The more you know, the more you dare®

Master in International Finance

Do Prediction Markets Misprice Equity Threshold Events? The Role of Liquidity in the Polymarket–Options Gap

Authors	Ferran Capdevila Oliver Marc Tabernero Riera
Supervisor	Daniel SCHMIDT
Programme	Master in International Finance
Academic Year	2025–2026
Submission Date	1st June, 2026
Keywords	Polymarket; prediction markets; options; liquidity; mispricing; threshold events

Ferran Capdevila Oliver. Perú Street 17, 08195, Sant Cugat del Vallès, Spain. Tel: +34 667 409 941.
Email: ferran.capdevila@hec.edu

Marc Tabernero Riera. Torrent de Can Caralleu Street 24, 08017, Barcelona, Spain. Tel: +34 628 524 637.
Email: marc.tabernero-riera@hec.edu

Do Prediction Markets Misprice Equity Threshold Events? The Role of Liquidity in the Polymarket–Options Gap

Ferran Capdevila Oliver and Marc Tabernero Riera

Abstract

Prediction markets have grown rapidly, but it remains unclear whether their prices align with option-implied benchmarks. This paper compares Polymarket equity-threshold prices with option-implied probabilities for eight large-cap US equities between November 2025 and March 2026, using 5,376 contract-day observations and two-way fixed effects regressions. We find that disagreement widens non-linearly as contracts approach resolution, especially in the final three days, while higher Polymarket trading activity is associated with smaller gaps at longer horizons. We interpret mispricing as relative divergence from a risk-neutral options benchmark, rather than proof that either market is objectively mispriced.[1]

[1]The authors used large language model tools, specifically ChatGPT and Claude, to assist with editing and improving some sentences in this paper. All analysis, data collection, empirical design, and substantive content are the authors' own work.

Contents

1. Introduction

Prediction markets have grown rapidly in recent years. Polymarket, the world's largest decentralised prediction market platform, recorded trading volume of approximately \$9 billion in 2024, up from just \$73 million in 2023 — a 120-fold increase in a single year. Combined volume across the two leading platforms, Polymarket and Kalshi, approached \$40 billion in 2025.[1] This rapid expansion has attracted growing interest from academics, practitioners, and regulators alike, raising the question: how do these market-implied probabilities compare with benchmarks derived from established financial markets such as options markets?

A natural benchmark for evaluating this question is the options market. For any equity threshold event (such as whether a given stock will close above a specified price on a given date) a call spread on the underlying asset provides a theoretically grounded estimate of the risk-neutral probability of that event, derived directly from no-arbitrage principles (Breeden and Litzenberger, 1978). If prediction-market prices and option-implied benchmarks reflected the same information under similar pricing conditions, Polymarket contract prices and option-implied probabilities would be expected to align closely. Systematic divergences, by contrast, may reflect differences in market microstructure, frictions that prevent arbitrage from closing the gap, or participants with different information sets, although none of these mechanisms is directly tested in this paper. Understanding the drivers of these divergences is the central question of this paper.

Given this benchmark, the term "mispricing" is used in a relative empirical sense throughout the paper. We do not claim to observe the true physical probability of each event, nor do we test whether either market is objectively wrong. Instead, we study whether Polymarket prices diverge systematically from a feasible option-implied, risk-neutral benchmark constructed from listed option prices.

We address this question using a panel of daily observations on equity threshold contracts listed on Polymarket over the period November 2025 to March 2026, covering eight large-cap US equities: Apple, Alphabet, Microsoft, Meta, Tesla, NVIDIA, Netflix, and Palantir. For each contract-day observation, we construct an option-implied probability using the finite-difference approximation of Breeden and Litzenberger (1978) applied to call option prices retrieved from LSEG Workspace, and compare it to the corresponding Polymarket contract price. Our main dependent variable is the absolute difference between the two measures in logit space. We relate this gap to proxies for liquidity on both markets, time-to-resolution, and contract moneyness, using a two-way fixed effects panel regression with contract and date fixed effects and standard errors clustered at the contract level. The panel contains 5,376 contract-day observations across 971 unique contracts.

We document three main patterns. First, time-to-resolution is the strongest predictor of the gap: as contracts approach their resolution date, the disagreement between Polymarket and option-implied probabilities widens substantially. The effect is non-linear and concentrated in the final days before expiry, with contracts in the 0–3 day range exhibiting a gap that is, on average, approximately one logit

[1] Volume figures are sourced from Sacra (2025) and Phemex News (2025). Note that Polymarket volume metrics may vary depending on the measurement methodology; Paradigm (2025) document that some dashboards double-count trades by summing both sides of each transaction, which can inflate reported figures by approximately a factor of two.

unit larger than contracts with more than 15 days remaining. This finding is stable across a wide range of specifications, fixed effects structures, and sample definitions. Second, higher trading activity on Polymarket is associated with a smaller gap, particularly for longer-horizon contracts. The interaction between volume and time-to-resolution suggests that the gap-reducing association of trading activity is stronger when sufficient time remains before resolution. The option-side liquidity result is weaker. The relative bid-ask spread of the options used to construct the benchmark is negatively associated with the gap in the main specification, but this result is sensitive to the treatment of boundary observations and does not hold in the alternative sample where extreme probabilities are removed. We therefore treat this finding as secondary rather than as a core result of the paper.

To our knowledge, this is the first systematic panel comparison of Polymarket equity-threshold probabilities with individual stock option-implied probabilities, based on a purpose-built dataset that matches prediction-market contracts to listed option quotes at the same strike and expiry. A recent project by researchers at Stevens Institute of Technology applies a related methodology to Kalshi contracts on the S&P 500 index (Stevens Financial Strategy Club, 2025), but focuses on index-level contracts rather than individual equities and does not examine the role of liquidity frictions in explaining discrepancies. Our paper extends this line of work to individual equity contracts and examines how liquidity and time-to-resolution are related to divergences between Polymarket prices and an option-implied benchmark. This allows us to study whether the gap between the two markets varies systematically over the life of a contract.

The paper connects to several strands of the finance literature. It relates to work on prediction markets and the interpretation of market prices as probabilities (Wolfers and Zitzewitz, 2006; Manski, 2006), and to evidence on liquidity frictions and limits to arbitrage (Shleifer and Vishny, 1997; Tetlock, 2008). On the options side, it builds on the extraction of risk-neutral probabilities from option prices (Breeden and Litzenberger, 1978; Jackwerth and Rubinstein, 1996) and the no-arbitrage foundations of derivatives pricing (Merton, 1973).

The remainder of the paper is organised as follows. Section 2 reviews the related literature. Section 3 describes the data and sample construction, including the matching of individual equity option chains to Polymarket contracts at identical strike levels and expiry dates. Section 4 presents the probability measures and empirical design. Section 5 reports the empirical results. Section 6 concludes.

2. Literature Review

This paper lies at the intersection of two distinct literatures: prediction markets as mechanisms for probability aggregation, and the extraction of risk-neutral probabilities from option prices. We review the key contributions in each strand and identify the gap that motivates our analysis.

2.1 Prediction Markets as Probability Aggregation Mechanisms

A key result in the prediction market literature is that contract prices can be read as market-implied probabilities, as long as certain assumptions about risk aversion and differences in beliefs hold (Wolfers and Zitzewitz, 2006). However, this interpretation is not exact. Manski (2006) shows that the link between prices and probabilities depends on traders' risk preferences and beliefs, meaning that deviations from risk neutrality can introduce systematic biases in how prices reflect probabilities. Together, these contributions establish both the theoretical appeal of prediction market prices as probability estimates and the conditions under which that interpretation breaks down.

One well-documented breakdown is the *favourite-longshot bias*: low-probability contracts tend to be systematically overpriced relative to their realised frequencies. Snowberg and Wolfers (2010) provide evidence that this bias is driven by probability misperceptions consistent with prospect theory (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992), making it a structural feature rather than a liquidity-driven effect. For our purposes, this is relevant because it implies that prediction market prices near the extremes of the probability distribution may diverge from option-implied probabilities even in the absence of liquidity frictions; a point that motivates how we handle boundary observations described in Section 4.

The role of liquidity in prediction market efficiency provides a second motivation for our empirical design. Tetlock (2008) finds, in the context of political prediction markets, that mispricing is more prevalent in contracts with lower trading activity, a result consistent with the limits-to-arbitrage framework of Shleifer and Vishny (1997). If arbitrage between Polymarket and options markets is constrained by trading frictions on either side, the gap between the two probability measures need not be eliminated even when it is observable. This provides the mechanism through which liquidity enters our empirical analysis as a key explanatory variable.

More recent empirical work has looked directly at Polymarket. Reichenbach and Walther (2025) analyse over 124 million trades and find that, on average, prices line up well with realised probabilities, with no strong evidence of a longshot bias. They do, however, document some tendency to overtrade the default and "Yes" outcomes.

Akey et al. (2026) study the complete transaction history of Polymarket from 2022 to 2026, covering over 2.4 million users and $67 billion in trading volume. They find that prices are well-calibrated on aggregate, but that low-volume markets exhibit systematic S-shaped deviations from perfect calibration. This finding is directly relevant to our setting: it suggests that the degree to which Polymarket prices track a risk-neutral benchmark may depend on the level of trading activity in a given contract, which is one of the central hypotheses we test. Akey et al. (2026) also document that traders at extreme prices (below 10 cents or above 90 cents) account for a disproportionate share of losses, a pattern consistent with the favourite-longshot bias and relevant to how we handle boundary observations in Section 4.

026 compare Polymarket, Kalshi, PredictIt, and Robinhood during the 2024 US election and find that more liquid prediction markets produce more accurate prices, and that Polymarket leads Kalshi in price discovery. This cross-platform evidence reinforces the view that liquidity is a key determinant of how closely prediction-market prices track the true probability of an event. Gomez Cram et al. (2026) add a further layer of nuance: using the complete Polymarket transaction history from 2023 to 2025, they find that roughly 3% of traders account for the bulk of price discovery, while the remaining majority provides volume but little information. This suggests that aggregate calibration, as documented by Reichenbach and Walther (2025) and Akey et al. (2026), may mask substantial heterogeneity at the contract level: contracts with limited participation may lack the informed traders needed to keep prices aligned with any external benchmark.

For our purposes, this suggests that Polymarket prices contain useful information about event probabilities. However, being informative does not necessarily mean they are fully efficient when compared to other benchmarks, such as option-implied probabilities derived from no-arbitrage pricing.

The broader accuracy of prediction markets across settings is documented by Berg et al. (2008), who survey a decade of evidence from the Iowa Electronic Markets and find that market prices consistently outperform opinion polls as probability forecasts.

2.2 Option-Implied Probabilities and the Risk-Neutral Density

The extraction of risk-neutral probabilities from option prices is based on the framework of Breeden and Litzenberger (1978), who show that the cross-section of European call prices contains the full risk-neutral distribution of the underlying asset. In practical terms, the probability that the asset ends above a given threshold can be approximated from the slope of the call price curve at that point, typically using a call spread. This result forms the basis for how we construct our option-implied probability measure.

Later empirical work has built on this framework, improving both the methods and how they are applied in practice. Jackwerth and Rubinstein (1996) apply it to S&P 500 options and find clear departures from the usual lognormal assumption in the implied distributions. Aït-Sahalia and Lo (1998) introduce nonparametric methods to deal with sparse strike data, and Figlewski (2010) offers practical guidance on implementation choices, including how to handle boundary observations which is an issue directly relevant in our setting.

One implementation consideration is the use of American-style options. For call options on non-dividend-paying stocks, early exercise is never optimal (Merton, 1973). The equities in our sample either pay no dividends (TSLA, NVDA, META, GOOGL, NFLX, PLTR) or have annual dividend yields below 0.7% (AAPL, MSFT), making the early exercise premium economically negligible for the short-dated contracts we consider. Our sample consists exclusively of contracts with a maximum of 32 calendar days to expiry on large-cap US equities with low dividend yields, and we follow the standard practice of treating observed call prices as European for the purpose of the Breeden–Litzenberger approximation.

2.3 The Intersection: Prediction Markets and Option-Implied Probabilities

Despite the parallel development of these two literatures, they have rarely been brought together in a single empirical framework, as there is limited evidence comparing prediction market probabilities with option-implied probabilities for individual equity threshold events. Wolfers and Zitzewitz (2006) briefly suggest that options markets could provide a natural benchmark for prediction market prices, but do not explore this idea empirically, as their focus is mainly on political and macroeconomic prediction markets.

The closest antecedent to our paper is a recent project by researchers at Stevens Institute of Technology (Stevens Financial Strategy Club, 2025), which applies the Breeden–Litzenberger framework to compare Kalshi event contract prices with SPX option-implied probabilities at the index level and documents systematic differences between the two measures, particularly in low-liquidity environments. While methodologically related, that study focuses on aggregate index-level contracts, does not examine individual equity threshold events, and does not use panel regression methods to isolate the separate roles of liquidity and time-to-resolution.

A closely related contribution is Gomez Cram et al. (2025), who use Polymarket's earnings beat contracts to compare prediction-market prices with analyst forecasts for the same underlying event. They find that Polymarket prices are more accurate than analyst consensus one day before earnings announcements, and that prediction-market implied expectations contain incremental information for stock returns around earnings dates. While their benchmark is analyst forecasts rather than option-implied probabilities, their approach shares the core logic of our paper: using an established financial market measure to evaluate the informational content of Polymarket prices for equity-related events. Our paper extends this comparison to option-implied risk-neutral probabilities and examines how the gap between the two measures varies with liquidity and time to resolution.

Our paper addresses these gaps in three ways. First, we construct a matched panel dataset comparing Polymarket equity-threshold probabilities with option-implied probabilities from individual stock options across eight underlying equities and five months of daily observations. Second, we use two-way fixed effects panel regressions to isolate the roles of Polymarket liquidity, options market liquidity, time-to-resolution, and contract moneyness. Third, we find that liquidity and time-to-resolution interact: the gap-narrowing effect of trading activity is strongest for contracts with more time remaining. This helps explain how the relative efficiency of the two markets changes over the life of a contract.

3. Data and Sample Construction

The analysis uses two main data sources. Polymarket trading data covering daily prices, trading volumes, and contract-level details, were collected through the platform's public API using a custom extraction script.[1] Option prices for the corresponding underlying equities were obtained from LSEG Workspace via the `@RDP.Data` Excel add-in. Together, these datasets provide market-based probabilities for the same equity threshold events, but from two different market environments, which is the core comparison of this study.

An important feature of the study is that the dataset had to be constructed from raw market-level sources rather than from a consolidated academic database. Polymarket equity-threshold contracts are not available in standard financial datasets, and the matching to listed equity options required a purpose-built data pipeline combining API-based extraction, contract-level metadata validation, and strike-by-strike retrieval from LSEG Workspace. The empirical contribution of the paper therefore lies not only in the regression analysis, but also in the construction of a matched Polymarket–options panel for a class of contracts that is not directly available in off-the-shelf financial databases.

The dataset covers equity threshold contracts on eight US equities: Apple (AAPL), Alphabet (GOOGL), Microsoft (MSFT), Meta (META), Tesla (TSLA), NVIDIA (NVDA), Netflix (NFLX), and Palantir (PLTR), over the period November 2025 to March 2026. These eight names represent all equities for which Polymarket listed equity threshold contracts during the sample period with sufficient coverage on both sides of the comparison. An initial universe of nine equities was identified, but Opendoor (OPEN) was excluded due to a substantially smaller number of contracts and limited availability of option quotes at the required bracket strikes.

The remainder of this section is organised as follows: Section 3.1 describes the Polymarket data and the sample construction on the prediction-market side, Section 3.2 covers the LSEG option data, and Section 3.3 describes the matching procedure and presents the final estimation sample.

3.1 Polymarket Data

Polymarket is a decentralised prediction market built on the Polygon blockchain where participants buy and sell binary contracts on the outcomes of real-world events. Each contract settles at $1 if the specified event occurs and at $0 otherwise, so its price can be read directly as a market-implied probability. Contracts span seven broad categories: Sports, Crypto, Politics, Finance, Tech, Culture, and Weather. Sports is the largest category by number of markets and accounts for 37.9% of total platform volume, while Politics generates 31.5% from a much smaller number of markets. Finance, Tech, Culture, and Weather together account for less than 12% of total volume (Akey et al., 2026), reflecting the more specialised nature of these contract types.

This paper focuses on equity threshold contracts, which fall within the Finance category and represent a recent addition to the platform. An equity threshold contract resolves based on whether a given stock closes above a specified price level K on a predetermined resolution date T, paying $1 if the stock closes

[1] The extraction codebase is documented in Appendix B.

above K and \$0 otherwise. Contracts are launched regularly across a set of large-cap US equities at multiple strike levels and resolution horizons, with new contracts typically appearing each week for the upcoming daily, weekly, and month-end expiries. Figure 1 shows a representative example of how these contracts appear on the platform, including the evolution of the implied probability over the life of the contract and the prevailing bid–ask spread.

The Polymarket-implied probability for contract i on date t, denoted $p_{i,t}^{PM}$, captures the market's assessment of that event. The next section describes how this measure is transformed into the dependent variable used in the regressions.

3.1.1 Sample period

An initial data collection focused on February and March 2026. Reviewing that extract showed that contracts of this type had first appeared on the platform in November 2025, and that restricting the sample to later months could not only induce selection bias but also be affected by month-specific circumstances. The extraction was therefore repeated from November 2025 onwards. The final sample spans five calendar months, from 4 November 2025 to 27 March 2026.

3.1.2 Contract types and universe

Contracts in the sample fall into three categories according to their resolution horizon. Daily contracts resolve at the end of a single trading day (e.g. *Will Microsoft (MSFT) close above \$490 on January 9?*). Weekly contracts resolve on the last trading day of a given week (e.g. *Will Microsoft (MSFT) finish the week of January 5 above \$490?*). Month-end contracts resolve at the end of a calendar month (e.g. *Will Microsoft (MSFT) close above \$490 at end of January?*). All contracts retained in the final sample resolve on a Friday, a restriction that comes from the matching requirement on the options side: since the matched option chain uses Friday expiries, only Friday-resolving Polymarket contracts can be paired with a same-date option quote.

New contracts are launched on a rolling basis, with fresh daily, weekly, and month-end contracts typically appearing at the start of each week across a set of strike levels around the prevailing stock price. The maximum time to resolution in the final sample is 32 calendar days, reflecting the short-dated nature of equity threshold contracts relative to other Polymarket categories. During the sample period, equity threshold contracts were listed for eight large-cap US equities: Apple (AAPL), Alphabet (GOOGL), Microsoft (MSFT), Meta (META), Tesla (TSLA), NVIDIA (NVDA), Netflix (NFLX), and Palantir (PLTR). These represent all equities for which Polymarket listed equity threshold contracts with sufficient coverage on both sides of the comparison during the sample period.

Contract resolution is determined by reference to the official closing price of the underlying stock on the resolution date, as reported by standard market data providers and verified through Polymarket's decentralised oracle mechanism. A contract resolves to \$1 if the stock closes strictly above the threshold K on date T, and to \$0 otherwise. Once the outcome is confirmed by the oracle, positions are settled automatically on the blockchain and funds are distributed to holders of the winning side.

3.1.3 Data extraction

Because Polymarket does not provide a standardised downloadable dataset for equity-threshold contracts, the extraction proceeded through an iterative two-stage API pipeline. In the first, the **Gamma API** was used to identify the universe of eligible markets. A set of 13,878 candidate slugs was generated by combining the eight target tickers with candidate strike levels and resolution dates; probing these against the API returned 3,874 live or resolved equity threshold markets. For each market, the API provided a unique identifier (`conditionId`), the underlying ticker, the strike threshold, the resolution date, and the contract type.

In the second stage, the **Polymarket Data API** was queried contract by contract to retrieve the full transaction history for each of the 3,874 markets. A total of 251,401 individual trades were downloaded, each record containing a timestamp, a transaction price, and a traded quantity. These records were aggregated to the daily level by computing, for each contract-day, the volume-weighted average price (VWAP), total dollar volume, and number of transactions. The resulting raw daily panel was then passed through the cleaning pipeline described below.

This extraction process required repeated validation of contract identifiers, strike thresholds, resolution dates, and contract types, since Polymarket markets are identified through platform-specific slugs rather than through standardised financial identifiers. This is particularly relevant for the present study because the subsequent matching to options data depends mechanically on the correct identification of the underlying ticker, strike, and resolution date for each contract.

3.1.4 Data cleaning

The raw panel was cleaned through a sequential series of restrictions, summarised in Table 1 in the Tables section. Starting from 19,127 observations, the pipeline proceeds as follows. Observations on weekends and US market holidays are removed first (3,685 rows), aligning the panel with the options market trading calendar.[2] A small number of rows with missing essential fields are then removed, and a unique observation per contract-day is enforced by retaining the record with the highest trading volume where duplicates exist (1,990 rows removed across both steps).

Observations outside the active life of each contract are then removed. The active period is defined from the first recorded trade to the contract's resolution date. Listing dates were not consistently available in the Gamma API, so we use the first observed trade as the starting point. This is a reasonable approximation, as there is no price information before the first transaction and the contract has no real market activity prior to that.

Two additional filters come from the construction of the option-based measure. First, we remove contracts whose strike is not a multiple of five, since computing the option-implied probability requires call prices at $K - 5$ and $K + 5$, which only exist when K aligns with the discrete strike grid in the LSEG option chain (this removes 1,143 observations across 263 contracts). Second, we exclude contracts that do not resolve on a Friday. After these steps, the sample is reduced from 3,371 to 2,165 unique contracts and from 12,225 to **9,097 contract-day observations**.

[2] The main holidays within the sample window are Thanksgiving (27 November 2025), Christmas (25 December 2025), New Year's Day (1 January 2026), and Presidents' Day (16 February 2026).

3.1.5 Panel expansion

Polymarket only records days with trading activity, so days without trades are missing from the raw panel. Since option-implied probabilities can change on days without Polymarket trading, dropping these days would condition the panel on trading activity and remove periods in which a stale Polymarket price is compared with an updated option-implied benchmark. To make the two series comparable, we expand the panel to include all eligible weekdays within each contract's active life, even if no trades occurred. This is done at the contract level, defined by ticker, strike, resolution date, and contract type, as contracts with the same basic characteristics but different types are distinct markets.

After this step, the panel increases from 9,097 to **11,857 contract-day observations**, with 2,760 no-trade rows added (about 23% of the total). For these added observations, trading variables such as volume and number of trades are set to zero. The Polymarket probability is carried forward from the last observed daily VWAP, so no-trade days use the most recent traded Polymarket price rather than an interpolated or model-imputed probability. Time-to-resolution is then recalculated for each added date.

As a result, the gap on no-trade days reflects a comparison between a stale Polymarket price and a continuously updated option-implied probability. The implications of this choice are examined in Section 5.3.1 using a no-trade indicator.

3.2 LSEG Option Data

Option prices were retrieved from LSEG Workspace using the `@RDP.Data` add-in. Unlike consolidated academic option databases, this source required observation-level retrieval of the relevant option quotes at the bracket strikes needed for each Polymarket contract-day. For each contract-day observation in the expanded Polymarket panel, bid and ask quotes were collected for the two call options with strikes $K_i - 5$ and $K_i + 5$, using the standard LSEG RIC format. The $\Delta = 5$ bracket is the narrowest strike interval that could be retrieved consistently from LSEG Workspace across the relevant tickers, strikes, and expiry dates.[3] The mid-price at each bracket strike was computed as the average of the bid and ask. The resulting option-based measure should be interpreted as a risk-neutral rather than a physical probability. The discount factor applied in the finite-difference approximation uses a fixed risk-free rate of $r = 3\%$ per annum; given the short maturities in the sample (maximum 32 calendar days), the sensitivity of results to this assumption is negligible.

A note on price synchronisation: Polymarket trades continuously, while LSEG option quotes reflect US market hours (9:30AM–4:00PM ET). The daily Polymarket VWAP used in this paper aggregates trades over the full 24-hour period, and may therefore include activity that is not aligned with the timing of option quotes. This mismatch is a limitation of working with daily data and is common in studies that compare 24/7 decentralised markets with exchange-traded instruments.

Because all retained Polymarket contracts resolve on a Friday, each matched option observation is taken from the Friday expiry coinciding with the contract resolution date T_i. No interpolation across expiry dates is required.

Observations for which LSEG returns missing bid or ask quotes at either bracket strike are excluded.

[3] For example, the call on Apple stock expiring 16 January 2026 with strike $230 is retrieved under the ticker `AAPLC20260116000230000.U`.

This occurs primarily when the required strikes are far from the current stock price, where option quote availability is limited and the call-spread benchmark cannot be constructed directly. Section 4 describes how the option-implied probability $\hat{p}_{i,t}^{OPT}$ is constructed from these quotes and how values that fall outside the valid range are handled in the estimation sample.

3.3 Matching and Final Sample

Each Polymarket contract-day observation is matched to the corresponding LSEG option data along three dimensions: underlying equity, strike K_i, and resolution date T_i. For each matched observation, the option-implied probability $\hat{p}_{i,t}^{OPT}$ is constructed from call mid-prices at $K_i - 5$ and $K_i + 5$ using the finite-difference approximation described in Section 4. Since all retained Polymarket contracts resolve on a Friday and listed equity options expire on Fridays, the option expiry coincides exactly with the Polymarket resolution date for every matched observation.

3.3.1 Sample losses and final estimation sample

The merge is applied to the full expanded panel of 11,857 observations and reduces the sample in two data-construction steps, together removing 5,190 observations, or 43.8 per cent of the expanded panel. These matching-stage losses are distinct from the 8 singleton observations later dropped automatically by the two-way fixed effects estimator.

First step: observations without a valid option-implied benchmark. The finite-difference approximation can yield values outside the unit interval $[0, 1]$ when the contract is far from at-the-money or close to expiry. These are precisely the cases in which a fixed bracket of $\Delta = 5$ is least able to approximate the local slope of the call-price function. In addition, the LSEG Workspace add-in does not systematically provide complete and reliable bid-ask information at all required bracket strikes when options are very deep in- or out-of-the-money. Observations with $\hat{p}_{i,t}^{OPT} \notin [0, 1]$ are therefore excluded because the finite-difference benchmark does not yield a valid probability for those contract-days.

This exclusion reflects a data-availability and benchmark-construction constraint rather than discretionary cleaning choice. Retaining these observations would require imposing interpolation, extrapolation, or smoothing assumptions on the option surface, which would introduce an additional layer of model dependence. Given the objective of maintaining a transparent comparison between observed Polymarket prices and observed option quotes, we restrict the estimation sample to observations for which the option-implied probability can be constructed directly from available LSEG quotes.

The number of such observations is economically meaningful: 4,836 observations, approximately 41% of the expanded panel, are removed at this stage. This high share is consistent with the short-dated nature of the contracts and with the concentration of observations in regions where the finite-difference approximation is mechanically less reliable. Observations whose $\hat{p}_{i,t}^{OPT}$ lies within $[0, 1]$ but outside $[0.01, 0.99]$ are retained and clipped before estimation.

Second step: missing LSEG quotes. A further 354 observations are excluded because LSEG returns a null bid quote at one or both bracket strikes, making the mid-price unavailable. This is a limitation of the option data source rather than a discretionary cleaning choice: without both bid and ask quotes at the two bracket strikes, the call-spread benchmark cannot be constructed directly. The issue is concentrated in contracts where the required option strikes are far from the prevailing stock price, especially for Netflix (NFLX), where the bid side of the option market is frequently absent for the relevant strikes.

After these two data-construction steps, the raw matched dataset contains **6,667 contract-day observations** across **975 unique contracts**. Prior to estimation, we remove 1,283 exact duplicate rows, observations with identical values across all variables introduced during dataset assembly, leaving 5,384 unique observations. The final estimation sample contains **5,376 observations**, because eight observations form singleton groups under the two-way fixed effects structure and are dropped automatically by the estimation procedure. The total reduction from the expanded panel to the final estimation sample is therefore 6,481 observations, of which 5,190 are matching-stage losses, 1,283 are duplicate rows removed prior to estimation, and 8 are singletons dropped by the estimator.

Two samples are used in the analysis. The **full sample** ($N = 5,376$) is used for all main specifications. A smaller **restricted liquidity subsample** ($N = 3,301$) is used only in intermediate checks involving the Amihud price-impact measure, which is unavailable for the first observation of each contract and on no-trade days. Since this measure is not included in the preferred specification, the full sample remains the main estimation sample throughout the paper.

The composition of the final matched estimation sample by underlying equity is reported in Table 2 in the Tables section. The low representation of NFLX (110 observations, 2.0 per cent of the sample) is consistent with the higher rate of missing bid quotes, and the final sample representation varies across tickers because of contract availability on Polymarket. Table 2 also reports the no-trade fraction by ticker. The full-sample no-trade share is 21.4%, with substantial variation across names. NFLX has the highest no-trade fraction at 45.5%, consistent with its limited Polymarket trading activity and small representation in the final matched sample.

Table 3 summarises the contract-level characteristics of the final estimation sample. The average contract is observed for 5.5 contract-days (median 4), reflecting the short-dated nature of most equity threshold contracts. Contract active lives span from under one day to approximately 32 calendar days, with 57.1 per cent of contracts having an active life of two to five days. Month-end contracts account for 53.3 per cent of contract-day observations, weekly contracts for 42.0 per cent, and daily contracts for 4.7 per cent. In terms of moneyness, 33.0 per cent of observations are near at-the-money ($|\ln(K/S)| < 0.05$), 45.4 per cent are moderately in- or out-of-the-money, and 21.6 per cent are deep in- or out-of-the-money ($|\ln(K/S)| \geq 0.15$). Of the 5,376 observations, 1,148 (21.4 per cent) are no-trade days inserted during panel expansion.

4. Probability Measures and Empirical Design

This section describes how the two market-based probability measures are constructed from the raw data, defines the dependent and explanatory variables used in the regressions, and sets out the econometric framework. Section 4.1 covers the construction of the Polymarket and option-implied probabilities. Section 4.2 defines the gap measures and the transformations applied before estimation. Section 4.3 describes the explanatory variables. Section 4.4 presents the baseline regression specification, the research hypotheses, and the robustness strategy.

4.1 Construction of Probability Measures

4.1.1 Polymarket probability

The Polymarket-implied probability $p^{PM}_{i,t}$ for contract i on date t is defined as the volume-weighted average price of all transactions in that contract-day:

$$p^{PM}_{i,t} = \frac{\sum_{j \in \mathcal{J}_{i,t}} P_{i,t,j} \cdot Q_{i,t,j}}{\sum_{j \in \mathcal{J}_{i,t}} Q_{i,t,j}} \tag{1}$$

where $P_{i,t,j}$ is the price and $Q_{i,t,j}$ the quantity of trade j, and $\mathcal{J}_{i,t}$ is the set of all transactions in that contract-day. On no-trade days inserted during panel expansion, $p^{PM}_{i,t}$ is set to the most recently observed Polymarket VWAP, as described in Section 3. This means that the Polymarket side of the comparison is stale by construction on those days, a point revisited in the robustness analysis.

4.1.2 Option-implied probability

Theoretical foundation

The option-implied probability $\hat{p}^{OPT}_{i,t}$ is derived from listed call prices using the result of Breeden and Litzenberger (1978). Under no-arbitrage, the risk-neutral probability density of the underlying asset price at expiry can be recovered from the cross-section of European call prices as the second partial derivative of the call pricing function with respect to the strike:

$$q(S_T = K) = e^{r\tau} \frac{\partial^2 C(K,\tau)}{\partial K^2} \tag{2}$$

where $C(K,\tau)$ is the price of a European call with strike K and time-to-expiry τ, and r is the continuously compounded risk-free rate. The cumulative risk-neutral probability that the underlying asset exceeds K at expiry follows from integrating the density in equation (2) from K to infinity:

$$\mathbb{Q}(S_T > K) = \int_K^\infty q(x)\,dx = e^{r\tau} \int_K^\infty \frac{\partial^2 C(x,\tau)}{\partial x^2}\,dx = -e^{r\tau} \frac{\partial C(K,\tau)}{\partial K} \tag{3}$$

where the last equality uses the fact that $\partial C/\partial K \to 0$ as $K \to \infty$, since call prices vanish for arbitrarily high strikes. The intuition is straightforward: raising the strike by a small amount dK reduces the call payoff by dK in every state where $S_T > K$. The rate at which the call price falls as the strike increases therefore directly encodes the discounted risk-neutral probability that the stock finishes above K. This expression is the negative of the discounted first derivative of the call price with respect to the strike (Jackwerth and Rubinstein, 1996). This measure is risk-neutral and may embed a risk premium, so it should not be interpreted as the physical probability of the event.

Finite-difference approximation

Since option prices are quoted only at a discrete set of strikes, the derivative $\partial C/\partial K$ is not directly observable. Following the standard approach in the empirical literature on risk-neutral densities (Aït-Sahalia and Lo, 1998; Figlewski, 2010), we approximate it using a centred finite difference:

$$\hat{p}_{i,t}^{OPT} = e^{r\tau} \cdot \frac{C^{mid}(K - \Delta,\, \tau) - C^{mid}(K + \Delta,\, \tau)}{2\Delta} \tag{4}$$

where $C^{mid}(K \pm \Delta, \tau) = \frac{1}{2}[C^{bid}(K \pm \Delta, \tau) + C^{ask}(K \pm \Delta, \tau)]$ is the mid-price of the call at each bracket strike. This expression uses the price difference between two calls with strikes around the Polymarket threshold. Scaled by the distance between the two strikes, this difference approximates the local slope of the call price function with respect to the strike. Under the Breeden–Litzenberger framework, this slope is linked to the risk-neutral probability that the underlying stock finishes above the threshold. As Δ becomes small, the approximation converges to $\mathbb{Q}(S_T > K)$. With a finite strike interval, the accuracy of the measure depends on how smooth the call price function is around K, which is why the approximation is less precise for contracts that are far from at-the-money or close to expiry.

American options

The options in this study are American-style contracts. We apply the European framework of Breeden and Litzenberger (1978) directly, which is standard practice in the empirical risk-neutral density literature (Jackwerth and Rubinstein, 1996; Aït-Sahalia and Lo, 1998). For call options on non-dividend-paying stocks, early exercise is never optimal (Merton, 1973). Six of the eight equities in our sample (TSLA, NVDA, META, GOOGL, NFLX, PLTR) pay no dividends; the remaining two (AAPL, MSFT) have annual dividend yields below 0.7% over the sample period. For short-dated contracts with a maximum of 32 calendar days to expiry, the early exercise premium is economically negligible across all names in the sample.

Implementation

For each contract i on observation date t, call mid-prices are retrieved at strikes $K_i - 5$ and $K_i + 5$, so that $\Delta = 5$. The use of $\Delta = 5$ is imposed by the strike grid available consistently through LSEG Workspace for the relevant contracts. Narrower brackets could not be retrieved reliably across tickers, strikes, and expiry dates, so $\Delta = 5$ is the closest consistently available bracket around each Polymarket threshold rather than an optimised modelling choice.

This choice means that the option-implied probability should be interpreted as a feasible empirical benchmark constructed from observed option quotes, rather than as a model-free estimate of the exact risk-neutral tail probability. The limitation is most relevant for contracts far from at-the-money or close to expiry, where the call-price function is more curved and a fixed strike interval provides a less precise local approximation. This bracket width is determined by the minimum available strike spacing in the LSEG Workspace option chain for the equities in our sample, and is not a free choice. As a result, the approximation window represents different shares of the underlying price across stocks (approximately ±2–3% for higher-priced stocks and ±4–5% for lower-priced ones) introducing some variation in precision across tickers. The moneyness control $|\ln(K_i/S_t)|$ included in all specifications partly absorbs this cross-sectional variation. The discount factor is $e^{r\tau}$, where $r = 3\%$ per annum and $\tau = (T_i - t)/365$.

For this expression to yield a valid probability, two conditions must hold. First, $C^{mid}(K_i - 5) \geq C^{mid}(K_i + 5)$: call prices must be weakly decreasing in the strike, which follows from standard no-arbitrage. Second, the scaled spread must not exceed one: $e^{r\tau}[C^{mid}(K_i - 5) - C^{mid}(K_i + 5)] \leq 10$. When these conditions are not met—typically for deep in- or out-of-the-money contracts close to expiry, where the distribution is highly concentrated and a fixed $\Delta = 5$ becomes too coarse—the observation is dropped, as it does not provide a meaningful probability estimate. The frequency of these cases and how they are distributed across tickers are reported in Section 3.

4.2 Gap Measures and Transformations

4.2.1 The logit transformation

The analysis compares $p_{i,t}^{PM}$ and $\hat{p}_{i,t}^{OPT}$ in logit space rather than in probability levels. The reason is that a given difference in probability levels does not carry the same economic meaning across the unit interval: a gap of five percentage points near 50% represents very different disagreement from the same gap near 95%. The logit transformation,

$$\operatorname{logit}(p) = \ln\left(\frac{p}{1-p}\right) \tag{5}$$

maps $(0, 1)$ to $(-\infty, +\infty)$ and makes the scale more symmetric, so that equal logit differences correspond to more comparable levels of disagreement across different parts of the probability range.

4.2.2 Boundary treatment

The logit function becomes numerically unstable as p approaches zero or one. In the main specification, both $p_{i,t}^{PM}$ and $\hat{p}_{i,t}^{OPT}$ are therefore clipped to the interval $[0.01, 0.99]$ before the transformation is applied. We refer to this as the clipped-boundary sample. This is a numerical-stability choice: it prevents extreme transformed values from unduly influencing the regression estimates while retaining observations for which the option-implied probability has a valid interpretation within $[0, 1]$. The unclipped probabilities are preserved separately in the dataset throughout.

As a robustness check, we also estimate the main specifications on an alternative drop-boundary sample, in which observations outside $[0.01, 0.99]$ are removed entirely rather than clipped. The differences between the two approaches are documented in Section 5.3.

4.2.3 Gap measures

The main dependent variable is the absolute logit gap:

$$Y_{i,t} = \left| \text{logit}(\hat{p}_{i,t}^{OPT}) - \text{logit}(p_{i,t}^{PM}) \right| \tag{6}$$

A larger value of $Y_{i,t}$ indicates greater disagreement between Polymarket and the option-implied benchmark on the probability of the threshold event. This gap should be interpreted as a relative difference between two market-based probability measures, not as direct evidence that one of the two markets is incorrectly priced. Using the absolute value reflects the research question, whether the two markets disagree and by how much, rather than asking which one is higher.

To address the direction of disagreement, we also examine the signed gap:

$$SG_{i,t} = \text{logit}(\hat{p}_{i,t}^{OPT}) - \text{logit}(p_{i,t}^{PM}) \tag{7}$$

A negative value of $SG_{i,t}$ means that Polymarket assigns a higher probability to the threshold event than the options market. The signed gap is used in the robustness analysis rather than as the primary dependent variable, since the research question centres on the magnitude of disagreement rather than its direction.

4.3 Explanatory Variables

The explanatory variables fall into three groups: Polymarket-side liquidity measures, an options-side liquidity measure, and controls.

4.3.1 Polymarket liquidity

The main Polymarket liquidity proxy in the preferred specification is $\ln(\text{Volume}_{i,t})$, the natural logarithm of daily dollar trading volume. Higher volume is expected to be associated with a smaller gap, as more actively traded contracts are likely to incorporate information more quickly and track the option-implied benchmark more closely. An alternative activity proxy, the number of daily transactions ($n\_\text{trades}_{i,t}$), is used in a secondary specification to confirm that the volume effect is not sensitive to the choice of liquidity measure.

In the robustness analysis, continuous liquidity measures are replaced by binary indicators to assess whether the effect operates through the presence or absence of trading rather than its intensity. These include a no-trade dummy (equal to one when no transactions are recorded on a given day), a zero-move dummy (equal to one when the Polymarket probability is unchanged from the previous observation), and a low-volume dummy based on the bottom quartile of dollar volume. These variables are not included in the baseline specification.

4.3.2 Options-side liquidity

$\text{rel\_opt\_spread}_{i,t}$ is the relative bid-ask spread of the call options used to construct $\hat{p}_{i,t}^{OPT}$, defined as the average of the relative spreads at the two bracket strikes:

$$\text{rel\_opt\_spread}_{i,t} = \frac{1}{2}\left[\frac{C^{ask}(K-5) - C^{bid}(K-5)}{C^{mid}(K-5)} + \frac{C^{ask}(K+5) - C^{bid}(K+5)}{C^{mid}(K+5)}\right] \tag{8}$$

This variable allows us to distinguish disagreement arising from prediction-market frictions from disagreement that may reflect stale or imprecise option quotes. It is winsorised at the 1st and 99th percentiles to limit the influence of outliers. The expected sign is ambiguous: wider spreads indicate less liquid options and might produce noisier $\hat{p}^{OPT}_{i,t}$ values, which could widen the gap; but wide spreads may also coincide with periods of low information flow in both markets, where stale prices in both series reduce the measured gap. The empirical results address this ambiguity directly.

4.3.3 Control variables

$\ln(\text{TTR}_{i,t})$ is the natural logarithm of the continuous calendar time remaining until contract resolution, measured in days. This is the central variable of the analysis. Shorter time-to-resolution is expected to be associated with a larger gap, consistent with greater sensitivity of the option-implied probability near expiry. The associated economic interpretation is discussed further in Section 4.4.3.

$|\ln(\text{Moneyness}_{i,t})|$ is the absolute value of the log ratio of the strike to the current stock price, $|\ln(K_i/S_t)|$. It captures how far the contract is from at-the-money and controls for the fact that the finite-difference approximation is less precise for deep in- or out-of-the-money contracts, where the risk-neutral density is more curved around the threshold.

4.3.4 Note on the Amihud illiquidity measure

An Amihud price-impact measure was considered as an additional Polymarket-side liquidity proxy, defined as the ratio of the absolute daily price change to dollar volume:

$$\text{Amihud}_{i,t} = \frac{|\Delta p^{PM}_{i,t}|}{\text{Volume}_{i,t}} \tag{9}$$

In practice, this measure proved unstable across specifications: it was significant when entered alone but lost significance once dollar volume was included, and its sign reversed across different fixed effects structures. Given this instability and its limited suitability for prediction-market microstructure — where price impact is less directly interpretable than in order-book equity markets — the Amihud measure is excluded from the preferred specification. Where this measure is used, the estimation is restricted to observations for which it can be computed. This produces a restricted liquidity subsample ($N = 3,301$). Since the measure is unstable across specifications and is not included in the preferred model, results based on this subsample are treated as intermediate checks rather than as part of the main design.

4.4 Econometric Framework

4.4.1 Baseline specification

The baseline regression is estimated on a contract-day panel with two-way fixed effects. All regressions are estimated by ordinary least squares. Although the dependent variable $Y_{i,t}$ is non-negative by construction, OLS fitted values are positive throughout the estimation sample given the sample mean of 1.29 logit units, so the non-negativity constraint is not binding in practice. The baseline equation is:

$$Y_{i,t} = \beta_1 \ln(\text{Volume}_{i,t}) + \beta_2 \ln(\text{TTR}_{i,t}) + \beta_3 |\ln(\text{Moneyness}_{i,t})| + \beta_4\, \text{rel\_opt\_spread}_{i,t} + \alpha_i + \delta_t + \varepsilon_{i,t} \quad (10)$$

where α_i are contract fixed effects and δ_t are calendar-date fixed effects. Standard errors are clustered at the contract level.

Contract fixed effects control for all time-invariant differences across contracts, such as ticker characteristics, strike levels, and contract type. Date fixed effects capture shocks that affect all contracts on a given day, including overall market movements or macro news. As a result, identification comes from how the explanatory variables change over time within each contract, after accounting for common daily shocks.

Standard errors are clustered at the contract level to allow for arbitrary correlation of residuals within each contract over time. Since contracts vary in length and the panel is unbalanced, within-contract serial correlation is a natural concern. This clustering choice matches the entity dimension of the panel and the within-contract variation used for identification. Calendar-date fixed effects absorb shocks that are common across contracts on a given day.

4.4.2 Preferred specification

The preferred specification is estimated on the full sample ($N = 5,376$) and includes dollar volume, time-to-resolution, moneyness, and the relative option spread as regressors, with contract and date fixed effects. The Amihud price-impact measure is not included in the preferred specification. Specifications using this measure rely on a smaller restricted liquidity subsample ($N = 3,301$) and are treated only as intermediate checks. A reduced full-sample specification excluding the option spread is also estimated to assess whether the main volume and TTR associations depend on the inclusion of option-side liquidity controls.

4.4.3 Research hypotheses

Three hypotheses motivate the empirical design. Each is stated as an expected association between an explanatory variable and the gap rather than as a causal claim.

H1 (Time-to-resolution). The absolute gap is expected to be negatively associated with $\ln(\text{TTR}_{i,t})$: contracts closer to resolution should exhibit a larger gap. A plausible interpretation is that option-implied probabilities become more sensitive to small movements in the underlying asset price near expiry, which may widen the measured gap if Polymarket prices do not adjust at the same pace.

H2 (Polymarket volume). More actively traded Polymarket contracts are expected to exhibit a smaller gap, consistent with more frequent price updating.

H3 (Option spread). The sign of the option spread coefficient is theoretically ambiguous. Wider spreads indicate less liquid options and might produce noisier $\hat{p}_{i,t}^{OPT}$ values, which would widen the gap; alternatively, wide spreads may coincide with periods of low information flow in both markets, mechanically reducing measured disagreement. Rather than specifying a directional prior, H3 examines which association is more consistent with the data.

4.4.4 Robustness strategy

The main specification is assessed against a range of alternatives, each targeting a specific modelling choice.

Alternative liquidity proxies. The continuous volume measure is replaced by binary indicators of inactivity: a no-trade dummy, a zero-move dummy, and a low-volume dummy. These test whether the volume effect operates through the presence or absence of trading rather than its intensity.

Alternative fixed effects structures. The preferred two-way specification is compared against a contract-only fixed effects model and a ticker-plus-date fixed effects model. These alternatives test whether the results are sensitive to the level at which common shocks are absorbed.

Alternative dependent variables. The signed gap $SG_{i,t}$ is used to assess whether the disagreement has a systematic direction.

Persistence. A lagged dependent variable $Y_{i,t-1}$ is added to the preferred specification to control for serial correlation in the gap and assess whether the main associations remain after controlling for persistence.

Non-linearities in time-to-resolution. The continuous $\ln(\text{TTR})$ term is replaced by indicator variables for four maturity ranges: 0–3 days, 4–7 days, 8–14 days, and 15 or more days. This tests whether the effect is concentrated in the final days before resolution or distributed more evenly.

Volume–TTR interaction. An interaction term $\ln(\text{Volume}_{i,t}) \times \ln(\text{TTR}_{i,t})$ is added to test whether the gap-reducing effect of trading activity varies with the time remaining to resolution.

Quantile regressions. We also estimate the preferred specification at the 50th, 75th, and 90th percentiles of the conditional distribution of $Y_{i,t}$ to see whether the main effects are driven by typical levels of disagreement or by more extreme cases. Since standard quantile regression cannot handle the same two-way fixed effects as the baseline model, we use within-contract demeaned variables as an approximation. These results should therefore be seen as supportive evidence rather than directly comparable to the main specifications.

Drop-boundary sample. All main specifications are re-estimated on an alternative drop-boundary sample in which observations outside $[0.01, 0.99]$ are removed entirely rather than clipped, to assess whether the boundary treatment materially affects the results.

5. Empirical Results

This section presents the empirical results of the analysis. Section 5.1 reports descriptive statistics. Section 5.2 presents the baseline regression results and the preferred specification. Section 5.3 documents robustness checks across alternative liquidity proxies, fixed effects structures, and dependent variable transformations. Section 5.4 analyses heterogeneity across underlying equities and contract types. Section 5.5 reports additional analyses: gap persistence, non-linearities in the time-to-resolution effect, liquidity interactions, and quantile regressions.

5.1 Descriptive Statistics

Table 4 reports summary statistics for the main variables in the full estimation sample ($N = 5,376$). The dependent variable $Y_{i,t} = |\text{logit}(\hat{p}_{i,t}^{OPT}) - \text{logit}(p_{i,t}^{PM})|$ has a mean of 1.29 and a median of 1.03 logit units, indicating substantial disagreement between Polymarket and option-implied probabilities across a broad set of contract-day observations. To give economic content to these magnitudes, for probabilities around 50%, a logit gap of 1.29 corresponds roughly to a 25–30 percentage-point difference, although the mapping is non-linear and depends on the initial probability level. The distribution is right-skewed: the standard deviation of 1.08 is nearly as large as the mean of 1.29, and the maximum of 8.44 appears to be concentrated in near-expiry observations, where option-implied probabilities are especially sensitive to small movements in the underlying asset and measured differences are largest.

Among the explanatory variables, $\ln(\text{TTR})$ shows the strongest unconditional association with Y (Pearson correlation -0.44), providing a first descriptive indication of the importance of contract maturity. The unconditional correlation between dollar trading volume and the gap is modestly positive (+0.08): contracts with higher volume tend to be closer to resolution, where the gap is larger, so the raw correlation does not reflect the conditional association documented in the regressions. The coefficients on $\ln(\text{Volume})$ and $\ln(\text{TTR})$ are stable across specifications that include each variable separately and jointly, suggesting that multicollinearity between these two regressors is not a material concern. The relative option spread shows a negative unconditional correlation with Y (-0.16), consistent with the conditional negative relationship documented in Section 5.2.

Cross-sectional heterogeneity across tickers is visible in Panel B of Table 4. NVDA exhibits the largest mean gap (1.52), followed by GOOGL (1.44). NFLX has a substantially smaller mean gap (0.44), consistent with its limited trading activity on Polymarket and the small number of observations. The sources of this variation are explored further in Section 5.4.

Figure 2 plots the evolution of the daily cross-sectional mean of $Y_{i,t}$ over the period with regular matched-sample coverage, together with the interquartile range. The gap fluctuates around the full-sample mean of 1.29 logit units without a clear time trend, and the sharp increase concentrated in the final week of the sample may reflect the mechanical concentration of contracts in the near-expiry regime, where the gap is systematically larger.

5.2 Baseline Results and Preferred Specification

Table 5 presents the main regression results. All specifications are estimated on the full sample ($N = 5,376$) with contract and calendar-date fixed effects and standard errors clustered at the contract level. Column (1) includes dollar volume as the sole Polymarket-side liquidity proxy. Column (2) replaces volume with the number of daily trades to confirm that the liquidity result is not sensitive to the choice of proxy. Column (3) is the preferred specification, which adds the relative option bid-ask spread to Column (1).

5.2.1 Time-to-resolution

Figure 3 illustrates the unconditional relationship between $\ln(\text{TTR})$ and $Y_{i,t}$ across all contract-day observations. The negative association is visible in the raw data: the pooled OLS slope is −0.43 with a Pearson correlation of −0.44, the strongest unconditional association among all regressors in the sample (Table 4). Two features of the scatter are worth noting. First, the variance of $Y_{i,t}$ is markedly larger at low values of $\ln(\text{TTR})$, consistent with the concentration of extreme observations in the near-expiry regime. Second, the vertical clustering reflects the fact that time-to-resolution takes a limited set of repeated values in the daily panel, even though the variable is measured as continuous calendar time remaining in days. The unconditional slope of −0.43 is attenuated relative to the panel estimate of −0.785 in Table 5, which exploits within-contract variation after absorbing contract and date fixed effects.

The most consistent finding across specifications is the negative and highly significant coefficient on $\ln(\text{TTR})$. In the preferred specification (Column 3 of Table 5), the estimate is −0.785 ($p < 0.001$). The magnitude is stable across specifications and the coefficient remains significant at the one-percent level. As contracts approach their resolution date, the gap between Polymarket and option-implied probabilities widens substantially.

The regressions identify a strong empirical association between time-to-resolution and the size of the gap; they do not directly identify the underlying mechanism. One plausible economic interpretation runs through option gamma. Gamma measures how quickly an option's delta changes as the underlying price moves, and it scales inversely with the square root of time to expiry, diverging as expiry approaches for at-the-money contracts. As a result, when contracts have little time remaining, a small move in the stock price can shift the option-implied probability substantially: there is no longer enough time for the price to revert across the threshold, so the risk-neutral probability of the event is highly sensitive to where the stock is trading relative to the strike. If Polymarket prices respond to this environment more slowly, or if the two markets draw on different information sets in the final trading days, a widening gap would be the expected result. This interpretation is consistent with the empirical pattern.

5.2.2 Trading volume

The coefficient on $\ln(\text{Volume})$ is negative and significant in the preferred specification ($\hat{\beta} = -0.019$, $p = 0.032$). More actively traded Polymarket contracts are associated with a smaller gap relative to option-implied probabilities. One plausible interpretation is that active trading is associated with faster incorporation of information, though the regression does not identify this mechanism directly. A one-standard-deviation increase in $\ln(\text{Volume})$ (approximately 2.65 units) is associated with a reduction in the logit gap of approximately 0.05 units, relative to a sample mean of 1.29 — a reduction of approximately

4 per cent of the mean gap. The volume effect is therefore statistically robust but economically modest in magnitude, consistent with liquidity playing a secondary role relative to time-to-resolution. The number-of-trades specification in Column (2) of Table 5 confirms that the result is not driven by the specific liquidity measure chosen.

5.2.3 Options market illiquidity

The preferred specification also includes the relative bid-ask spread of the options used to construct $\hat{p}_{i,t}^{OPT}$. The coefficient is -1.078 ($p < 0.001$), indicating a negative association between the option spread and the gap in the clipped-boundary sample. This sign is not straightforward to interpret: wider bid-ask spreads normally indicate less liquid options and could therefore make the option-implied benchmark noisier, rather than mechanically reducing disagreement.

We therefore treat this coefficient primarily as a benchmark-side control rather than as a central result. The negative coefficient should not be read as evidence that less liquid options improve benchmark quality. A more plausible interpretation is that the spread variable partly captures features of option quotes close to the boundary, where the finite-difference approximation is less precise and the treatment of probabilities near zero or one matters most.

Consistent with this interpretation, the coefficient loses statistical significance and changes sign in the drop-boundary sample reported in Section 5.3.4. The option-spread result is therefore sample-sensitive and secondary to the two main empirical patterns of the paper: the time-to-resolution effect and the Polymarket trading activity association.

5.2.4 Contract moneyness

The coefficient on $|\ln(\text{Moneyness})|$ is not significant in any of the three baseline specifications. Given this consistent insignificance, we retain moneyness as a control but do not place interpretive weight on it. The variable becomes more important in the drop-boundary robustness sample and in the ticker-FE specification, as discussed in Section 5.3.

5.3 Robustness Checks

Table 6 presents robustness checks across three panels. Panel A uses binary liquidity proxies for Polymarket. Panel B considers alternative fixed effects structures. Panel C uses the signed gap as the dependent variable. A fourth robustness check, comparing the clipped-boundary sample to the drop-boundary sample in which boundary observations are removed entirely, is presented in Section 5.3.4.

5.3.1 The no-trade and zero-move results

The no-trade and zero-move dummies in Panel A of Table 6 are negative and significant (-0.114, $p = 0.001$; -0.110, $p = 0.001$): days characterised by inactivity in Polymarket are associated with a *smaller* measured gap relative to options, not a larger one. The low-volume Q25 dummy is not significant ($p = 0.191$).

It is important to state clearly what this result does and does not imply. It does not imply that inactive markets produce better-calibrated prices or that stale Polymarket probabilities are more informative. A more precise interpretation is that on inactive days inserted during panel expansion, the Polymarket probability is carried forward from the previous traded observation and remains anchored to a stale value. If the option-implied probability is also relatively stable on low-information days, the two series move in parallel and the measured gap remains small. By contrast, on days with active Polymarket trading, new order flow is incorporated, potentially reflecting intraday movements, short-term sentiment, or event-specific narratives not captured in the end-of-day option-implied probability. A small measured gap on inactive days indicates that neither market is moving, not that they genuinely agree.

5.3.2 Alternative fixed effects structures

Column (4) of Panel B in Table 6 retains only contract fixed effects, removing date fixed effects. The ln(TTR) coefficient falls from -0.785 to -0.401, consistent with the possibility that date fixed effects absorb part of the TTR variation that is common across contracts resolving on the same calendar day. Volume and the option spread remain strongly significant in this specification.

Column (5) replaces contract fixed effects with ticker fixed effects, so that identification exploits variation across contracts of the same underlying equity rather than within-contract time-series variation. TTR remains negative and highly significant (-0.648, $p < 0.001$). Moneyness becomes large and significant (2.678, $p < 0.001$) in this cross-sectional specification. The significance of moneyness in the ticker-FE specification contrasts with its consistent insignificance in the contract-FE specifications. This pattern suggests that moneyness is primarily a cross-sectional phenomenon: contracts with strikes far from the current stock price exhibit systematically larger gaps on average, but within a given contract the moneyness does not vary enough over the sample window to generate within-contract identification. The ticker-FE specification, by exploiting cross-contract variation, recovers this cross-sectional relationship.

5.3.3 Directional differences in probabilities: the signed gap

Column (6) of Table 6 uses the signed gap $\text{logit}(\hat{p}_{i,t}^{OPT}) - \text{logit}(p_{i,t}^{PM})$ as the dependent variable. The mean signed gap in the estimation sample is -0.176, meaning that, in logit terms, Polymarket assigns a higher probability to threshold events than options, on average. Volume is negative and strongly significant ($\hat{\beta} = -0.127$, $p < 0.001$): more actively traded contracts are associated with an even higher Polymarket probability relative to options. The ln(TTR) coefficient is not significant ($p = 0.637$), suggesting that the expiry effect primarily widens the absolute gap without generating a robust directional bias in the signed difference. The option spread is marginally significant (-0.389, $p = 0.068$).

The cross-sectional pattern in the signed gap is also notable. NVDA has the most negative mean signed gap (-0.633), consistent with Polymarket being substantially more bullish on NVDA threshold events than options. TSLA (+0.143) and PLTR (+0.120) have slightly positive mean signed gaps, indicating that for those tickers the options market assigns relatively higher probabilities.

One possible explanation for the negative mean signed gap is the risk premium embedded in option-implied probabilities. Because $\hat{p}_{i,t}^{OPT}$ is a risk-neutral measure, it may systematically differ from the physical probability of the event. For threshold contracts on equities that are positively correlated with aggregate market risk, risk-averse investors require compensation for holding the underlying asset, which causes the risk-neutral probability of an upward move to fall below the corresponding physical probability. If Polymarket prices reflect something closer to participants' physical beliefs about the outcome, a negative signed gap, where Polymarket assigns higher probabilities than options, would be the expected result. This interpretation is consistent with the direction of the signed gap documented here, though we cannot test it directly with the data available. The directional asymmetry therefore complements the main absolute-gap results by showing that disagreement is not purely symmetric, but it remains modest and sensitive to the treatment of boundary observations, as shown in the drop-boundary robustness exercise below.

5.3.4 Robustness to boundary treatment: the drop-boundary sample

Table 7 compares the preferred specification across the two probability boundary treatments described in Section 4. The clipped-boundary sample ($N = 5{,}376$) clips p^{PM} and $\hat{p}^{OPT}$ to $[0.01, 0.99]$ before the logit transformation. The drop-boundary sample ($N = 4{,}055$) removes observations outside $[0.01, 0.99]$ entirely.

The TTR and volume coefficients are stable in sign, magnitude, and significance across both samples, and the central findings are not sensitive to the boundary treatment. Two differences are worth noting. First, moneyness becomes strongly significant in the drop-boundary sample (+2.220, $p < 0.01$), where it is insignificant in the clipped-boundary sample. One plausible explanation is that extreme-probability observations are disproportionately deep in- or out-of-the-money, so clipping the probabilities compresses some of the variation that would otherwise load onto moneyness; removing those observations allows that variation to be more clearly identified. Second, the option spread loses significance in the drop-boundary sample (+0.302, $p = 0.645$) and the sign reverses. Observations with probabilities near zero or one tend to have wide bid-ask spreads and large logit gaps; removing them materially weakens the negative association between the spread and the gap. This suggests that the option-spread result in the clipped-boundary sample depends on the treatment of boundary observations. For this reason, we do not treat it as a robust finding. The main conclusions of the paper instead rest on the time-to-resolution and Polymarket trading activity results, which remain stable across both samples.

5.4 Heterogeneity Analysis

5.4.1 Cross-sectional heterogeneity by underlying equity

Table 8 reports the preferred specification estimated separately for each underlying equity. All specifications use the full sample.

The primary result is that ln(TTR) is negative and significant at the one-percent level for every ticker except NFLX, whose small sample (110 observations) precludes reliable inference. This near-universality is the most important takeaway: regardless of the underlying equity, contracts approaching expiry exhibit a substantially larger gap. Differences in the magnitude of the TTR coefficient across tickers may reflect genuine heterogeneity, but given the per-ticker sample sizes they should be treated as suggestive. PLTR (−1.352) and TSLA (−1.407) show the largest TTR effects, while the mega-cap equities cluster around −0.67.

Table 9 provides a descriptive grouping of the underlying equities. We separate AAPL, MSFT, GOOGL, and META as large technology platforms, and NVDA, TSLA, and PLTR as a higher-volatility group during the sample period. NFLX is excluded from both groups because its small sample size ($N = 110$) makes group-level inference less reliable. This grouping is exploratory and is not intended as a formal classification of the underlying equities.

The purpose of this grouping is to assess whether the main patterns differ between more established mega-cap equities and the higher-volatility names in the sample. The TTR coefficient is larger in absolute value for the high-volatility group (−1.015) than for the mega-cap group (−0.670). This pattern is consistent with stronger near-expiry widening among the more volatile names in the sample, but it should be interpreted cautiously because the sample is limited and the analysis does not formally test differences between group coefficients.

The option-spread coefficient is negative and significant only in the mega-cap group, while it is positive and insignificant in the high-volatility group. This reinforces the view that the option-spread result is not stable across samples and should not be treated as a core finding.

5.4.2 Heterogeneity by contract type

Table 10 reports the mean gap by contract type alongside selected regression coefficients. Panel A shows a clear descriptive pattern: the mean gap decreases from 1.98 for daily contracts to 1.60 for weekly contracts and 0.99 for month-end contracts. This pattern is consistent with the time-to-resolution result, because contract type largely determines the typical maturity profile observed in the panel: daily contracts are almost always observed close to resolution, while month-end contracts usually have more time remaining.

These differences should not be interpreted as an independent contract-type effect. Contract type is mechanically related to maturity, so daily, weekly, and month-end contracts differ in their typical time-to-resolution profile by construction. For this reason, the contract-type results are best read as descriptive evidence consistent with the maturity channel, rather than as evidence of an independent contract-type effect.

Panel B should also be interpreted with caution. ln(TTR) is not separately identified in the daily subsample, where each contract has only one observation by construction. In the month-end subsample, there is more maturity variation in principle, but once contract and date fixed effects are included, the remaining identifying variation is limited. In the weekly subsample, a coefficient can be estimated but is not statistically significant ($p = 0.669$). Volume is significant only in the month-end subsample (−0.050, $p < 0.001$), which suggests that the volume association is easier to detect when contracts remain active for longer. Overall, Table 10 is consistent with the main finding that shorter-maturity observations are associated with larger gaps, but it does not establish contract type as an independent source of variation in the gap.

5.5 Additional Analyses

5.5.1 Gap persistence: the lagged dependent variable

To assess whether the main results partly reflect serial correlation in the gap, Table 11 augments the preferred specification with $Y_{i,t-1}$. The coefficient on $Y_{i,t-1}$ is 0.447 ($p < 0.001$): approximately 45 per cent of the previous day's gap carries forward into the following day, indicating that the gap exhibits substantial persistence across adjacent trading days. The ln(TTR) and rel_opt_spread coefficients remain significant after controlling for persistence, suggesting that those effects are not solely driven by persistence in the gap. Volume becomes marginally significant ($p = 0.065$) once the lagged gap is included, suggesting that part of the volume association may operate through persistence. This does not overturn the main volume finding, but indicates that the contemporaneous effect is attenuated once past gap levels are controlled for.

We note that including a lagged dependent variable in a two-way fixed effects model with a short panel introduces Nickell bias (Nickell, 1981). Given the short contract panels in our sample, with an average of 5.5 observations per contract and a median of 4.0, this specification should be interpreted as a robustness exercise rather than a preferred dynamic panel model. The purpose of including the lagged dependent variable is therefore not to estimate a full dynamic panel model, but to check whether the main associations remain after accounting for persistence in the gap. The stability of the TTR and volume coefficients after controlling for the lagged gap suggests that the main findings are not driven by serial correlation in the gap.

5.5.2 Non-linearities in time-to-resolution

To assess the functional form of the TTR effect, Table 12 replaces ln(TTR) with indicator variables for four maturity ranges, with 15 or more days as the reference category. The descriptive statistics reveal a clear monotonically decreasing pattern from the 0–3 day bin (mean 1.807) to the 15+ day bin (mean 0.802). The regression confirms that being in the final three days before expiry is associated with a gap that is 0.947 logit units larger than at horizons of 15 or more days ($p < 0.001$), and the 4–7 day bin also shows a significant positive coefficient (+0.414, $p < 0.001$). On the full sample, both near-expiry ranges are separately identified, providing clearer evidence of non-linearity than in earlier restricted-sample checks. The 8–14 day dummy cannot be separately identified given limited within-contract variation in that range. Taken together, the bin analysis suggests that the gap widens sharply in the final week before resolution and accelerates further in the last three days.

Table 13 shows that this near-expiry widening is pervasive across all tickers, with all equities except NFLX exceeding a mean gap of 1.3 logit units in the 0–3 day bin.

Figure 4 visualises the cross-sectional dimension of this non-linearity. The near-expiry widening is pervasive across all non-NFLX names, with the broadly decreasing pattern from shorter to longer horizons consistent across most tickers. NFLX remains the exception, with mean gap values below 0.60 logit units across all bins, consistent with its limited Polymarket trading activity and small per-bin sample sizes.

5.5.3 The direction of the gap near expiry

Table 14 and Figure 6 report the mean signed gap by TTR bin. The raw bin means are negative across all maturity groups, indicating that Polymarket assigns higher event probabilities than options on average in each bin. The most negative values appear in the 15+ day bin (mean −0.226) and the 0–3 day bin (mean −0.197), with a smaller absolute signed gap in the intermediate ranges (4–7 days: −0.116; 8–14 days: −0.096). The 8–14 day bin is not statistically distinguishable from zero at conventional significance levels, as reflected in the confidence interval crossing zero in Figure 6. This evidence is descriptive and modest:

the directional pattern does not appear concentrated near expiry, but it is sensitive to the treatment of boundary observations, as shown in Section 5.3.4.

5.5.4 The interaction between volume and time-to-resolution

Table 15 augments the preferred specification with the interaction term $\ln(\text{Volume}) \times \ln(\text{TTR})$. The interaction coefficient is -0.043 ($p < 0.001$). The total marginal effect of volume on the gap is:

$$\frac{\partial Y_{i,t}}{\partial \ln(\text{Volume}_{i,t})} = \hat{\beta}_{\text{vol}} + \hat{\beta}_{\text{int}} \cdot \ln(\text{TTR}_{i,t}) = +0.041 - 0.043 \cdot \ln(\text{TTR}_{i,t}) \tag{11}$$

For the majority of observations, $\ln(\text{TTR})$ is positive, so the negative interaction implies that the gap-reducing association of volume is stronger when TTR is larger. In other words, trading activity is most strongly associated with a narrower gap for contracts with more time remaining. At the median value of $\ln(\text{TTR}) = 1.35$, the total marginal effect of volume is $+0.041 - 0.043 \times 1.355 = -0.017$, indicating that volume remains associated with a reduction in the gap at typical horizons.

It is important to note that in the interaction specification, the main effect of $\ln(\text{Volume})$ is $+0.041$ ($p = 0.003$), which is positive — a sign reversal relative to the preferred specification coefficient of -0.019. This is not a contradiction: in the interaction model, the main effect is evaluated at $\ln(\text{TTR}) = 0$, i.e., approximately one day before resolution. A positive volume coefficient at this point is economically plausible: in the final day before resolution, additional trading activity may reflect disagreement rather than information convergence, as traders take positions on the outcome. The interaction term then restores the negative volume effect at all typical TTR values above zero. Near expiry, when $\ln(\text{TTR})$ approaches zero, the volume effect attenuates and variation in the gap is more closely associated with time-to-resolution itself.

Figure 5 plots the total marginal effect of $\ln(\text{Volume})$ on $Y_{i,t}$ as a continuous function of $\ln(\text{TTR})$, together with its 95% confidence band. The marginal effect crosses zero at $\ln(\text{TTR}) = 0.95$, corresponding to approximately 2.6 calendar days before resolution: to the right of this threshold volume is associated with a smaller gap, while to the left it reverses. The confidence band widens at both extremes of the distribution, reflecting limited observations at very short and very long horizons. Across much of the economically relevant range, the marginal effect is negative, consistent with the preferred specification result of -0.019 (Table 5).

This result connects the two central empirical patterns of the paper. The negative association between Polymarket trading activity and the gap is most visible for longer-horizon contracts, where the market has more time to incorporate information before resolution. Near expiry, additional volume provides less incremental benefit and the time-to-resolution pattern becomes more pronounced.

5.5.5 Quantile regressions

Table 16 reports quantile regression estimates at the median, 75th, and 90th percentiles. These regressions use within-contract demeaned variables as an approximation for contract fixed effects, since standard quantile regression does not accommodate two-way fixed effects directly. They should be read as supporting evidence rather than at the same level of identification as the main panel specifications. The within-contract demeaning approximation introduces bias that vanishes asymptotically in T but may be non-negligible in short panels; this limitation is common in the literature (Canay, 2011).

Two findings are worth noting. First, $\ln(\text{TTR})$ is negative and significant at all three quantiles, with the absolute magnitude increasing monotonically from -0.370 at the median to -0.631 at the 90th percentile. The expiry effect is therefore not limited to average contracts; it is especially pronounced in episodes of extreme disagreement. Second, $\ln(\text{Volume})$ is negative and significant across all quantiles, with a larger coefficient in absolute terms at the 90th percentile, suggesting that the negative association between trading activity and disagreement is strongest in the upper tail of the gap distribution.

The option-spread coefficient is also negative and significant across the quantiles reported in Table 16. However, given the boundary-treatment sensitivity documented in Section 5.3.4, this pattern should be read as supportive evidence on benchmark-side quote conditions rather than as an additional core finding.

Summary of Findings

The empirical results can be summarised in five points. First, time-to-resolution is the strongest and most stable predictor of the gap: the association is strong, stable across all specifications, and non-linear, with the sharpest widening concentrated in the final three days before resolution. Second, Polymarket trading volume is associated with a reduction in the gap, particularly at longer horizons; the interaction with TTR indicates that active trading is most effective at narrowing disagreement when contracts have more time to run. Third, the relative bid-ask spread of the benchmark options is negatively associated with the gap in the clipped-boundary sample, but this association is not robust to the alternative boundary treatment. We therefore treat it as a secondary, benchmark-side result rather than as part of the core empirical narrative. Fourth, cross-sectional heterogeneity by ticker and contract type is broadly consistent with the TTR narrative, though these patterns are secondary to the main full-sample findings. Fifth, the signed-gap evidence suggests a modest tendency for Polymarket to assign higher probabilities than options across maturity ranges, although this directional asymmetry is descriptive and sensitive to the treatment of extreme-probability observations.

6. Conclusion

This paper studies whether Polymarket contract prices and option-implied probabilities assign similar probabilities to the same equity threshold events, and documents how the gap between them varies with time-to-resolution, trading activity, and benchmark-side quote conditions. Using a panel of 5,376 contract-day observations across eight large-cap US equities, we relate the absolute logit gap to measures of time-to-resolution, trading activity, and liquidity on both sides of the comparison.

Main findings

Two main findings and one secondary result emerge from the analysis. First, time-to-resolution is the strongest predictor of the gap. As contracts approach expiry, the disagreement between Polymarket and option-implied probabilities widens non-linearly, with the sharpest increase in the final three days before resolution. This result is stable across specifications and holds for nearly every underlying equity in the sample, with the exception of NFLX, where the sample is small.

Second, Polymarket trading activity is associated with a smaller gap, especially for longer-horizon contracts. The interaction between volume and time-to-resolution shows that this association is stronger when more time remains before resolution. Near expiry, the time-to-resolution pattern becomes more pronounced, and additional trading activity is less strongly associated with gap reduction.

The secondary result concerns option-side spreads. In the main specification, the relative bid-ask spread of the options used to construct the benchmark is negatively associated with the gap. However, this result does not hold in the alternative sample where boundary observations are removed. We therefore treat it as a secondary finding and do not draw strong conclusions from it. The signed-gap evidence also suggests a modest directional asymmetry, but this pattern is descriptive and not central to the paper.

Economic interpretation

Taken together, the results suggest that Polymarket prices and option-implied probabilities do not move as close substitutes in this setting. The gap is especially large near resolution, when small movements in the underlying stock price can lead to large changes in the option-implied probability. Polymarket prices, by contrast, depend on discrete trading activity and may adjust at a different pace.

This interpretation is consistent with differences in market structure, participant base, and trading frictions across the two markets. However, the regressions do not directly identify the mechanism behind the gap. The results should therefore be read as evidence of systematic cross-market divergence, rather than as proof of arbitrage opportunities or objective pricing errors.

Implications

The results contribute to the growing literature comparing prediction-market prices with derivative-based probability measures. They show that the Polymarket–options gap is not random: it varies systematically with time-to-resolution and, to a lesser extent, with Polymarket trading activity. This suggests that the relationship between prediction markets and options markets depends on the life cycle of the contract, rather than being constant across all maturities.

From an academic perspective, the findings show that prediction-market prices can be informative while still differing from option-implied benchmarks in systematic ways. This distinction is important because prediction markets and options markets do not necessarily reflect the same type of probability. Polymarket prices reflect the price of a binary event contract traded by prediction-market participants, while option-implied probabilities are risk-neutral measures extracted from listed option prices. The gap between the two should therefore be interpreted as a cross-market difference in pricing, not as direct evidence that one market is always more accurate than the other.

From a practical perspective, the results suggest that users of Polymarket prices should pay attention to contract maturity and trading activity. Prices appear closer to option-implied probabilities when contracts have more time remaining and when trading activity is higher. Near resolution, the two measures diverge more strongly, which means that prediction-market prices may be less directly comparable to option-implied benchmarks in the final days before settlement. This is relevant for interpreting cross-market discrepancies, and for researchers or practitioners who use prediction-market prices as inputs in forecasting or valuation exercises.

Limitations

Several limitations are worth noting. First, the option-implied probability is a risk-neutral measure and may embed a risk premium. For this reason, differences between Polymarket prices and option-implied probabilities should not be interpreted mechanically as evidence that either market is incorrectly priced.

Second, we treat American-style equity options as European options when applying the Breeden–Litzenberger framework. This is a standard approximation for short-dated calls on low-dividend stocks, but it may still introduce some measurement error.

Third, the option-implied benchmark is constrained by the availability of LSEG option quotes and by the discrete strike grid. The fixed $\Delta = 5$ bracket is the closest interval that is consistently available across the sample. This approximation is less precise for contracts that are very far in- or out-of-the-money, and for contracts close to expiry. As a result, some observations cannot be retained because the option-implied probability cannot be constructed with a valid probabilistic interpretation. This limits the external validity of the results in the most extreme parts of the contract space.

Fourth, the Polymarket probability is constructed using daily VWAPs, and no-trade days are filled using the most recent observed price. This carry-forward rule is transparent and avoids dropping inactive days, but it also means that some Polymarket prices are stale. As a result, measured gaps on no-trade days may understate the true degree of disagreement between the two markets.

Fifth, the sample is restricted to eight large-cap US equities over a five-month period. The results should therefore be interpreted as evidence for this specific setting, rather than as a general statement about all prediction markets or all option-implied probability benchmarks.

Finally, the regression framework identifies systematic associations, not causal mechanisms. The results show that the gap is strongly related to time-to-resolution and, to a lesser extent, trading activity, but they do not directly establish why these relationships arise.

Future research

Several extensions would be useful. First, applying the same framework to other types of contracts would help assess whether the time-to-resolution pattern documented in this paper is specific to equity threshold events. Similar comparisons could be conducted for macroeconomic events, interest-rate thresholds, cryptocurrency prices, or index-level contracts. This would show whether the widening of the gap near resolution is a general feature of prediction-market contracts or a pattern that is specific to individual equities.

Second, future work could use higher-frequency data. The daily frequency used in this paper is a practical choice given the structure of the available data, but it cannot fully capture the timing difference between a continuously traded prediction market and exchange-traded options quoted during US market hours. Intraday data would allow a more precise comparison between Polymarket price updates, stock price movements, and changes in option-implied probabilities, especially in the final days before resolution.

Third, future research could study whether the observed gaps are tradable in practice. This paper measures cross-market divergences, but it does not test arbitrage profitability. A full trading analysis would need to account for transaction costs, bid-ask spreads on both sides, execution timing, market access, collateral requirements, and the practical difficulty of trading Polymarket contracts and listed options simultaneously. This would be necessary before interpreting the observed gaps as exploitable opportunities.

Fourth, access to a more complete option database could allow narrower strike brackets or interpolation across strikes, making it possible to test whether the main results are sensitive to the benchmark construction.

Finally, a longer sample period would help test whether the results remain stable as prediction markets develop. As Polymarket lists more equity contracts and trading activity grows, future work could examine whether the gap narrows over time and whether market participants learn to align prediction-market prices more closely with option-implied benchmarks.

Closing remark

This paper provides a systematic panel comparison of Polymarket equity-threshold probabilities and individual stock option-implied probabilities. The results show that the gap between the two markets is strongly related to time-to-resolution and, more modestly, to Polymarket trading activity. The clearest pattern is that disagreement widens as contracts approach resolution, especially in the final three days before settlement.

The paper should therefore be read as an initial step in understanding how prediction markets and options markets differ when pricing similar binary equity events. The evidence does not show that either market is objectively wrong, nor that the observed gaps are exploitable after transaction costs and timing frictions. It does show that the two markets do not always move in line with each other, and that these differences vary systematically with maturity and trading activity. As prediction markets continue to grow, understanding when their prices align with established financial-market benchmarks will become increasingly important.

References

Yacine Aït-Sahalia and Andrew W. Lo. Nonparametric estimation of state-price densities implicit in financial asset prices. *Journal of Finance*, 53(2):499–547, 1998.

Pat Akey, Vincent Grégoire, Nicolas Harvie, and Charles Martineau. Who wins and who loses in prediction markets? evidence from polymarket. Working Paper, ESSEC Business School and University of Toronto, 2026.

Joyce Berg, Robert Forsythe, Forrest Nelson, and Thomas Rietz. Results from a dozen years of election futures markets research. In Charles R. Plott and Vernon L. Smith, editors, *Handbook of Experimental Economics Results*, volume 1, pages 742–751. Elsevier, Amsterdam, 2008.

Douglas T. Breeden and Robert H. Litzenberger. Prices of state-contingent claims implicit in option prices. *Journal of Business*, 51(4):621–651, 1978.

Ivan A. Canay. A simple approach to quantile regression for panel data. *Econometrics Journal*, 14(3): 368–386, 2011.

Stephen Figlewski. Estimating the implied risk neutral density. In Tim Bollerslev, Jeffrey Russell, and Mark Watson, editors, *Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle*. Oxford University Press, Oxford, 2010.

Roberto Gomez Cram, Yunhan Guo, Theis Ingerslev Jensen, and Howard Kung. Financial prediction markets: A new measure of earnings expectations. Working Paper, London Business School and Yale University. Available at SSRN: `https://ssrn.com/abstract=5933475`, 2025.

Roberto Gomez Cram, Yunhan Guo, Theis Ingerslev Jensen, and Howard Kung. Prediction market accuracy: Crowd wisdom or informed minority? Working Paper, London Business School and Yale University. Available at SSRN: `https://ssrn.com/abstract=6617059`, 2026.

Jens Carsten Jackwerth and Mark Rubinstein. Recovering probability distributions from option prices. *Journal of Finance*, 51(5):1611–1631, 1996.

Daniel Kahneman and Amos Tversky. Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–292, 1979.

Charles F. Manski. Interpreting the predictions of prediction markets. *Economics Letters*, 91(3):425–429, 2006.

Robert C. Merton. Theory of rational option pricing. *Bell Journal of Economics and Management Science*, 4(1):141–183, 1973.

Stephen Nickell. Biases in dynamic models with fixed effects. *Econometrica*, 49(6):1417–1426, 1981.

Paradigm. Polymarket volume is being double-counted, 2025. Available at: `https://www.paradigm.xyz/2025/12/polymarket-volume-is-being-double-counted`.

Phemex News. Polymarket and Kalshi hit record $40b trading volume in 2025, 2025. Available at: `https://phemex.com/news/article/polymarket-and-kalshi-achieve-record-40b-trading-volume-in-2025-49434`.

Felix Reichenbach and Martin Walther. Exploring decentralized prediction markets: Accuracy, skill, and bias on Polymarket. Working Paper, Technische Universität Berlin and German International University Berlin. Available at SSRN: `https://ssrn.com/abstract=5910522`, 2025.

Sacra. Polymarket, 2025. Available at: `https://sacra.com/c/polymarket/`.

Andrei Shleifer and Robert W. Vishny. The limits of arbitrage. *Journal of Finance*, 52(1):35–55, 1997.

Erik Snowberg and Justin Wolfers. Explaining the favourite-longshot bias: Is it risk-love or misperceptions? *Journal of Political Economy*, 118(4):723–746, 2010.

Stevens Financial Strategy Club. Event contract mispricing via options-implied probabilities, 2025. Stevens Institute of Technology.

Paul C. Tetlock. Liquidity and prediction market efficiency. Working Paper, Columbia University, 2008.

Amos Tversky and Daniel Kahneman. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4):297–323, 1992.

Justin Wolfers and Eric Zitzewitz. Interpreting prediction market prices as probabilities. Working Paper 12200, National Bureau of Economic Research, 2006.

Tables

Table 1. Polymarket data cleaning

Step	Description	Obs. before	Obs. after	Dropped
1. Weekends & holidays	NYSE calendar filter	19,127	15,442	3,685
2. Missing fields	Drop rows with missing essentials	15,442	15,440	2
3. Duplicate contract-days	One observation per contract-day	15,440	13,452	1,988
4. Outside active life	Drop obs. before first trade or after expiry	13,452	13,429	23
5. Non-positive strikes	Drop $K \leq 0$	13,429	13,368	61
6. Non-multiple-of-five strikes	Option chain compatibility ($K \pm 5$)	13,368	12,225	1,143
7. Non-Friday expiries	Option expiry alignment	12,225	9,097	3,128
Final cleaned panel	2,165 unique contracts		**9,097**	

Notes: Each row reports the number of contract-day observations before and after the indicated restriction. The cleaned panel covers 2,165 unique contracts across 8 tickers over the period 4 November 2025 to 27 March 2026. The cleaning pipeline begins from an intermediate panel of 19,127 rows.

Table 2. Final matched estimation sample

Ticker	Contract-day obs.	Share of sample (%)	No-trade obs. (% of ticker total)
AAPL	830	15.4	25.4
GOOGL	880	16.4	19.9
META	861	16.0	21.6
MSFT	918	17.1	24.6
NFLX	110	2.0	45.5
NVDA	715	13.3	12.4
PLTR	163	3.0	28.8
TSLA	899	16.7	18.2
Total	5,376	100.0	21.4

Notes: The sample covers 4 November 2025 to 27 March 2026. All contracts resolve on a Friday. No-trade observations are contract-days inserted during panel expansion on which Polymarket recorded zero transactions; the Polymarket probability is carried forward from the most recently observed daily VWAP on these days. The no-trade fraction is reported as a percentage of each ticker's total contract-day observations. Total no-trade observations across the full sample: 1,148 (21.4% of all contract-day observations). NFLX exhibits the highest no-trade fraction (45.5%), consistent with its limited trading activity on Polymarket documented throughout the analysis.

Table 3. Contract-level characteristics of the estimation sample

Characteristic	**Value**
Panel A: Sample composition	
Total contract-day observations	5,376
Unique contracts	971
Underlying equities	8
Panel B: Observations per contract	
Mean	5.5
Median	4.0
Standard deviation	6.0
Minimum	1
Maximum	23
Panel C: Contract active life (calendar days)	
Mean	6.4
Median	3.0
Standard deviation	8.7
Maximum	32
Share with active life ≤ 1 day (%)	24.1
Share with active life 2–5 days (%)	57.1
Share with active life 6–14 days (%)	3.4
Share with active life 15+ days (%)	15.4
Panel D: Contract type (share of obs., %)	
Month-end	53.3
Weekly	42.0
Daily	4.7
Panel E: Moneyness $\lvert \ln(K/S) \rvert$ (share of obs., %)	
Near at-the-money ($\lvert \ln(K/S) \rvert < 0.05$)	33.0
Moderate OTM/ITM ($0.05 \leq \lvert \ln(K/S) \rvert < 0.15$)	45.4
Deep OTM/ITM ($\lvert \ln(K/S) \rvert \geq 0.15$)	21.6
Panel F: No-trade observations	
No-trade contract-days	1,148
Share of total obs. (%)	21.4

Notes: Full matched estimation sample, $N = 5{,}376$ contract-day observations across 971 unique contracts, covering 4 November 2025 to 27 March 2026. Contract active life is defined as the number of calendar days from the first recorded trade to the resolution date. Moneyness is measured as $\lvert \ln(K_i/S_t) \rvert$ where K_i is the contract strike and S_t is the prevailing stock price. No-trade observations are rows inserted during panel expansion on which Polymarket recorded zero transactions.

Table 4. Descriptive statistics

Variable	N	Mean	Median	Std. Dev.	Min	Max	Corr.(Y)
Panel A: Summary statistics							
$Y_{i,t}$ (logit gap, absolute)	5,376	1.293	1.025	1.076	0.000	8.439	+1.000
ln(Volume)	5,376	3.202	3.214	2.656	0.000	13.121	+0.084
ln(TTR)	5,376	1.609	1.355	1.100	−0.182	3.493	−0.440
| ln(Moneyness)|	5,376	0.099	0.080	0.081	0.000	0.517	−0.004
rel_opt_spread	5,376	0.053	0.020	0.117	0.004	0.833	−0.163
n_trades	5,376	6.364	2.000	13.683	0.000	259.0	+0.033
Panel B: Mean gap by underlying equity							
AAPL	830	1.267	1.074	0.993			
GOOGL	880	1.444	1.202	1.132			
META	861	1.245	0.971	1.095			
MSFT	918	1.271	1.065	1.041			
NFLX	110	0.443	0.225	0.590			
NVDA	715	1.517	1.434	1.096			
PLTR	163	1.275	0.908	1.125			
TSLA	899	1.168	0.824	1.058			

Panel C: Pairwise correlations between regressors

	ln(Vol)	ln(TTR)	| ln(Mon)|	spread	n_trades
ln(Volume)	1.000				
ln(TTR)	−0.150	1.000			
| ln(Moneyness)|	−0.341	+0.186	1.000		
rel_opt_spread	−0.278	−0.077	+0.453	1.000	
n_trades	+0.494	−0.186	−0.121	−0.103	1.000

Notes: Full matched estimation sample, $N = 5{,}376$, covering November 2025 to March 2026. Panel A: Corr.(Y) denotes the Pearson correlation with $Y_{i,t}$; rel_opt_spread is winsorised at the 1st–99th percentile before all calculations. Panel C: pairwise Pearson correlations between the main regressors. The modest negative correlation between ln(Volume) and ln(TTR) (−0.150) indicates that multicollinearity between these two key regressors is not a material concern.

Table 5. Baseline and preferred specification

	(1) Volume only	(2) Trades only	(3) Preferred
ln(Volume)	-0.018^{**}		-0.019^{**}
	(0.037)		(0.032)
n_trades		-0.002^{**}	
		(0.043)	
ln(TTR)	-0.780^{***}	-0.772^{***}	-0.785^{***}
	(0.000)	(0.000)	(0.000)
$\vert\ln(\text{Moneyness})\vert$	-0.839	-0.705	-0.390
	(0.287)	(0.373)	(0.630)
rel_opt_spread			-1.078^{***}
			(0.000)
N	5,376	5,376	5,376
Contract FE	Yes	Yes	Yes
Date FE	Yes	Yes	Yes

Notes: Dependent variable is $Y_{i,t} = |\text{logit}(\hat{p}_{i,t}^{OPT}) - \text{logit}(p_{i,t}^{PM})|$. All specifications include contract and calendar-date fixed effects. Standard errors are clustered at the contract level; p-values are reported in parentheses. Column (3) is the preferred specification. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 6. Robustness checks

	Panel A: Alt. liquidity proxies			*Panel B: Alt. FE*		*Panel C*
	No trade	Zero move	Low vol	Contract FE	Ticker FE	Signed gap
	(1)	(2)	(3)	(4)	(5)	(6)
ln(Volume)				-0.017^{**}	-0.010	-0.127^{***}
				(0.026)	(0.215)	(0.000)
ln(TTR)	-0.775^{***}	-0.775^{***}	-0.767^{***}	-0.401^{***}	-0.648^{***}	0.043
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.637)
$\lvert\ln(\text{Moneyness})\rvert$	-0.413	-0.413	-0.520	-0.533	2.678^{***}	-1.024
	(0.599)	(0.600)	(0.512)	(0.483)	(0.000)	(0.324)
rel_opt_spread				-1.098^{***}	-1.921^{***}	-0.389^{*}
				(0.000)	(0.000)	(0.068)
No-trade dummy	-0.114^{***}					
	(0.001)					
Zero-move dummy		-0.110^{***}				
		(0.001)				
Low vol Q25 dummy			-0.044			
			(0.191)			
N	5,376	5,376	5,376	5,376	5,376	5,376
Contract FE	Yes	Yes	Yes	Yes	No	Yes
Ticker FE	No	No	No	No	Yes	No
Date FE	Yes	Yes	Yes	No	Yes	Yes

Notes: Dependent variable in Columns (1)–(5) is $Y_{i,t} = |\text{logit}(\hat{p}^{OPT}_{i,t}) - \text{logit}(p^{PM}_{i,t})|$. Column (6) uses the signed gap, $\text{logit}(\hat{p}^{OPT}_{i,t}) - \text{logit}(p^{PM}_{i,t})$. All specifications are estimated on the full sample. Standard errors are clustered at the contract level; p-values are reported in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 7. Robustness to boundary treatment: clipped-boundary vs drop-boundary

	Clipped-boundary sample (N = 5,376)	Drop-boundary sample (N = 4,055)
ln(Volume)	-0.019^{**}	-0.019^{**}
	(0.032)	(0.026)
ln(TTR)	-0.785^{***}	-0.751^{***}
	(0.000)	(0.000)
$\lvert\ln(\text{Moneyness})\rvert$	-0.390	$+2.220^{***}$
	(0.630)	(0.006)
rel_opt_spread	-1.078^{***}	$+0.302$
	(0.000)	(0.645)
N	5,376	4,055
Contract FE	Yes	Yes
Date FE	Yes	Yes

Notes: Preferred specification with contract and date fixed effects. The clipped-boundary sample clips both p^{PM} and $\hat{p}^{OPT}$ to $[0.01, 0.99]$ before applying the logit transformation. The drop-boundary sample removes observations outside $[0.01, 0.99]$ entirely. Standard errors are clustered at the contract level; p-values are reported in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 8. Preferred specification by underlying equity

	AAPL	GOOGL	META	MSFT	NFLX	NVDA	PLTR	TSLA
ln(Volume)	−0.011	−0.014	−0.019	−0.040**	−0.035	−0.005	+0.003	−0.028
	(0.533)	(0.531)	(0.456)	(0.031)	(0.740)	(0.810)	(0.976)	(0.151)
ln(TTR)	−0.607***	−0.671***	−0.706***	−0.692***	−0.178	−0.646***	−1.352***	−1.407***
	(0.002)	(0.000)	(0.002)	(0.000)	(0.274)	(0.000)	(0.004)	(0.000)
$\lvert \ln(\text{Moneyness}) \rvert$	−1.538	3.944**	0.858	−5.078***	—	−3.284	4.391	1.750
	(0.558)	(0.034)	(0.674)	(0.008)		(0.160)	(0.311)	(0.379)
rel_opt_spread	−0.874	−1.780***	−1.153**	−1.736*	+0.348	+0.703	+4.699	+3.943**
	(0.125)	(0.000)	(0.011)	(0.092)	(0.599)	(0.641)	(0.741)	(0.020)
Mean Y	1.267	1.444	1.245	1.271	0.443	1.517	1.275	1.168
N	830	880	861	918	110	715	163	899

Notes: Each column estimates the preferred specification on the subsample of the indicated ticker. Contract and date fixed effects throughout. Standard errors are clustered at the contract level; p-values are reported in parentheses. Moneyness is omitted where there is insufficient within-subsample variation. Ticker-level results should be interpreted as exploratory heterogeneity checks. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 9. Preferred specification by equity group

	Mega-cap (AAPL, MSFT, GOOGL, META) (1)	High-volatility (NVDA, TSLA, PLTR) (2)
ln(Volume)	-0.021^{*}	-0.016
	(0.052)	(0.285)
ln(TTR)	-0.670^{***}	-1.015^{***}
	(0.000)	(0.000)
$\lvert \ln(\text{Moneyness}) \rvert$	-0.622	-0.128
	(0.553)	(0.920)
rel_opt_spread	-1.035^{***}	$+1.208$
	(0.000)	(0.274)
N	3,489	1,777

Notes: Groups are defined descriptively to separate large technology platforms from a higher-volatility group during the sample period. NFLX is excluded from both groups due to its small sample size (N = 110). The grouping is exploratory and should not be interpreted as a formal classification of the underlying equities. Contract and date fixed effects are included throughout. Standard errors are clustered at the contract level; p-values are reported in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 10. Gap and selected regression results by contract type

	Daily	Weekly	Month-end
Panel A: Descriptive			
Mean Y	1.982	1.595	0.994
Median Y	1.879	1.437	0.762
N	255	2,257	2,864
Panel B: Selected regression coefficients			
ln(TTR)	(not identified)	−0.123	(not identified)
		(0.669)	
ln(Volume)	−0.031	+0.020	-0.050^{***}
	(0.750)	(0.148)	(0.000)
rel_opt_spread	−4.417	−0.183	-1.240^{***}
	(0.528)	(0.688)	(0.000)

Notes: Panel A reports descriptive statistics by contract type. Panel B reports selected coefficients from contract-type subsample regressions. Contract type is mechanically related to time-to-resolution, so these estimates should be interpreted as descriptive evidence consistent with the maturity pattern rather than as independent contract-type effects. Standard errors are clustered at the contract level; p-values are reported in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 11. Preferred specification with lagged dependent variable

	Preferred	With lagged Y
	(1)	(2)
$Y_{i,t-1}$		0.447^{***}
		(0.000)
ln(Volume)	-0.019^{**}	-0.015
	(0.032)	(0.065)
ln(TTR)	-0.785^{***}	-0.596^{***}
	(0.000)	(0.000)
$\lvert \ln(\text{Moneyness}) \rvert$	-0.390	-0.760
	(0.630)	(0.231)
rel_opt_spread	-1.078^{***}	-0.728^{***}
	(0.000)	(0.000)
N	5,376	4,405

Notes: Contract and date fixed effects throughout. Standard errors are clustered at the contract level; p-values are reported in parentheses. The lagged specification loses one observation per contract. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 12. Non-linear TTR effects: bin analysis

TTR bin	N	Mean Y	Median Y	Estimate	p-value
0–3 days	2,138	1.807	1.735	$+0.947^{***}$	0.000
4–7 days	1,062	1.171	0.939	$+0.414^{***}$	0.000
8–14 days	666	0.950	0.793	(not separately identified)	
15+ days	1,510	0.802	0.613	(reference)	
Other regressors in the bin specification					
ln(Volume)	−0.011 (0.234)				
rel_opt_spread	-1.058^{***} (0.000)				
$\lvert \ln(\text{Moneyness}) \rvert$	−0.220 (0.788)				

Notes: The regression replaces ln(TTR) with maturity-bin dummies; the reference category is 15+ days. The 8–14 day dummy is not separately identified in the two-way fixed effects model due to limited within-contract variation at that maturity. Contract and date fixed effects are included throughout. Standard errors are clustered at the contract level. The column labelled p-value reports coefficient p-values for the maturity-bin indicators. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 13. Mean logit gap by TTR bin and underlying equity

TTR bin	AAPL	GOOGL	META	MSFT	NFLX	NVDA	PLTR	TSLA
0–3 days	1.754	1.972	1.870	1.743	0.589	1.944	1.688	1.775
4–7 days	1.260	1.361	1.099	1.107	0.357	1.279	0.920	1.070
8–14 days	0.849	1.221	0.825	0.874	0.254	1.086	1.001	0.949
15+ days	0.846	0.922	0.719	0.864	0.339	1.107	0.766	0.547

Notes: Mean of $Y_{i,t}$ for each ticker–TTR bin combination. Full estimation sample (N = 5,376).

Table 14. Mean signed gap by TTR bin

TTR bin	N	Mean signed gap	Median signed gap
0–3 days	2,138	−0.197	−0.038
4–7 days	1,062	−0.116	0.000
8–14 days	666	−0.096	−0.011
15+ days	1,510	−0.226	−0.136

Notes: Signed gap = $\text{logit}(\hat{p}_{i,t}^{OPT}) - \text{logit}(p_{i,t}^{PM})$. A negative value indicates that Polymarket assigns a higher probability to the threshold event than options. Raw bin means; not regression-adjusted. Full estimation sample (N = 5,376).

Table 15. Volume–TTR interaction

	Preferred (1)	With interaction (2)
ln(Volume)	-0.019^{**}	$+0.041^{***}$
	(0.032)	(0.003)
ln(TTR)	-0.785^{***}	-0.573^{***}
	(0.000)	(0.000)
ln(Volume) × ln(TTR)		-0.043^{***}
		(0.000)
| ln(Moneyness)|	-0.390	-0.232
	(0.630)	(0.776)
rel_opt_spread	-1.078^{***}	-0.750^{***}
	(0.000)	(0.000)
N	5,376	5,376

Notes: Contract and date fixed effects are included throughout. Standard errors are clustered at the contract level; p-values are reported in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 16. Quantile regressions

	Median (1)	75th pct (2)	90th pct (3)
ln(Volume)	-0.017^{***}	-0.018^{**}	-0.027^{**}
	(0.000)	(0.011)	(0.015)
ln(TTR)	-0.370^{***}	-0.522^{***}	-0.631^{***}
	(0.000)	(0.000)	(0.000)
$\lvert \ln(\text{Moneyness}) \rvert$	-0.290	-1.174^{**}	-1.154
	(0.396)	(0.037)	(0.200)
rel_opt_spread	-0.928^{***}	-0.764^{***}	-0.671^{**}
	(0.000)	(0.001)	(0.035)
N	5,376	5,376	5,376

Notes: Regressions use within-contract demeaned variables as an approximation for contract fixed effects. p-values are reported in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Figures

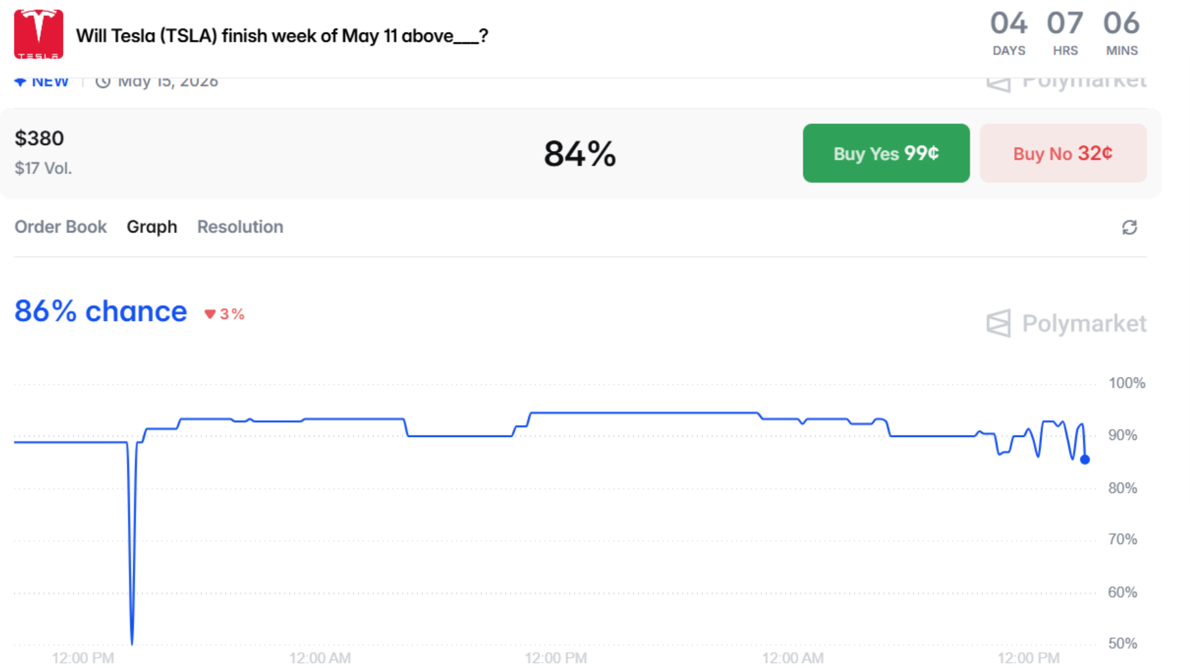

Figure 1. Illustrative example of a Polymarket equity threshold contract. The contract asks whether Tesla (TSLA) will finish the week of 11 May above $380, with resolution on 15 May 2026. The y-axis shows the implied probability of the "Yes" outcome over the life of the contract. The flat segments reflect periods without trading activity, during which the probability is carried forward from the last observed price, a feature of the panel structure described in Section 3.1.5. This contract is shown for illustrative purposes; it postdates the estimation sample, which covers November 2025 to March 2026.

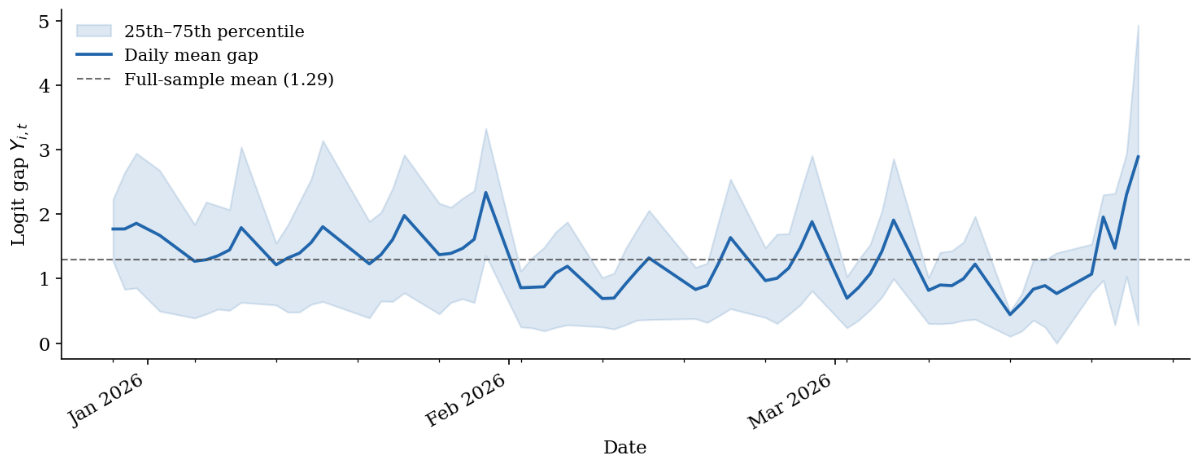

Figure 2. Evolution of the daily cross-sectional mean of the logit gap $Y_{i,t}$, together with the interquartile range, over the period with regular matched-sample coverage. The gap fluctuates around its full-sample mean without a clear time trend over most of the plotted period. The noticeable increase at the end of the period reflects a concentration of observations with larger gaps, consistent with the higher dispersion visible in the interquartile band.

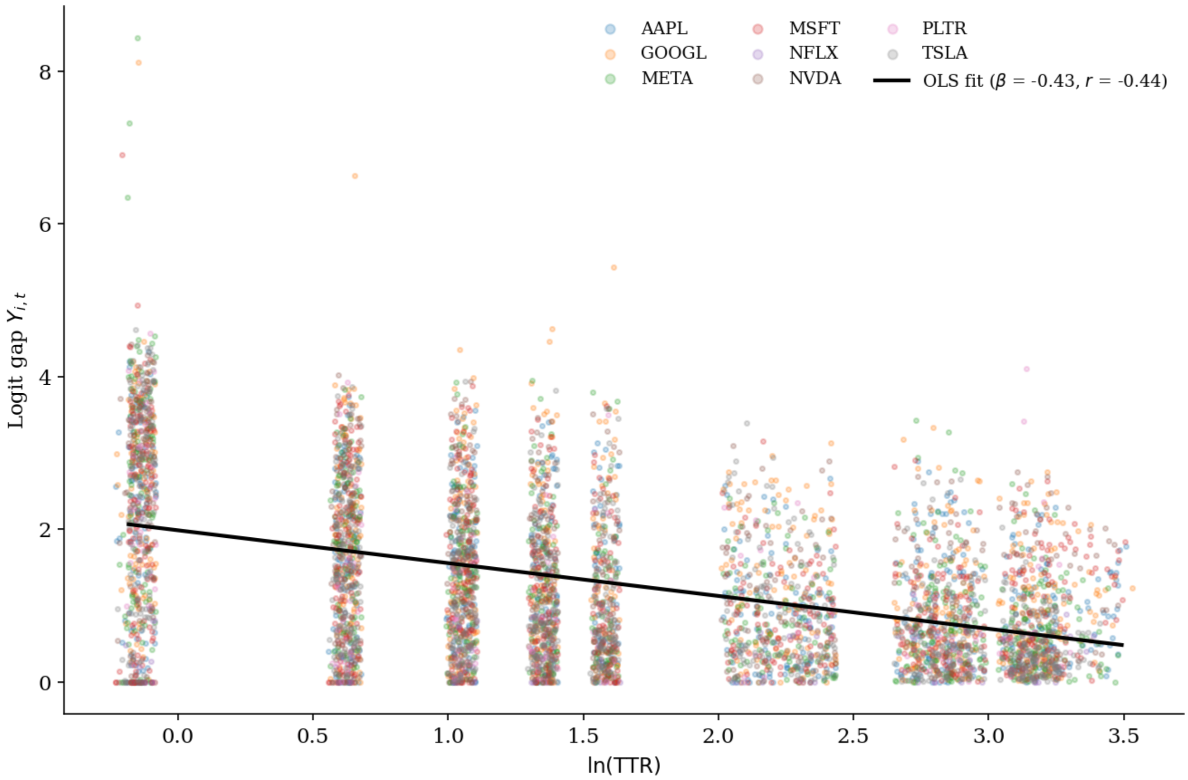

Figure 3. Relationship between the logit gap $Y_{i,t}$ and $\ln(\text{TTR})$ across all contract-day observations. The negative slope indicates that disagreement increases as contracts approach resolution. The dispersion of the gap is substantially larger at low values of $\ln(\text{TTR})$, showing that extreme disagreements are concentrated near expiry. The vertical clustering reflects the limited set of repeated time-to-resolution values generated by the daily panel structure.

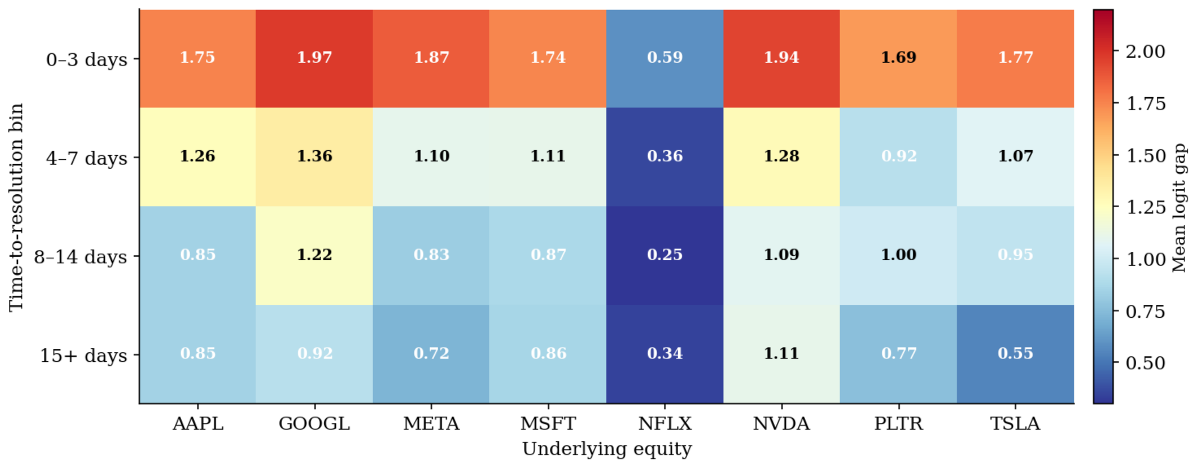

Figure 4. Mean logit gap across time-to-resolution bins and underlying equities. The gap declines systematically as maturity increases for most tickers, indicating that near-expiry observations exhibit larger disagreement. While the pattern is consistent across equities, the magnitude varies across names, with some tickers displaying stronger near-expiry effects than others.

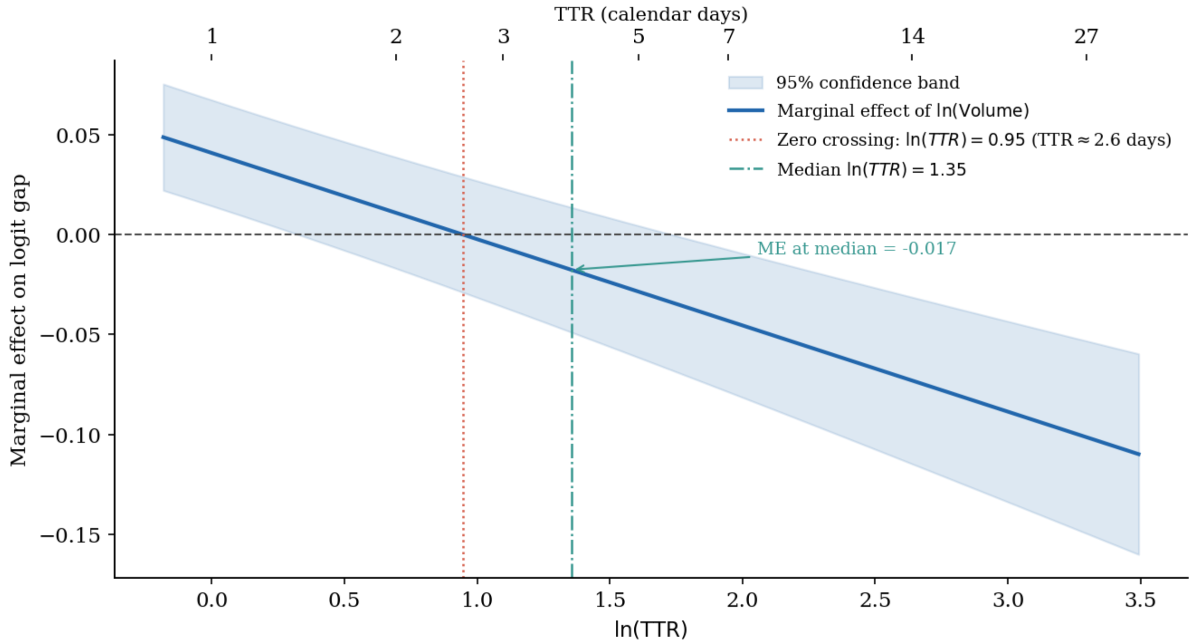

Figure 5. Marginal effect of ln(Volume) on the logit gap as a function of ln(TTR), together with its 95% confidence band. The marginal effect becomes more negative as time-to-resolution increases, indicating that trading activity is more strongly associated with a reduction in the gap at longer horizons. The effect approaches zero and may reverse close to expiry, consistent with different dynamics in the final days before resolution.

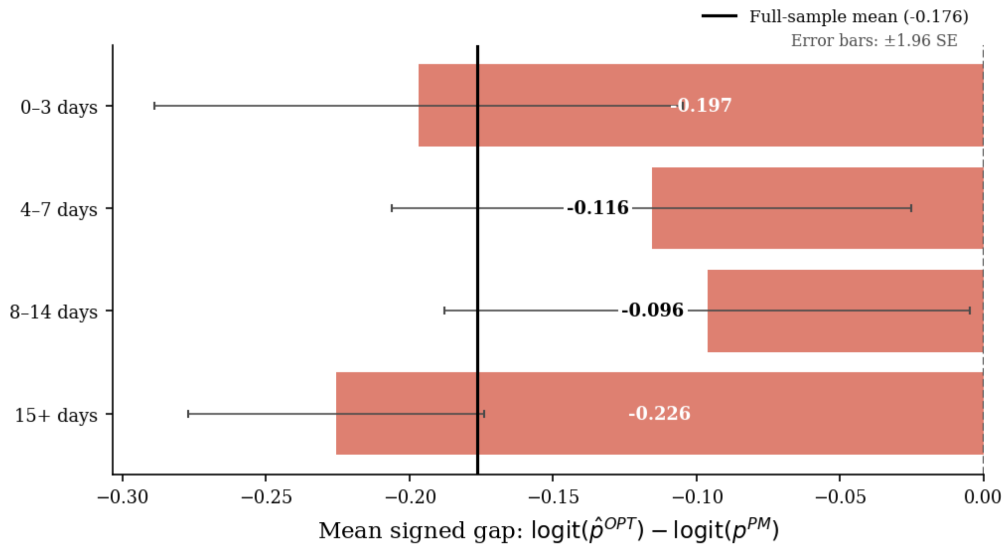

Figure 6. Mean signed gap across time-to-resolution bins, together with confidence intervals. The signed gap is negative across all maturity ranges, indicating that Polymarket assigns higher probabilities than options on average in this sample. The magnitude of this asymmetry is relatively modest and does not appear to be concentrated in a specific maturity range, with some bins not statistically distinguishable from zero.

A. Description of Dataset Variables

Table A1 summarises the main variables included in the final matched estimation panel, covering both the Polymarket and option-implied probability sides of the dataset.

Table A1. Description of Panel Variables

Variable	**Description**
Panel A: Contract identifiers and metadata	
`Condition_Id`	Unique identifier for each Polymarket contract, corresponding to the `conditionId` field in the Gamma API output. Used as the entity identifier for contract fixed effects in the panel regressions.
`Date_clean`	Observation date for the contract-day row, formatted as a datetime object. Forms the time dimension of the panel.
`Ticker`	Underlying stock ticker associated with the contract (e.g. `AAPL`, `GOOGL`, `MSFT`, `META`, `TSLA`, `NVDA`, `NFLX`, `PLTR`).
`Contract Type`	Type of contract. Possible values are `daily`, `weekly`, and `month_end`, reflecting the resolution horizon of the contract.
`strike`	Threshold price level K used in the contract question. Must be a multiple of five to be compatible with the LSEG option strike grid.
`question`	Full Polymarket market question as retrieved from the Gamma API. Retained for audit purposes and manual verification.
`Date_Expiration_Clean`	Contract resolution date, formatted as a datetime object. All retained contracts resolve on a Friday.
Panel B: Polymarket probability measures	
`p_vwap`	Volume-weighted average YES price of the day, computed across all transactions in the contract-day. This is the main Polymarket probability measure $p_{i,t}^{PM}$ used throughout the analysis. On no-trade days inserted during panel expansion, it is carried forward from the most recently observed daily VWAP.

Variable	Description
`p_trim`	Clipped version of `p_vwap`, constrained to the interval $[0.01, 0.99]$ before the logit transformation to avoid infinite values. The variable name is retained from the original codebase, but the operation is probability clipping rather than sample trimming. Observations outside this range are clipped in the main clipped-boundary sample and removed in the drop-boundary robustness sample.
`logit_p`	Logit transformation of the clipped Polymarket probability: $$\text{logit}(p_{i,t}^{PM}) = \ln\left(\frac{\texttt{p\_trim}}{1 - \texttt{p\_trim}}\right)$$
Panel C: Option-implied probability measures	
`p_OPT`	Option-implied probability $\hat{p}_{i,t}^{OPT}$, constructed from the finite-difference approximation of Breeden and Litzenberger (1978) applied to call mid-prices at strikes $K_i - 5$ and $K_i + 5$. Observations outside $[0, 1]$ are excluded; those outside $[0.01, 0.99]$ are clipped in the main sample.
`C_mid_Kminus5`	Mid-price of the call option at strike $K_i - 5$, computed as the average of the bid and ask retrieved from LSEG Workspace.
`C_mid_Kplus5`	Mid-price of the call option at strike $K_i + 5$, computed as the average of the bid and ask retrieved from LSEG Workspace.
`rel_opt_spread`	Relative bid-ask spread of the options used to construct $\hat{p}_{i,t}^{OPT}$, defined as the average of the relative spreads at the two bracket strikes: $$\text{rel\_opt\_spread}_{i,t} = \frac{1}{2}\left[\frac{C^{ask}(K-5) - C^{bid}(K-5)}{C^{mid}(K-5)} + \dots\right]$$ $$\left[\dots + \frac{C^{ask}(K+5) - C^{bid}(K+5)}{C^{mid}(K+5)}\right]$$ Captures options market illiquidity on the benchmark side.
`rel_opt_spread_w`	Winsorised version of `rel_opt_spread`, capped at the 1st and 99th percentiles to limit the influence of outliers in the option-spread variable. This is the spread measure used in the regressions.
Panel D: Dependent variables	

Variable	Description
`Logit_basis`	Main dependent variable. Absolute logit gap between the option-implied and Polymarket probabilities: $$Y_{i,t} = \left\| \text{logit}(\hat{p}_{i,t}^{OPT}) - \text{logit}(p_{i,t}^{PM}) \right\|$$ A larger value indicates greater disagreement between the two markets.
`logit_basis_signed`	Signed version of the logit gap: $$SG_{i,t} = \text{logit}(\hat{p}_{i,t}^{OPT}) - \text{logit}(p_{i,t}^{PM})$$ A negative value indicates that Polymarket assigns a higher probability than options. Used in the signed-gap robustness analysis.
`raw_basis`	Raw (non-logit) absolute difference between the two probability measures: $$\|p_{i,t}^{OPT} - p_{i,t}^{PM}\|$$ Used as an alternative dependent variable in robustness checks.
Panel E: Time-to-resolution variables	
`TTR_days`	Continuous calendar time remaining until contract resolution, measured in days. The variable is positive for all retained observations and may be below one on the resolution date, which allows the logarithmic transformation to be defined throughout the estimation sample.
`Log_TTR`	Natural logarithm of continuous time-to-resolution: $$\ln(\texttt{TTR\_days})$$ Central explanatory variable in all main specifications.
`ttr_bin_0to3`	Dummy equal to 1 if `TTR_days` ≤ 3. Used in the TTR bin non-linearity analysis.
`ttr_bin_4to7`	Dummy equal to 1 if $4 \leq$ `TTR_days` ≤ 7. Used in the TTR bin non-linearity analysis.
`ttr_bin_8to14`	Dummy equal to 1 if $8 \leq$ `TTR_days` ≤ 14. Used in the TTR bin non-linearity analysis. Reference category is 15+ days.
Panel F: Polymarket liquidity variables	
`volume_usd_clean`	Cleaned version of raw USD trading volume, with missing values replaced and numeric formatting standardised. Intermediate construction variable used to compute `log_volume_usd`.

Variable	Description
log_volume_usd	Natural logarithm of daily USD trading volume on Polymarket: $$\ln(\texttt{volume\_usd\_clean} + 1)$$ Main Polymarket liquidity proxy in the preferred specification.
n_trades_PM	Number of individual trades recorded for the contract on that date. Used as an alternative Polymarket liquidity proxy to confirm that the volume result is not sensitive to the choice of measure.
no_trade	Dummy equal to 1 if the contract recorded zero trades on that day, including rows inserted during panel expansion. Used as a binary liquidity proxy in Panel A of the robustness checks.
zero_move	Dummy equal to 1 if the Polymarket probability did not change relative to the previous observation: $$\mathbf{1}\{\lvert p^{PM}_{i,t} - p^{PM}_{i,t-1}\rvert = 0\}$$ Used as a binary liquidity proxy in Panel A of the robustness checks.
low_volume_usd_q25	Dummy equal to 1 if the USD trading volume is below the 25th percentile of the sample distribution. Used as a binary liquidity proxy in Panel A of the robustness checks.
vol_x_ttr	Interaction term between log volume and log time-to-resolution: $$\texttt{vol\_x\_ttr} = \ln(\texttt{volume\_usd}) \times \ln(\texttt{TTR\_days})$$ Used in the volume–TTR interaction analysis.
log_illiq_amihud_usd	Winsorised Amihud illiquidity proxy using USD trading volume: $$ILLIQ^{USD}_{i,t} = \frac{\lvert p^{PM}_{i,t} - p^{PM}_{i,t-1}\rvert}{\texttt{volume\_usd}_{i,t}}$$ Included in intermediate specifications as an additional liquidity control. Excluded from the preferred specification due to instability across fixed effects structures.
amihud_missing_due_to_lag	Dummy equal to 1 when the Amihud measure is missing because no lagged observation exists for that contract (i.e. first observation of each contract). Used to define the restricted liquidity subsample ($N = 3,301$) in checks involving the Amihud price-impact measure.

Variable	Description
Panel G: Control variables	
`Abs_log_moneyness`	Absolute value of the log ratio of the strike price to the current stock price: $$\lvert \ln(K_i/S_t) \rvert$$ Controls for how far the contract is from at-the-money. Larger values indicate deeper in- or out-of-the-money contracts, where the finite-difference approximation is less precise.
Panel H: Additional analysis variables	
`Logit_basis_lag`	Lagged dependent variable $Y_{i,t-1}$, constructed by shifting `Logit_basis` by one period within each contract. Used in the persistence analysis to control for serial correlation in the gap.
`is_imputed_no_trade_day`	Dummy equal to 1 if the row was inserted during panel expansion because Polymarket recorded no trades on that day. The variable name is retained from the codebase, but the operation is a carry-forward rule for inserted no-trade rows rather than a statistical imputation model.
`question`	Full Polymarket market question as retrieved from the Gamma API. Retained for audit purposes and manual verification.

B. Python Codebase Structure

This appendix documents the Python codebase used to construct the matched Polymarket–options dataset and to run the empirical analysis reported in the paper. The appendix describes the structure and role of each script. The complete code files are available in the shared OneDrive folder for download or review.

The codebase consists of six scripts: three are dedicated to data collection and cleaning from Polymarket — `Data_Collection_Script_vf.py` queries the Polymarket Gamma and Data APIs to download the full transaction history; `Data_cleaning_and_liquidity_proxies_vf.py` applies the sequential cleaning restrictions and constructs the liquidity proxies; and `No_trade_days_vf.py` expands the panel to include no-trade days with carried-forward prices. The remaining three scripts implement the empirical analysis. `Regression_results_vf.py` runs the main specifications without the Amihud price-impact measure, using the full clipped-boundary sample. `Regression_results_notrim_vf.py` replicates the analysis on the drop-boundary robustness sample, where observations outside $[0.01, 0.99]$ are removed rather than clipped. `Generate_figures_vf.py` produces all figures reported in the paper from the final estimation sample.