskedastic by construction. Log-odds linearise the price scale, produce approximately symmetric increments in the interior of the price range, and ensure that a one-unit move at a mid-probability of 0.5 is treated consistently with a one-unit move at 0.7, which is not the case when working in price space. Second, trades are aggregated into hourly bins. Finer granularity, such as five-minute or trade-by-trade panels, increases statistical power in principle but substantially raises the noise-to-signal ratio on a market where many bins would contain only one or two trades; coarser bins, such as daily, discard intraday variation that is the primary source of identifying variation for the Kyle regression. Hourly bins represent a practical compromise between these extremes, yielding a manageable panel size while preserving meaningful within-day dynamics. Third, order flow imbalance is measured in USD notional rather than contract units. Contracts on different Polymarket markets trade at very different prices, meaning that a one-contract buy at a price of 0.10 represents a different amount of capital commitment than a one-contract buy at 0.70. USD-denominated flow normalises this difference and produces price impact coefficients that are directly comparable across markets and price regimes.

We therefore decomposed the flow per persona and estimated $\Delta p_t = \alpha + \sum_{k=1}^K \lambda_k x_{k,t} + \varepsilon_t$ where $x_{k,t}$ is the net signed USD flow (in \$1k units) attributable to persona k in bin t , and λ_k is the persona-specific price impact. In the canonical updated specification, we extend this to include a one-period lag, $\Delta p_t = \alpha + \sum_k \lambda_k x_{k,t} + \sum_k \gamma_k x_{k,t-1} + \varepsilon_t$, following Hasbrouck (1991) [2]. The contemporaneous λ_k captures immediate impact, γ_k captures slow price discovery, and the sum $\lambda_k + \gamma_k$ is the total permanent impact per persona. Estimation is by OLS with Newey-West [5] HAC standard errors using lags $L = \max(1, \lfloor 4(N/100)^{2/9} \rfloor)$ to handle residual autocorrelation and heteroskedasticity. Flow regressors are winsorised at the 1% level per market to dampen the influence of a handful of whale trades, bins with fewer than 10 trades are dropped (the v1 to v3 path showed that thinner bins are dominated by quoting noise), and pooled regressions across Trump and Harris carry market fixed effects so the estimated λ_k are not contaminated by between-market level differences.

Estimating this on the post-Sep-1 hourly panel, the canonical specification produces a per-market R^2 of 0.134 on Trump and 0.096 on Harris with the lag-1 term included, and a pooled $R^2 = 0.051$ across the two markets. The contemporaneous-only baselines are roughly half those values, and the gap is what justifies the lag-1 extension: a meaningful fraction of price impact arrives one bin late, consistent with slow incorporation of information by the market maker. On Trump, the High-Frequency and Conviction-Bettor personas carry positive and statistically significant total impact, while the Inactive / One-off bucket is indistinguishable from zero, which is exactly what theory predicts: traders who post directional, sustained flow move prices, and small one-off bettors do not. Crucially, no persona on the two large markets carries a wrong-signed λ at conventional significance levels. The "Informed Bettor $\lambda < 0$ " result that haunted v1 and v2 is gone, and we attribute its disappearance entirely to

All the views expressed are opinions of Bocconi Students Investment Club members and can in no way be associated with Bocconi University. All the financial recommendations offered are for educational purposes only. Bocconi Students Investment Club declines any responsibility for eventual losses you may incur implementing all or part of the ideas contained in this website. The Bocconi Students Investment Club is not authorised to give investment advice. Information, opinions, and estimates contained in this report reflect a judgment at its original date of publication by Bocconi Students Investment Club and are subject to change without notice. The price, value of and income from any of the securities or financial instruments mentioned in this report can fall as well as rise. Bocconi Students Investment Club does not receive compensation and has no business relationship with any mentioned company.

dropping the `signed\_log` flow transform and to building persona labels from pre-Aug-31 trades rather than the full sample.

The headline numbers should be read with appropriate calibration. As noted earlier, R^2 values in the 0.05 to 0.15 band are what the high-frequency microstructure literature routinely reports, because hourly price changes are dominated by news arrivals, inventory rebalancing, and microstructure noise that order flow alone cannot explain. The economically interesting quantities here are the persona-level coefficients: their signs, their relative magnitudes, and their statistical significance. By those criteria, the canonical pipeline produces a clean, theory-consistent decomposition of price impact across trader types on the two markets where Kyle's λ is identified well enough to be trusted.

Optimal Market Making: The Avellaneda–Stoikov Framework

Market making is the process of providing liquidity to a financial market by continuously quoting two prices: a bid price at which the market maker is willing to buy, and an ask price at which the market maker is willing to sell. The difference between these two prices is the bid–ask spread. This spread is the market maker's main source of revenue, but it is not risk-free. A market maker earns the spread by standing ready to trade with other participants, which exposes it to inventory risk and adverse selection risk.

Inventory risk occurs when the market maker accumulates a large position on one side of the market. For example, if the strategy repeatedly buys Trump YES tokens and the Trump price then falls, the market maker loses money on its inventory. Adverse selection occurs when informed traders trade against the market maker's quotes shortly before the price moves. In that case, the market maker may buy just before the price falls or sell just before the price rises. Therefore, successful market making is not only about earning the spread, but also about managing inventory and avoiding being picked off by better-informed flow.

This is especially relevant on Polymarket because the contracts are binary. A YES token settles at 1 if the event happens and 0 if it does not. Therefore, the price of a YES token can be interpreted as the market's implied probability of the event occurring. In the 2024 US Presidential Election markets, a Trump YES price of 0.60 means that the market is pricing Trump as having approximately a 60% chance of winning. A market maker in this setting provides liquidity by posting buy and sell orders around this implied probability.

Polymarket uses an order book structure. An order book is a live record of all outstanding limit orders in a market. The bid side contains orders from traders willing to buy the contract, while the ask side contains orders from traders willing to sell it. The highest bid is called the best bid, and the lowest ask is called the best ask. The difference between them is the spread, and the midpoint between them is often used as a simple estimate of the current fair price. When a trader wants to trade immediately, they cross the spread by buying from the best ask or selling into the best bid. The market maker's role is to keep these bids and ask quotes available so that other traders can trade without waiting for a direct counterparty.

In our project, we did not have full historical snapshots of Polymarket's limit order book. Instead, we reconstructed the market's microstructure from trade-level data across the 2024 election markets. We used the trade data to

All the views expressed are opinions of Bocconi Students Investment Club members and can in no way be associated with Bocconi University. All the financial recommendations offered are for educational purposes only. Bocconi Students Investment Club declines any responsibility for eventual losses you may incur implementing all or part of the ideas contained in this website. The Bocconi Students Investment Club is not authorised to give investment advice. Information, opinions, and estimates contained in this report reflect a judgment at its original date of publication by Bocconi Students Investment Club and are subject to change without notice. The price, value of and income from any of the securities or financial instruments mentioned in this report can fall as well as rise. Bocconi Students Investment Club does not receive compensation and has no business relationship with any mentioned company.

calculate YES-implied prices and built an hourly VWAP mid-price series. This mid-price served as the anchor around which our simulated market makers posted limit orders. The backtester then checked whether incoming trades would have crossed the market maker's bid or ask, and if so, recorded a fill. This allowed us to approximate the role of a liquidity provider even without observing the full historical order book.

To improve on a simple fixed-spread strategy, we implemented the Avellaneda–Stoikov market-making model. The Avellaneda–Stoikov framework [3] is a quantitative model for setting optimal bid and ask quotes when a market maker faces both price risk and inventory risk. Unlike a naive market maker, which always quotes the same spread around the mid-price, the Avellaneda–Stoikov model adjusts its quotes dynamically depending on the market maker's current inventory, volatility, liquidity conditions, and trading horizon.

The central concept in the Avellaneda–Stoikov model is the reservation price. The reservation price is the price at which the market maker is indifferent between buying and selling, given its current inventory.

It is defined as: $r = S - q\gamma\sigma^2T$ where r is the reservation price, S is the current mid-price, q is the current inventory, γ is the risk-aversion parameter, σ^2 is the variance of price changes, and T is the trading horizon. If the market maker has positive inventory, meaning it is long tokens, the reservation price falls below the mid-price. This makes the strategy more willing to sell and less willing to buy, helping reduce the long position. If the market maker is short, the reservation price rises above the mid-price, encouraging the strategy to buy back inventory.

After calculating the reservation price, the model places bid and ask quotes around it. The spread depends on risk aversion, volatility, and a liquidity parameter k , which represents how quickly order arrival rates decline as quotes move further away from the mid-price. In our implementation, the baseline half-spread was approximately one percentage point, producing a two-percentage point total spread at zero inventory. As inventory moved toward the limit, the A-S model shifted both quotes up or down to manage exposure. Neither gamma nor k is directly observable from trade data. Gamma was treated as a free parameter reflecting the strategy's aversion to holding unhedged inventory in a binary contract; k was estimated from the empirical distribution of trade arrival rates at different distances from the mid-price, measured over the pre-election training window. In practice, we calibrated both parameters jointly via a grid search over the Week 4 backtest sample, selecting the combination that minimised inventory variance subject to achieving a target fill rate. Sensitivity analysis confirmed that the strategy's advantage over the naive benchmark is robust across a wide range of parameter choices, with the main effect of misspecification being a shift in the level of inventory risk rather than a reversal of the qualitative results.

We adapted the standard Avellaneda–Stoikov model to fit Polymarket's binary prediction markets. First, because Polymarket prices are bounded between 0 and 1, all quotes were clipped to remain inside valid price limits. Second, instead of using time-to-resolution as the horizon, we used a rolling one-day horizon. This was more appropriate for our setting because the market was driven by daily political news cycles rather than by a simple final liquidation deadline. Third, we estimated daily variance from hourly mid-price changes, so that the model's inventory skew reflected the observed volatility of the Trump and Harris markets.

All the views expressed are opinions of Bocconi Students Investment Club members and can in no way be associated with Bocconi University. All the financial recommendations offered are for educational purposes only. Bocconi Students Investment Club declines any responsibility for eventual losses you may incur implementing all or part of the ideas contained in this website. The Bocconi Students Investment Club is not authorised to give investment advice. Information, opinions, and estimates contained in this report reflect a judgment at its original date of publication by Bocconi Students Investment Club and are subject to change without notice. The price, value of and income from any of the securities or financial instruments mentioned in this report can fall as well as rise. Bocconi Students Investment Club does not receive compensation and has no business relationship with any mentioned company.

The market-making strategies were implemented in `src/strategies/market_maker.py`. We created two strategy classes: `NaiveMM` and `AvellanedaStoikovMM`. The naive benchmark posted a fixed bid and ask around the mid-price, with no inventory-based skew. The A-S strategy used the reservation-price formula to adjust quotes dynamically. Both strategies used the same quote size, inventory limits, backtest window, and execution engine, allowing for a clean comparison.

We tested the strategies on the two most liquid Polymarket election markets: Donald Trump and Kamala Harris, from September 1 to November 4, 2024. We deliberately stopped before election day because prices close to resolution move sharply toward either 0 or 1, creating boundary effects that require additional circuit breakers. The Week 4 implementation therefore focused on the high-liquidity, news-driven pre-election period.

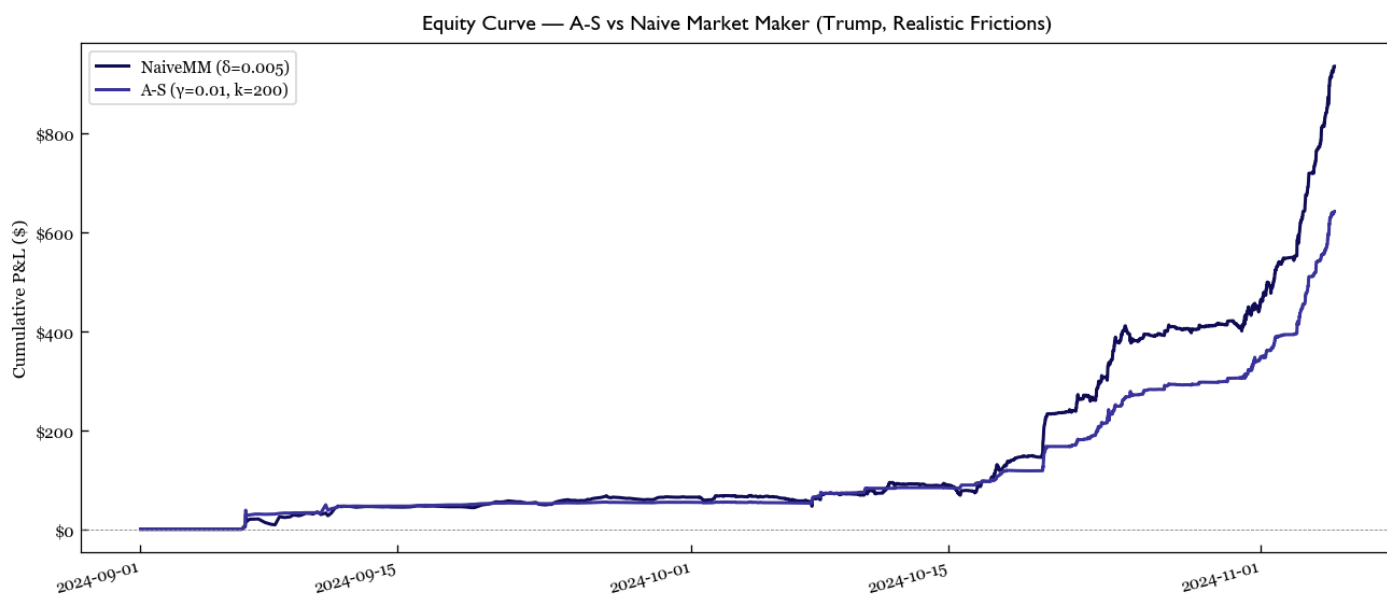
In the initial backtests, the Avellaneda–Stoikov strategy outperformed the naive benchmark on both markets. On the Trump market, A-S generated \$886.22 of P&L, compared with \$772.86 for the naive strategy. On the Harris market, A-S generated \$935.06, compared with \$855.89 for the naive strategy. A-S also produced better risk-adjusted performance, with higher Sharpe ratios and lower maximum drawdowns. This showed that the inventory-skew mechanism was working as intended: the strategy captured the spread while reducing exposure when inventory became too large.

In Week 5, we made the back-test more realistic by adding trading frictions. The first adjustment was a queue model, where only 20% of would-fill orders were assumed to actually fill. This reflected the fact that real market makers are often behind other orders at the same price. The second adjustment was a 10 basis point fee assumption. The third adjustment was a one-hour markout calculation, which measured whether fills were followed by favourable or unfavourable price movements. This helped test whether the strategy was being adversely selected.

The realistic-friction results showed that passive market making remained profitable across the Trump and Harris markets. The naive strategy with a tighter fixed spread was competitive in raw adjusted P&L, but the Avellaneda–Stoikov model remained valuable because it controlled risk better. Its main advantage was not simply higher return, but better inventory management and lower drawdowns. This distinction is important: the A-S model earned its complexity by improving the stability of the strategy rather than mechanically increasing profits in every configuration.

Overall, the market-making section of the project shows that liquidity provision was a strong strategy in the 2024 Polymarket election markets. The order-flow environment was largely driven by retail sentiment and short-term noise, which allowed passive market makers to capture the spread without being consistently picked off by informed traders. Compared with aggressive directional strategies, market making was more robust because it did not require forecasting the winner of the election. Instead, it profited from the microstructure of the market itself: traders' demand for immediacy, repeated crossing of the spread, and short-term price mean reversion.

All the views expressed are opinions of Bocconi Students Investment Club members and can in no way be associated with Bocconi University. All the financial recommendations offered are for educational purposes only. Bocconi Students Investment Club declines any responsibility for eventual losses you may incur implementing all or part of the ideas contained in this website. The Bocconi Students Investment Club is not authorised to give investment advice. Information, opinions, and estimates contained in this report reflect a judgment at its original date of publication by Bocconi Students Investment Club and are subject to change without notice. The price, value of and income from any of the securities or financial instruments mentioned in this report can fall as well as rise. Bocconi Students Investment Club does not receive compensation and has no business relationship with any mentioned company.



Execution Frictions and Strategy Robustness

The initial Week 4 results showed extremely strong performance for both the naive and Avellaneda–Stoikov market makers, with annualized Sharpe ratios in the range of 7 to 12. While these results confirmed that the strategy logic was sound, they were not realistic from a market microstructure perspective. The backtester in Week 4 assumed that any limit order would be filled immediately whenever a trade printed at or through the quoted price. In practice, this assumption is too optimistic because it ignores queue position, transaction costs, and adverse selection. As a result, Week 5 focused on making the simulation more realistic and stress-testing the strategies under a defensible execution model.

The first major adjustment was the introduction of a queue model. In real order books, a market maker is rarely first in line at a given price level. Instead, it is typically competing with other resting orders, and only a fraction of incoming trades will actually execute against its quotes. To approximate this, we introduced a fill probability parameter, $\text{fill_probability} \in (0,1]$, applied as a Bernoulli filter on potential fills. A value of 1.0 reproduces the Week 4 setup where all eligible orders fill, while lower values simulate being deeper in the queue. In our realistic specification, we set $\text{fill_probability} = 0.20$, which corresponds to roughly being behind four other orders at the same price level. This adjustment significantly reduced the number of fills and therefore the apparent profitability of high-frequency quoting strategies.

The second adjustment was the inclusion of trading fees. Although Polymarket does not charge traditional exchange fees, trading incurs blockchain-related costs that are non-negligible at scale. We approximated this by introducing a fee rate of 10 basis points ($\text{fee_rate} = 0.001$) applied to each trade. This penalizes excessive turnover and makes strategies with very tight spreads or high trading frequency less attractive. In particular, it forces a more realistic trade-off between capturing the spread and paying execution costs.

All the views expressed are opinions of Bocconi Students Investment Club members and can in no way be associated with Bocconi University. All the financial recommendations offered are for educational purposes only. Bocconi Students Investment Club declines any responsibility for eventual losses you may incur implementing all or part of the ideas contained in this website. The Bocconi Students Investment Club is not authorised to give investment advice. Information, opinions, and estimates contained in this report reflect a judgment at its original date of publication by Bocconi Students Investment Club and are subject to change without notice. The price, value of and income from any of the securities or financial instruments mentioned in this report can fall as well as rise. Bocconi Students Investment Club does not receive compensation and has no business relationship with any mentioned company.

The third and most important addition was the incorporation of markout costs, which capture adverse selection. For each fill, we computed the difference between the execution price and the mid-price at a fixed horizon in the future (one hour in our baseline). This is implemented through the function `compute_markout_cost(fills, midprices, markout_seconds)`, which assigns a per-fill markout based on subsequent price movement. If the price moves against the position after the trade, the markout is negative, indicating that the market maker was “picked off” by informed flow. If the price reverts favorably, the markout is positive. Aggregating these markouts across all fills provides a direct measure of whether the strategy is systematically trading against better-informed participants.

Together, these three components (queue position, fees, and markout) transform the backtest from an optimistic execution model into a defensible approximation of real trading conditions. Importantly, they affect different aspects of the strategy. The queue model reduces fill frequency, fees reduce net profitability per trade, and markout captures the informational disadvantage of passive liquidity provision.

To evaluate robustness, we ran a parameter sweep across both the Avellaneda–Stoikov and naive strategies. The grid consisted of 14 unique parameter configurations (varying risk aversion γ , liquidity parameter k , spread δ , and inventory limits), applied to both the Trump and Harris markets, for a total of 28 backtests under each fill model. The realistic grid was generated using the standalone script `notebooks/_run_week5_realistic.py`, which takes approximately 5–10 minutes to execute.

The results are stored in `data/processed/week5_grid_results_realistic.csv` and analyzed in `week5_mm_tuning.ipynb`, which loads the cached results and produces comparison tables almost instantly.

The key outcome of this exercise is that the ranking of strategies changes once realistic frictions are introduced. Under the optimistic Week 4 model, very tight spreads and aggressive quoting (e.g., high k in the A-S model) dominated because every quote was assumed to fill. Under the realistic model, however, these strategies suffer from lower effective fill rates and higher adverse selection. As a result, a simpler naive strategy with a slightly wider spread becomes competitive or even optimal in terms of adjusted P&L.

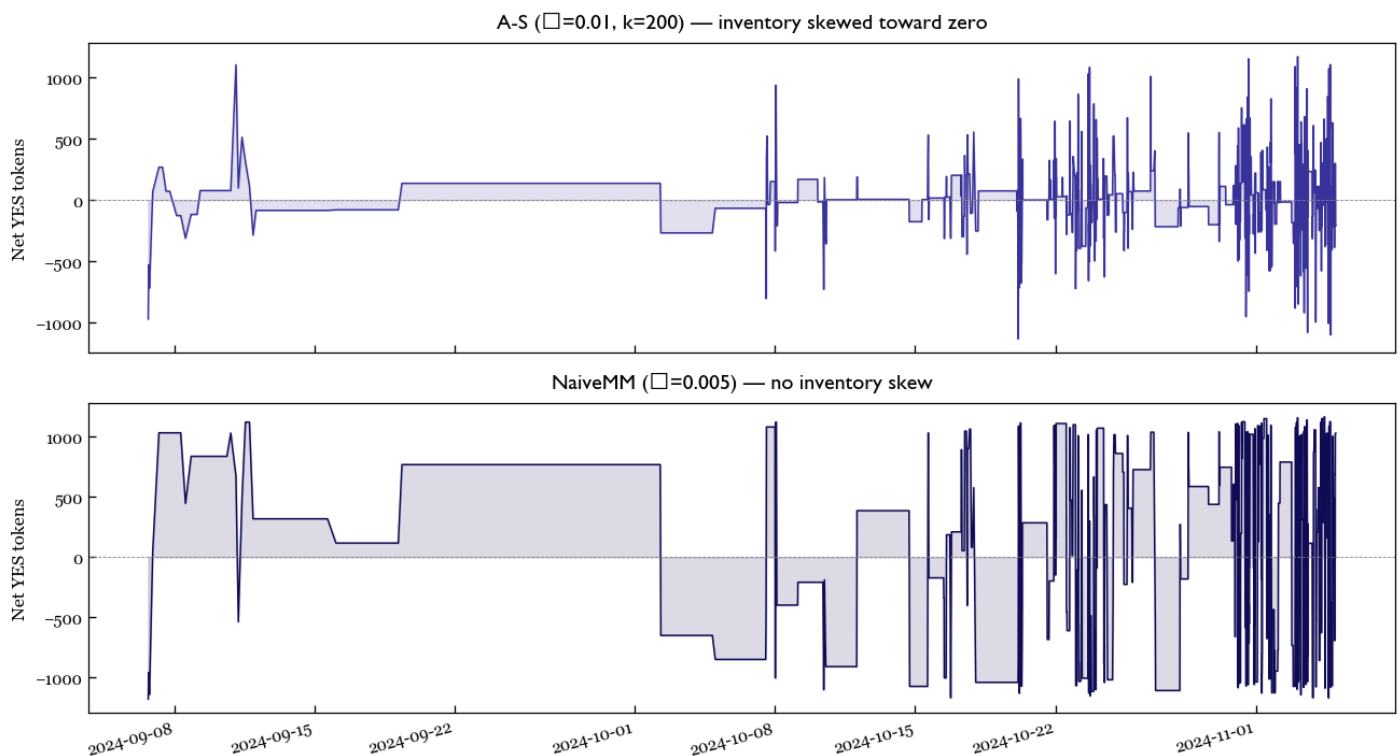
At the same time, the Avellaneda–Stoikov model continues to demonstrate its value in terms of risk management. Across the parameter grid, A-S configurations consistently exhibit lower maximum drawdowns and more stable inventory paths compared to the naive benchmark. This highlights an important distinction: while naive market making can match or exceed A-S in raw returns under certain conditions, the A-S framework provides a more disciplined approach to controlling risk and avoiding extreme inventory imbalances.

Another important finding is related to markout. Contrary to the standard concern that market makers are systematically adversely selected, our results show that markout is often positive in this dataset. This implies that fills are more frequently driven by short-term noise rather than informed trading, and that prices tend to revert after trades. In other words, passive liquidity provision in the Polymarket election markets benefits from mean reversion rather than being consistently exploited by better-informed participants. This result is consistent with the broader microstructure finding that retail-driven flow dominates these markets.

All the views expressed are opinions of Bocconi Students Investment Club members and can in no way be associated with Bocconi University. All the financial recommendations offered are for educational purposes only. Bocconi Students Investment Club declines any responsibility for eventual losses you may incur implementing all or part of the ideas contained in this website. The Bocconi Students Investment Club is not authorised to give investment advice. Information, opinions, and estimates contained in this report reflect a judgment at its original date of publication by Bocconi Students Investment Club and are subject to change without notice. The price, value of and income from any of the securities or financial instruments mentioned in this report can fall as well as rise. Bocconi Students Investment Club does not receive compensation and has no business relationship with any mentioned company.

Overall, Week 5 establishes a more credible evaluation framework for market-making strategies. By incorporating realistic execution frictions and systematically exploring the parameter space, we move from an idealized proof-of-concept to a defensible assessment of strategy performance. The main conclusion is that both A-S and naive market makers remain profitable under realistic assumptions, but their relative strengths differ: naive strategies can perform well in terms of raw returns, while the Avellaneda–Stoikov model provides superior control over inventory risk and drawdowns.

Inventory Paths — Trump Market (Sep–Nov 2024)



Source: Polymarket trades_election_2024 dataset, Sep–Nov 2024

Directional Strategies: Mean Reversion and Wallet Copying

In Week 6, the focus of the project shifted from liquidity provision to testing whether simple, mechanically defined trading strategies could extract directional alpha from the Polymarket election markets. Unlike the market-making approaches in Weeks 4 and 5, which relied on limit orders and earned the bid–ask spread, these strategies used market orders exclusively and therefore paid the spread and transaction costs. This distinction is important: whereas market makers benefit from providing liquidity, these “edge strategies” attempt to profit from predicting price movements, making them a stricter test of whether exploitable signals exist in the data.

All the views expressed are opinions of Bocconi Students Investment Club members and can in no way be associated with Bocconi University. All the financial recommendations offered are for educational purposes only. Bocconi Students Investment Club declines any responsibility for eventual losses you may incur implementing all or part of the ideas contained in this website. The Bocconi Students Investment Club is not authorised to give investment advice. Information, opinions, and estimates contained in this report reflect a judgment at its original date of publication by Bocconi Students Investment Club and are subject to change without notice. The price, value of and income from any of the securities or financial instruments mentioned in this report can fall as well as rise. Bocconi Students Investment Club does not receive compensation and has no business relationship with any mentioned company.

The first strategy, the mean-reversion strategy, was based on the idea that large trades might temporarily push prices away from their equilibrium levels, creating short-term opportunities for contrarian trading. The implementation detects large individual trades using a USD notional threshold, set at \$500 by default, and then takes the opposite position via a market order. Specifically, the strategy goes short after a large YES buy and goes long after a large NO buy, with the expectation that prices will revert toward their previous levels. Positions are managed using simple exit rules: a take-profit condition is triggered if the mid-price moves back in the expected direction by a predefined amount, while a stop-loss is triggered if the price continues moving against the position. To maintain simplicity and control risk, the strategy allows only one open position at a time.

The second strategy, wallet copying, was designed to exploit potential informational advantages held by certain traders. Building on the wallet classification developed earlier in the project, we identified a subset of “Directional Bettor” wallets, specifically participants who consistently trade in a particular direction, and attempted to replicate their trades. The strategy observes trades made by these wallets and, after a configurable delay, executes a market order in the same direction. The delay reflects realistic latency in observing and reacting to on-chain activity. Unlike the mean-reversion strategy, the wallet-copying strategy does not explicitly close positions; instead, it accumulates exposure over time, effectively betting that these traders possess persistent informational advantages.

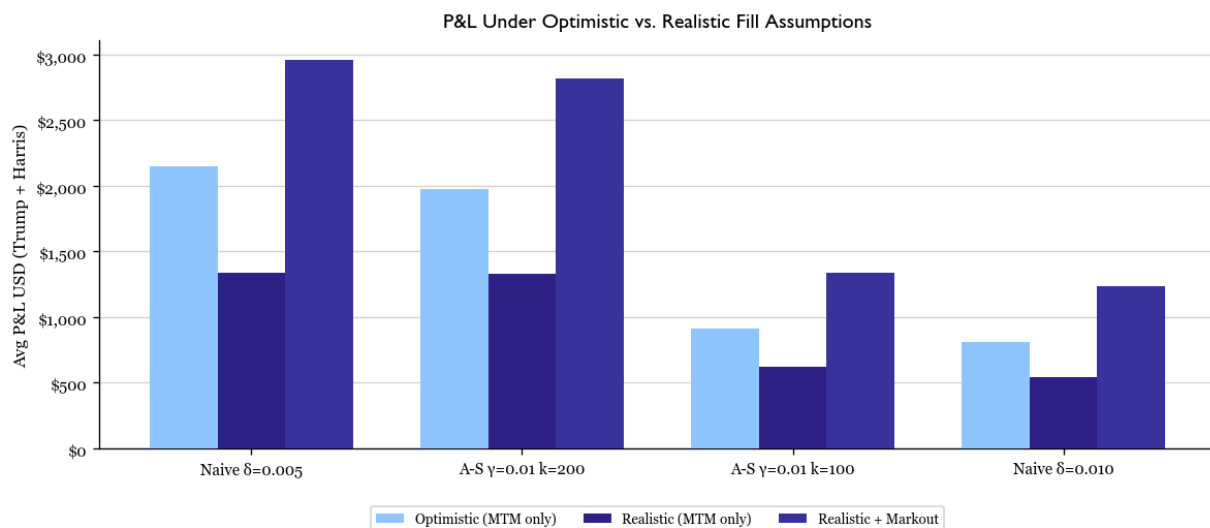
Both strategies were evaluated using the same backtesting framework as the market-making models, ensuring comparability. We conducted a parameter sweep across twelve configurations for each strategy, varying parameters such as trade size thresholds, lag times, and exit conditions. These configurations were tested on both the Trump and Harris markets, resulting in twenty-four total backtests. The results were generated using a standalone script and stored in a structured dataset, allowing for fast analysis and direct comparison with the market-making benchmarks.

The results of this analysis are unambiguous: both edge strategies underperform the market-making strategies. The mean-reversion strategy generates near-zero profits across most configurations, indicating that large trades do not, in fact, create temporary price dislocations that revert quickly. Instead, the data shows that large trades tend to be followed by continued price movement in the same direction. This finding is consistent with the positive Kyle’s λ estimates obtained in Week 3, which indicate that order flow has a persistent impact on prices. In this environment, contrarian trading is systematically disadvantaged. The underlying mechanism can be understood through three interconnected observations. The market is dominated by retail participants whose trades carry limited private information; this generates the noise-driven flow that allows passive market makers to earn positive markout rather than being adversely selected. That same positive markout is evidence of short-term price reversion around individual trades, which might seem to favour a mean-reversion strategy. However, the Kyle framework identifies positive and persistent aggregate price impact: the market maker’s price update is not a temporary dislocation but a genuine update to the consensus probability. Large trades, in particular, tend to reflect either informed conviction or coordinated directional flow, and both produce continuation rather than reversal. The failure of mean reversion is therefore consistent with the positive and persistent price impact identified in the Kyle framework: the reversion observed at the individual-fill level does not survive at the strategy level once the cost of crossing the spread is taken into account.

All the views expressed are opinions of Bocconi Students Investment Club members and can in no way be associated with Bocconi University. All the financial recommendations offered are for educational purposes only. Bocconi Students Investment Club declines any responsibility for eventual losses you may incur implementing all or part of the ideas contained in this website. The Bocconi Students Investment Club is not authorised to give investment advice. Information, opinions, and estimates contained in this report reflect a judgment at its original date of publication by Bocconi Students Investment Club and are subject to change without notice. The price, value of and income from any of the securities or financial instruments mentioned in this report can fall as well as rise. Bocconi Students Investment Club does not receive compensation and has no business relationship with any mentioned company.

The wallet-copying strategy performs even worse. Although it generates a high volume of trades, it consistently loses money across configurations. The primary reason is the cost of execution. Because the strategy relies on market orders, it pays the bid-ask spread and incurs slippage on every trade. While the directional wallets being copied may exhibit some predictive power, this signal is too weak to overcome the transaction costs. As a result, the strategy effectively transfers value to liquidity providers rather than capturing it.

Taken together, these results reinforce a central insight of the project: in this dataset, providing liquidity is fundamentally more profitable than consuming it. The edge strategies fail not because they are poorly designed, but because the market environment does not support their underlying assumptions. Price movements are driven more by short-term noise and gradual information incorporation than by predictable reversals or easily exploitable signals.



Source: Polymarket trades_election_2024 dataset, Sep–Nov 2024

Conclusion

This project examined the microstructure of Polymarket’s 2024 US Presidential Election markets and evaluated a range of trading strategies, from passive market making to active directional approaches. Across the different stages of analysis, several consistent and defensible conclusions emerge.

A key early finding is that trader classification must be handled with care. Initial attempts to classify traders based on ex-post realized profits led to misleading results, including incorrect signs in the estimation of Kyle’s λ . This was due to look-ahead bias, where future information was inadvertently incorporated into the classification. By shifting to a purely behavioural classification based on observable trading patterns, we obtained more reliable and theoretically consistent estimates of price impact across trader types.

All the views expressed are opinions of Bocconi Students Investment Club members and can in no way be associated with Bocconi University. All the financial recommendations offered are for educational purposes only. Bocconi Students Investment Club declines any responsibility for eventual losses you may incur implementing all or part of the ideas contained in this website. The Bocconi Students Investment Club is not authorised to give investment advice. Information, opinions, and estimates contained in this report reflect a judgment at its original date of publication by Bocconi Students Investment Club and are subject to change without notice. The price, value of and income from any of the securities or financial instruments mentioned in this report can fall as well as rise. Bocconi Students Investment Club does not receive compensation and has no business relationship with any mentioned company.

The project also highlights the importance of correctly specifying order flow in empirical models. The use of a signed-log transformation introduced a mismatch in units, effectively turning a linear price-impact model into an elasticity model. This significantly reduced explanatory power. Reverting to linear, winsorised order flow restored the model's performance, increasing R^2 substantially on the most liquid markets. This underscores the importance of aligning empirical implementation with theoretical foundations.

Another important insight is that the relationship between order flow and price changes is highly regime-dependent. When the analysis is restricted to the period from September to early November 2024, characterised by intense news flow and heightened trading activity, the explanatory power of the Kyle's λ model increases dramatically. This confirms that prediction markets are strongly influenced by external information events, and that microstructure dynamics are most pronounced during such periods.

The implementation of the Avellaneda–Stoikov market-making model demonstrates that its theoretical mechanisms translate effectively into practice. The model dynamically adjusts quotes in response to inventory and volatility, and across all tested configurations it respects inventory limits and behaves as expected. While its advantage over a naive fixed-spread strategy in terms of raw returns is modest under realistic conditions, it consistently delivers better risk management, with lower drawdowns and more stable inventory profiles.

The most robust result of the project is that passive market making is consistently profitable in these markets. When realistic trading frictions are introduced, including queue position, fees, and markout costs, both the naive and Avellaneda–Stoikov strategies generate positive adjusted returns across all configurations. Moreover, the markout analysis reveals that adverse selection is limited. In fact, market makers are more often filled by short-term noise than by informed traders, resulting in favourable post-trade price movements. This finding suggests that the market is dominated by retail-driven flow rather than by strong informational asymmetries.

The comparison between optimistic and realistic execution models further emphasizes the importance of proper backtesting assumptions. Under optimistic conditions, aggressive quoting strategies with very tight spreads appear highly profitable. However, once realistic frictions are introduced, these strategies lose their advantage, and simpler approaches with slightly wider spreads become competitive or superior. This demonstrates that naive backtests can produce misleading conclusions if they fail to account for execution constraints.

Finally, the failure of the edge strategies reinforces the broader narrative. Strategies that attempt to predict price movements, whether through mean reversion or by following specific traders, do not outperform once transaction costs are taken into account. Large trades tend to signal continuation rather than reversal, and the informational advantage of individual wallets is insufficient to overcome the cost of crossing the spread. As a result, active trading strategies underperform even simple market-making benchmarks.

Overall, the evidence supports a clear conclusion: during the 2024 election cycle, Polymarket behaved as a market where liquidity provision was structurally rewarded, while directional trading was challenging to execute profitably. The most effective strategy was not to predict the outcome of the election, but to exploit the microstructure of the market itself by earning the bid–ask spread and managing inventory risk.

All the views expressed are opinions of Bocconi Students Investment Club members and can in no way be associated with Bocconi University. All the financial recommendations offered are for educational purposes only. Bocconi Students Investment Club declines any responsibility for eventual losses you may incur implementing all or part of the ideas contained in this website. The Bocconi Students Investment Club is not authorised to give investment advice. Information, opinions, and estimates contained in this report reflect a judgment at its original date of publication by Bocconi Students Investment Club and are subject to change without notice. The price, value of and income from any of the securities or financial instruments mentioned in this report can fall as well as rise. Bocconi Students Investment Club does not receive compensation and has no business relationship with any mentioned company.

Limitations and Future Work

Despite the robustness of the results, several limitations remain. The queue model used in the backtester is a simplified Bernoulli approximation and does not fully capture order book depth or priority dynamics. The use of a one-hour markout window may miss shorter-term adverse selection effects due to data limitations. The analysis focuses only on the most liquid markets and excludes low-probability contracts, where boundary effects complicate modelling. Additionally, the study is limited to a single election cycle, which restricts the ability to generalize the findings across different market conditions.

Future work could address these limitations by incorporating higher-frequency data, developing more realistic order book simulations, and extending the analysis to other prediction markets and time periods. Improvements in trader classification and more precise identification of informed trading would also enhance the understanding of market dynamics. Nonetheless, within the scope of the available data and methods, the results presented here provide a coherent and well-supported account of how liquidity, information, and trading strategies interact in prediction markets.

References

- [1] Kyle, A. S., “Continuous Auctions and Insider Trading”, *Econometrica*, Vol. 53, No. 6, 1985
- [2] Hasbrouck, J., “Measuring the Information Content of Stock Trades”, *Journal of Finance*, Vol. 46, No. 1, 1991
- [3] Avellaneda, M. and Stoikov, S., “High-frequency trading in a limit order book”, *Quantitative Finance*, Vol. 8, No. 3, 2008
- [4] Lloyd, S. P., “Least Squares Quantization in PCM”, *IEEE Transactions on Information Theory*, Vol. 28, No. 2, 1982
- [5] Newey, W. K. and West, K. D., “A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix”, *Econometrica*, Vol. 55, No. 3, 1987
- [6] SII-WANGZJ, “Polymarket_data”, HuggingFace repository, 2024. Available at: https://huggingface.co/datasets/SII-WANGZJ/Polymarket_data

TAGS: Polymarket, Prediction Markets, Market Microstructure, Market Making, Avellaneda-Stoikov, Kyle’s Lambda, Order Flow Imbalance, Algorithmic Trading, Decentralised Finance, Blockchain

All the views expressed are opinions of Bocconi Students Investment Club members and can in no way be associated with Bocconi University. All the financial recommendations offered are for educational purposes only. Bocconi Students Investment Club declines any responsibility for eventual losses you may incur implementing all or part of the ideas contained in this website. The Bocconi Students Investment Club is not authorised to give investment advice. Information, opinions, and estimates contained in this report reflect a judgment at its original date of publication by Bocconi Students Investment Club and are subject to change without notice. The price, value of and income from any of the securities or financial instruments mentioned in this report can fall as well as rise. Bocconi Students Investment Club does not receive compensation and has no business relationship with any mentioned company.