Price Discovery and Trading in Modern Prediction Markets*

Hunter Ng Lin Peng Yubo Tao Dexin Zhou

This Version: March 17, 2026

Abstract

This study provides the first evidence on price discovery dynamics across modern prediction markets. We study a unique dataset of common contracts traded on leading prediction markets, Polymarket, Kalshi, PredictIt, and Robinhood, during the period leading up to the 2024 U.S. presidential election. We find that more liquid prediction markets substantially outperform polls in predicting subsequent election results, yet there are significant price disparities across platforms. Polymarket leads Kalshi in price discovery, particularly when liquidity and trading activity are high, implying economically meaningful arbitrage opportunities. Moreover, net order imbalance from large trades strongly predicts subsequent returns, and the market experiencing greater directional order flow from large trades tends to lead price discovery. These results underscore how platform structure, liquidity, and informed trading interact to shape trading and prices.

Keywords: Prediction Markets, Election Betting, Price Discovery, Polymarket, Kalshi, Arbitrage, Law of One Price, Return Prediction.

JEL Classification: D47, G14, D83, C58, H11.

*First draft, November 19, 2025. Hunter Ng, Zicklin School of Business, Baruch College, the City University of New York, hng@gc.cuny.edu; Lin Peng, Zicklin School of Business, Baruch College, the City University of New York, lin.peng@baruch.cuny.edu; Yubo Tao, Faculty of Social Sciences and Faculty of Business Administration, University of Macau, yubotao@um.edu.mo; Dexin Zhou, Zicklin School of Business, Baruch College, the City University of New York, dexin.zhou@baruch.cuny.edu. We thank seminar participants at Baruch College, University of Macau, and Hong Kong Polytechnic University, Jay Cao, Thomas Fujiwara, Sebastiano Manzan, Michael Richter, Yanbin Wu, and Yajun Wang, for their helpful comments.

1 Introduction

Prediction markets are a relatively new and rapidly evolving class of financial markets in which participants trade contracts that pay off based on the outcomes of future events. While these markets were historically small in scale and limited in participation, their scope and economic relevance have expanded markedly in recent years.[1] In particular, the emergence of blockchain-settled platforms, together with judicial clarification of the Wire Act's restrictions on interstate betting communications, has substantially lowered barriers to participation and accelerated market growth.[2]

As a result, modern prediction markets now operate at meaningful scale across multiple continents, attracting both retail traders and sophisticated investors who use these platforms not only for speculation but also as hedging instruments within broader portfolios. Their growing prominence is further reflected in the increased attention from market participants and regulators, driven by the markets' ability to successfully forecast high-profile outcomes.[3] Despite their growing economic and political salience, however, how these markets function and what contributes to price formation and trading dynamics remains poorly understood.

We address this question by studying how market design, regulatory boundaries, and user composition shape price discovery and trading activities in modern prediction markets. Our empirical setting exploits the coexistence of prediction markets that list overlapping contracts but differ sharply in institutional structure and governance,

[1] One of the historically well-known examples is the University of Iowa's political prediction market, originally created to forecast the 1988 presidential election. A range of web-based markets subsequently emerged, operated by firms offering trading or gambling services, with notional volumes ranging from tens of thousands of dollars (e.g., the Iowa markets) to hundreds of millions of dollars (e.g., economic derivatives offered by Goldman Sachs and Deutsche Bank), as well as virtual-currency markets awarding prizes such as a television. See, e.g., Wolfers and Zitzewitz (2004) for an overview.

[2] The 1961 Wire Act prohibited the use of interstate communications for betting to curb organized crime. The DOJ's 2011 opinion limited its scope to sports wagering. The Trump administration's 2019 Office of Legal Counsel (OLC) memorandum questioned that view, and federal courts in 2021 reaffirmed the sports-only interpretation, thereby supporting the expansion of state-regulated online wagering and prediction markets. See Cabot and Cloward (2021).

[3] Recent examples includes the 2024 U.S. Presidential Election, the 2025 New York City Mayoral Election, and the 2025 Nobel Peace Prize winner (Osipovich, 2024; Fernandez, 2025; Osipovich, 2025). Prediction markets have also attracted controversy, as illustrated by suspiciously well-timed bets on the potential downfall of Venezuela's president, Nicolás Maduro, shortly before President Trump ordered a U.S. military strike on January 3, 2026 (Stewart, 2026).

including Polymarket, Kalshi, PredictIt, and Robinhood. For example, U.S. investors remain barred from participating in Polymarket, while non-U.S. users did not have access to Kalshi, a CFTC-regulated and U.S.-based platform, until October 2025.[4] These regulatory and access frictions generate important differences in price discovery and market dynamics, creating a natural laboratory to discuss the extent to which these institutional factors contribute to these market outcomes.

We hypothesize that market design, regulatory boundaries, and user composition jointly shape price discovery across prediction platforms. When identical payoffs trade in segmented venues, price gaps can persist because arbitrage is limited by participation constraints, funding frictions, and restrictions on moving capital across platforms. This logic echoes the location-of-trade effects emphasized by Froot and Dabora (1999) and the segmentation mechanisms documented in cryptocurrency markets by Makarov and Schoar (2020). A related implication from market microstructure theory is that information is incorporated through trading, so venues with deeper liquidity and lower trading costs should adjust more quickly to news (e.g., Kyle, 1985; Glosten and Milgrom, 1985; Hasbrouck, 1995; O'Hara, 1995).

Motivated by these theories, we expect the more liquid and actively traded venue to lead price discovery, and we expect regulatory and access frictions to generate deviations from the law of one price and economically meaningful arbitrage opportunities. Finally, differences in oversight and transparency can affect the scope for informed trading. To the extent that less regulated venues offer greater camouflage for informed traders, large-trade order flow should be particularly informative and help determine which platform incorporates information first.

To facilitate this comparison, we compile a novel transaction-level dataset covering the days leading to the 2024 U.S. presidential election, one of the most liquid events in prediction-market history. We collect data at both daily and intraday frequencies through platform APIs and real-time scraping via virtual servers, covering four major and highly

[4] Polymarket is reportedly seeking to reenter the U.S. market (Beyoud, Natarajan, and White, 2025). Kalshi began offering global access with know-your-customer (KYC) requirements on October 9, 2025 (Kalshi, 2025).

liquid venues: Polymarket, Kalshi, PredictIt, and Robinhood. The presidential election was the single most actively traded outcome on both Kalshi and Polymarket. The total Polymarket volume for the election reached \$3.2 billion.[5] Our transaction-level dataset captures \$1.1 billion in Polymarket transactions over a 15-day interval and \$518 million in Kalshi transactions over the contract's entire duration.

Using this high-frequency, cross-platform dataset from a highly salient and liquid event, we first provide evidence that prediction market returns, particularly on Polymarket, substantially outperform polls in predicting Donald Trump winning the presidential election. Returns on Polymarket election contracts significantly predict subsequent poll levels and poll changes, whereas poll levels and poll changes do not significantly predict Polymarket returns.

However, not all prediction markets excel at price discovery. We document substantial cross-platform discrepancies and find that the law of one price, a central implication of market efficiency, fails to hold in this setting. These pricing wedges give rise to economically meaningful arbitrage opportunities. We then focus on cross-market price discovery and ask which platform incorporates information first and whether variation in liquidity and trading activity, especially the presence of large trades, helps explain the speed and direction of price discovery.

We first document significant and persistent price discrepancies across platforms. Because these contracts pay 1 if the candidate wins and 0 otherwise, prices are bounded between 0 and 1. Contracts favoring Kamala Harris (KH contracts) on Kalshi and PredictIt consistently trade at higher prices than on Polymarket, while the opposite holds for contracts favoring Donald Trump. Specifically, the KH contract price on Kalshi exceeds the Polymarket price on all 33 of overlapping trading days, and on PredictIt, it exceeds Polymarket on 304 out of 306 days. The average price difference is also economically and statistically significant: for KH contracts, Polymarket prices are 3.27 cents lower than on Kalshi; for DJT contracts, Polymarket prices are 2.59 cents higher than on Kalshi and 3.16 cents higher than on PredictIt. Robinhood, despite a shorter

[5] For more details, see Polymarket Analytics (2025b).

trading history, also displays systematically higher KH prices than both Kalshi and Polymarket.

To evaluate the economic significance of these price gaps, we turn to our high-frequency data from Kalshi and Polymarket. At each time interval, we examine the trading price differences across the two platforms and compute the realizable arbitrage value by multiplying the price discrepancies by the contemporaneous trading volume, while avoiding the double-counting of overlapping opportunities. When we adopt a stringent one-second arbitrage interval, we find that total arbitrage profits exceed $\$130,000$ over the 15-day period and remain substantial even after imposing a 2% arbitrage cost. When we lengthen the arbitrage interval to five minutes, potential arbitrage profits rise to over $2.2 million.

Given the significant price discrepancies documented above, we next assess the relative efficiency and speed of price discovery across these platforms. Our analysis focuses on the lead-lag relationship in returns to identify where information is first incorporated into prices. Using 30-minute trading windows, we find a clear asymmetry: Polymarket returns significantly lead those of Kalshi, but not vice versa. Specifically, our VAR analyses indicate that a 1 percentage point return increase on Polymarket corresponds to a subsequent 0.26 percentage point increase on Kalshi's DJT contract returns and a 0.25 percentage point increase for KH contracts within the next 30 minutes. Impulse-response functions confirm this asymmetric relationship, showing that Kalshi responds meaningfully to Polymarket shocks, whereas Polymarket does not respond to shocks originating from Kalshi.[6] Together, these findings point to Polymarket as the dominant venue in the price discovery process.

Given our hypothesis that liquidity facilitates faster incorporation of information into prices, we postulate that Polymarket's leadership in price discovery is associated with its superior liquidity. In market microstructure models, information is incorporated through trading, so venues with deeper liquidity and lower effective trading costs tend to adjust more quickly to news. Accordingly, Polymarket's higher liquidity may attract more

[6]Throughout this analysis, we interpret lead–lag relationships as indicating where prices respond first to common information shocks, rather than as evidence of exogenous causal effects across platforms.

informed traders and support greater risk-bearing by liquidity providers, accelerating price adjustment. To evaluate this mechanism, we split the sample by relative liquidity using the Amihud (2002) measure. We find that Polymarket's lead is concentrated in periods when its liquidity is high relative to Kalshi's. During such periods, a 1 percentage point return increase on Polymarket corresponds to roughly a 0.36 percentage point subsequent increase in Kalshi returns, an economically meaningful magnitude, during high liquidity periods. In contrast, there is no evidence supporting that Polymarket leads Kalshi when its relative liquidity is low.

We further segment the sample by the relative trading-volume ratio, defined as Polymarket volume divided by Kalshi volume. Consistent with the liquidity-based explanation, Polymarket returns significantly lead Kalshi returns only when Polymarket accounts for a larger share of aggregate trading activity. During high relative volume periods, a 1 percentage point increase in Polymarket returns predicts a 0.37 percentage point increase in Kalshi's next-interval returns, as demonstrated in our VAR analysis. These results indicate that price discovery tends to occur on Polymarket during periods when its relative liquidity is higher, consistent with liquidity shaping the speed of price adjustment.

Building on evidence that large investors have stronger incentives to acquire information (Boehmer and Kelley, 2009) and that trade size can help distinguish retail from institutional trading in U.S. equity markets (prior the shift to decimal pricing in 2001, see, e.g. Lee and Radhakrishna 2000), we use trade size as a proxy for informed trading by large investors and examine the contribution of large trades to price discovery. Consistent with anecdotal accounts that "whales" on prediction markets may be highly informed (see, e.g., Osipovich 2024, 2025), we construct separate order-imbalance measures for whale and non-whale trades, defining whale trades as those in the top 1% of the trade-size distribution on each platform-day. We find that net buying by whales strongly and consistently predicts future returns on both Kalshi and Polymarket: a one-standard-deviation increase predicts higher 30-minute returns by 20 to 32 basis points. In contrast, order imbalance from non-whale trades exhibits lower predictive power.

These findings indicate that large, likely informed, trades are strongly linked to price discovery, particularly in highly liquid prediction markets. Importantly, in cross-market comparisons, returns on the platform with greater large-trade net order imbalance tend to lead returns on the other platform, consistent with the view that directional order flow from large trades helps determine where information is incorporated first.

Taken together, these findings speak directly to our central hypothesis that market design, regulatory boundaries, and user composition jointly shape price discovery across prediction platforms. The evidence that Polymarket leads price discovery when liquidity and large-trade activity are higher is consistent with the view that information is incorporated through trading and that venues with deeper liquidity adjust more quickly. At the same time, the persistent price dispersion across platforms highlights the role of regulatory and access frictions in limiting arbitrage and allowing deviations from the law of one price. Finally, the strong predictive content of large-trade order flow suggests that differences in platform oversight and trader composition affect the extent to which informed trading is revealed in prices. Together, these results underscore how institutional features of prediction markets influence both the speed and the location of information revelation.

Several institutional differences across platforms offer plausible explanations for the observed variation in liquidity and price discovery. Polymarket's blockchain-based architecture, 24/7 trading, rapid capital recycling, and absence of binding position limits likely support deeper liquidity and more aggressive informed trading, facilitating faster information incorporation. By contrast, Kalshi's status as a CFTC-regulated Designated Contract Market, while enhancing legal clarity and investor protection, entails stricter onboarding and settlement through traditional financial infrastructure, which may slow capital reallocation and limit immediacy during high-volume periods. PredictIt and Robinhood impose even tighter constraints, through position caps, fees, or simplified trading mechanisms, that restrict depth and the scale of informed participation. At the same time, regulatory segmentation prevents capital from flowing freely across platforms, allowing persistent price dispersion and limiting arbitrage. Together, these institutional differences suggest that price discovery concentrates in venues where liquidity is deeper and participation constraints are weaker, while segmentation and regulatory frictions shape both the speed and the location of information revelation.

A natural caveat of our analysis is its focus on a single event, the 2024 U.S. presidential election, which may raise questions about external validity. Nevertheless, this setting offers a uniquely informative laboratory for several reasons. First, the U.S. presidential election is among the most salient global events, with broad political and economic implications. Second, it generated unprecedented trading volume,[7] providing visibility into liquid modern prediction markets. Finally, our cross-platform comparison, spanning markets with different regulatory regimes and user bases, offers insights with potential policy relevance, especially amid ongoing debate over the role and regulation of prediction markets.

Our results contribute to several strands of literature. First, traditional prediction markets, often small and participation-restricted, have been shown to incorporate useful information about future events (see, for example, Wolfers and Zitzewitz, 2004; Snowberg, Wolfers, and Zitzewitz, 2007; Arrow et al., 2008; Cowgill and Zitzewitz, 2015; Atanasov et al., 2017). Modern blockchain-based platforms and regulated event-contract exchanges have expanded both trading volume and participation, moving prediction markets into a very different institutional and microstructure environment. Recent work has begun to use prices established in these modern prediction markets to capture macroeconomic expectations, voter preferences, or firm-level earnings expectations (e.g., Eichengreen et al., 2025; Taillard and Zeng, 2025; Swanson, Renxuan, and Wu, 2025; Chernov, Elenev, and Song, 2025; Gómez-Cram et al., 2025).

Whereas this recent literature primarily takes prediction-market prices as inputs for measuring expectations, this paper is the first to study the platforms themselves and the price formation process across venues. Our paper contributes to this emerging literature by shifting attention from the informational content of prices from these platforms to the institutional determinants of price formation across platforms. This focus matters

[7]The 2024 U.S. Presidential Election contract remains the highest-volume market in the history of these platforms (see e.g., Polymarket Analytics, 2025a).

because, as we show, not all modern prediction markets excel at price discovery, so the informativeness of prediction-market prices depends on the platform on which those prices are formed. More broadly, because informed trading, liquidity provision, and market design can differ materially in today's larger and more open markets than in earlier settings, understanding market structure is essential for interpreting the context under which modern prediction-market prices are formed and the information they convey.

We also contribute to the literature on the limits of arbitrage, which analyzes the various forces that permit persistent violations of the law of one price (see, e.g., De Long et al., 1990; Shleifer and Vishny, 1997; Gromb and Vayanos, 2010). While most of the evidence has been documented in equity or closed-end funds, recent studies indicate that deviations from the law of one price are widespread across other asset classes. For example, Makarov and Schoar (2020) find significant price dispersions across crypto exchanges, and Rhode and Strumpf (2004) examine historical U.S. presidential betting markets, showing how legal restrictions, information frictions, and market organization affected the efficiency and limited arbitrage opportunities, thereby influencing price convergence. Our setting is similar to that of Froot and Dabora (1999) and Makarov and Schoar (2020), as the same underlying asset is traded simultaneously across distinct markets. Our evidence indicates that trading costs account for only part of the observed price deviations, while other frictions, such as legal constraints, also likely play a significant role in limiting arbitrage.

Relatedly, we contribute to the literature on price discovery across multiple markets. Prior work on cross-listed securities shows that price discovery often occurs jointly across venues (e.g., Eun and Sabherwal, 2003), with the dominant market typically being the one with greater trading activity (Karolyi, 2006). Evidence for prediction markets, however, remains scarce. An earlier exception is Tetlock (2008), who examines financial and sport event outcomes and finds that liquidity does not always improve price discovery when naïve traders dominate. By examining contracts linked to the same outcome across several modern prediction platforms, we provide the first empirical evidence on cross-market information discovery in this setting. Unlike earlier work, we find a strong positive relationship between liquidity and price leadership: the platform with deeper liquidity and more active trading consistently leads others in incorporating information.

2 Data

This section provides an overview of the prediction market data collected from four platforms. We first provide an overview of the regulatory background that precedes the emergence of election prediction markets, as recent regulatory changes have contributed to the rise of modern election prediction markets. Then, we detail the specific data collection process and variable construction processes. Finally, we discuss the summary statistics.

2.1 Regulatory Background

Election betting was, for a long time, in a legal grey zone, with many federal and state-level laws banning or restricting such activities.[8] However, recent regulatory developments substantially relaxed constraints on U.S.-based platforms to offer political betting. This subsection provides an outline of the regulatory backgrounds for the two main prediction markets we examine.

Kalshi's ability to offer political contracts stems from its status as a Designated Contract Market (DCM), a designation it received from the Commodity Futures Trading Commission (CFTC) in 2020. As a DCM, Kalshi must comply with core principles designed to ensure market integrity, as outlined in the Commodity Exchange Act (CEA). However, the CFTC initially denied Kalshi's request to list election contracts, citing concerns that they constituted "gaming" and could undermine public trust in the electoral

[8] The legal status of election betting markets in the United States is complex and has evolved over time. Many state and federal laws explicitly ban such gambling, but the enforcement has been lax, allowing election betting markets to thrive from the 1880s to the 1940s. The Federal Wire Act of 1961 prohibits the use of wire communications for bets across state lines, which implicitly bans election markets operating on a national level. The Unlawful Internet Gambling Enforcement Act of 2006 bans online gambling on a broad level. Individual state laws also ban election gambling, with a recent notable example being the West Virginia Lottery, which authorized and then rescinded its participation after a short time in the 2020 elections, citing a 1868 state law that prohibited such wagers and stated they could unduly influence the elections (Sayre, 2020).

process (KalshiEx LLC v. Commodity Futures Trading Commission, 2024). This decision was overturned by a federal court ruling in September 2024, which vacated the CFTC's denial and established the legal basis for CFTC-regulated exchanges to list election-related event contracts. Kalshi subsequently listed and began trading these election contracts in early October 2024, which is the operational milestone reflected in Figure 1 and the start of our Kalshi election contract sample. We therefore refer to September 2024 when discussing the court decision and to October 2024 when describing the start of trading in these contracts.[9]

Contracts on political outcomes are a type of event-based contract. Historically, the CFTC has been cautious in approving such markets, weighing whether they serve a legitimate risk-management purpose or are closer to gambling. The 2024 court ruling in favor of Kalshi represented a landmark clarification for this market subclass.

Unlike Kalshi, which pursued and received CFTC approval as a designated contract market, Polymarket has never sought formal regulatory authorization to operate within the United States. In fact, following a 2022 settlement with the CFTC, Polymarket agreed to cease offering markets to U.S. users and implement geoblocking measures (Hochstein, 2024). As of June 2025, Polymarket remains unavailable in the U.S. due to its failure to seek legal status as a Designated Contract Market (DCM), though it continues to operate in many other countries, subject to local financial and gambling laws. Polymarket has since sought to return to the U.S. markets (Beyoud, Natarajan, and White, 2025).

Given the regulatory landscape, at the time of the study, Kalshi and PredictIt were available only to U.S. residents, whereas Polymarket catered to a global audience, excluding U.S. residents. PredictIt operated under CFTC staff no action relief granted in 2014, and the 2016 entry in our institutional timeline refers to PredictIt opening markets for the 2016 United States elections under that same framework, rather than to a separate no action letter. Robinhood is available only to users in the United States, the United Kingdom, the European Union, and Australia.

[9]The CFTC dropped its appeal of this decision in May 2025. ForecastEx LLC is also a CFTC-regulated Designated Contract Market. Its regulatory designation is separate from its later distribution arrangement with Robinhood. After the Kalshi litigation clarified the legal pathway, ForecastEx partnered with Robinhood to offer election contracts to Robinhood customers.

In Figure 1, we show a comprehensive timeline of the history of DCMs in our sample. Appendix A provides a more comprehensive overview of the historical developments in the prediction markets and regulations.

2.2 Data Collection and Variable Construction

2.2.1 Data collection

The details of our data collection are presented in Table A1, which also includes the sample period, data source, granularity, and the legal market participants.[10] We collected Kalshi trading data from its API, where we were able to retrieve all transactions relating to the election prediction market from its market start on October 4, 2024, to November 6, 2024. We also collected Polymarket data by setting up a virtual server that tracked Polymarket trades as they were conducted on its platform. We started data collection on October 23, 2024, and continued until 6th November 2024.

PredictIt does not provide any transaction data and only lists the prices of the opposing positions. Robinhood entered the market relatively late on October 28, 2024, and trading was limited to a graphical user interface (GUI) on its Android and iOS platforms. We stop our data sample at 20:00 GMT on November 6, 2024.[11] This is also the time when Polymarket has resolved its markets, declaring Donald Trump's victory. Additionally, transaction-level data, such as transaction size, is available via Kalshi and Polymarket APIs. Robinhood's transaction flow was manually reconstructed from screen recordings due to the lack of an accessible transaction API (see Appendix B.4 for details and screenshots).

Figure 2 shows the different active periods of the four platforms. While we were able

[10] Appendix B outlines the data collection procedure on all four platforms and provides a snapshot of the collected data for Kalshi and Polymarket. Daily Prices for Kalshi can be retrieved from https://kalshi.com/markets/pres/presidential-elections. Polymarket's can be retrieved from https://polymarket.com/event/presidential-election-winner-2024/will-donald-trump-win-the-2024-us-presidential-election. PredictIt's https://www.predictit.org/markets/detail/7456/Who-will-win-the-2024-US-presidential-election. Robinhood's interface has been completely removed, and it has since stopped offering betting on political outcomes after the 2024 Presidential Election.

[11] There were already no trades into Polymarket at this point, so even though we collected data till the Unix epoch time of 1730923200, the last recorded timestamp in our dataset is 1730922336 in Unix epoch time.

to obtain the complete transaction data for Kalshi from its market start on October 4th, in contrast, we were only able to collect two weeks of transaction data from Polymarket. Robinhood then entered the market only a week before the election results on November 6th. Thus, we can compare cross-market activity in the crucial 2-week overlapping period before November 4th, where we have full data on Kalshi and Polymarket, supplemented by Robinhood's. We mainly focus on the 15-day period leading up to the elections, during which we have transaction-level data for both Polymarket and Kalshi platforms.

2.2.2 Variable Construction

We construct price and return data at several time intervals. First, data is most widely available at the daily frequency. We obtain prices for each outcome from all markets at the daily market close, defined as 23:59 UTC. For Kalshi, we use the quoted price as of 23:59 UTC each day. For Polymarket, we calculate the midpoint of the bid-ask spread at 23:59 UTC. PredictIt similarly reports bid and ask prices, and we use the spread midpoint at 23:59 UTC. For Robinhood, which offers a binary pricing system via a graphical interface, we use the prevailing quoted price at 23:59 UTC.

To capture intraday price movements, we construct time series at 30-minute intervals. Returns are calculated using the logarithmic difference of prices between adjacent intervals:

$$R_t = \log(P_t) - \log(P_{t-1}),$$

where P_t denotes the price observed at the end of interval t.[12]

All timestamps are standardized to UTC to ensure cross-platform comparability. Table A1 summarizes the data sources, including platform-specific features, sample periods, and levels of data granularity.

[12]For the 30-minute return series, we evaluate prices on a fixed 30-minute grid using the most recent transaction price at or before each interval end. If no trade occurs within the previous 60 minutes on a platform, the price is treated as missing. Consequently, Polymarket API outage intervals (and returns spanning those outages) are excluded from the following regression analysis.

2.3 Summary Statistics

2.3.1 Daily Data

Panels A and B of Table 1 report daily prices from all four prediction markets. Panel A reports the prices for DJT-Win contracts, and Panel B reports prices for the KH-win contracts. The market with the longest sample is PredictIt, followed by Polymarket. We have a much shorter sample from Kalshi and Robinhood. For Kalshi and Robinhood, the prices of DJT-Win and KH-Win contracts add up to over 100 cents, and the portion exceeding 100 cents is attributable to spreads. For Polymarket and PredictIt, the prices of DJT-Win and KH-Win add up to less than 100 cents, as the data from these two markets covers a period before Joe Biden dropped out of the race and Kamala Harris became the candidate of the Democratic Party.

2.3.2 Intraday Data

Panels C and D of Table 1 present summary statistics for intraday volume and returns for the DJT-Win and KH-Win contracts, respectively. All statistics are based on 30-minute trading intervals. We note that there is more trading on Polymarket compared to Kalshi. For example, the average trading volume of DJT-Win on Polymarket is close to four times higher compared to Kalshi. Additionally, the DJT contracts tend to be more actively traded compared to KH contracts.

Figure 3 shows the daily trading volumes on the two markets. Consistent with the overall statistics, we find that Polymarket consistently shows higher trading volumes on each of the 15 trading days. Both markets become more active as the election day approaches, with volume peaking on the day of the election. On this day, more than 100 million contracts are traded on Kalshi and more than 300 million on Polymarket.

Panels C and D also show summary statistics for returns. The mean return for the DJT contract is slightly negative, while the mean return for the KH contract is slightly positive. This is because our analysis focuses on the sample ending at market close on November 4, 2024 (Kalshi's last full trading day), which is before the election-night resolution, to avoid capturing discontinuous price jumps as the final election results were announced. Therefore, the mean return reflects intraday drift over the pre-election window, which is downwardly biased in the DJT contract returns and upwardly biased in the KH returns.[13]The returns on Kalshi are more volatile compared to the same contract on Polymarket. For example, the DJT-Win contract return has a standard deviation of 120 basis points on Kalshi, and the same contract has a standard deviation of 81 basis points on Polymarket. This suggests there may be more active price discovery on Kalshi. Alternatively, this could also indicate that excessive uninformed trading leads to excessive volatility on Kalshi (e.g., Foucault, Sraer, and Thesmar, 2011).

3 Main Results

In our empirical analyses, we first assess the informativeness of prediction markets relative to polls by comparing their ability to predict the outcome of the 2024 U.S. presidential election. We then document daily price dispersion across four platforms. Next, we use intraday trading data from the two markets for which granular high-frequency data are available, Kalshi and Polymarket, to characterize the magnitude of arbitrage opportunities. Finally, we examine how liquidity and trading activity shape cross-platform price discovery.

3.1 Informativeness of Prediction Markets

Previous literature has established that prices on prediction markets tend to be highly informative (Wolfers and Zitzewitz, 2004; Arrow et al., 2008). Our analyses largely support this conclusion. For example, as shown in Figure 5, Polymarket seems to be willing to assign a higher likelihood than the average poll numbers to Donald Trump, the eventual winner in the election. We also conduct a VAR regression with poll numbers and DJT contract returns. As reported in Table 2, we find that prediction market returns significantly predict future poll numbers and changes in poll results. However, the poll numbers and changes do not significantly predict returns on Polymarket. Taken together,

[13]For the accuracy of the regression analysis, we will have to ensure the time series are stationary and free from the impact of several influential observations.

these results provide prima facie evidence that liquid prediction markets incorporated information about the 2024 presidential election ahead of polls.

3.2 Cross-Platform Price Discrepancies

Motivated by our hypothesis that liquidity and market design shape price discovery, we examine whether prices on each platform agree with one another. Since the contracts on all platforms track the same election outcomes, we expect that the prices across the platforms will follow the law of one price. However, substantial fragmentation exists due to regulations and legal statuses. For example, Polymarket operates only outside the United States, while Kalshi operates exclusively in the U.S. market. PredictIt operates internationally, but restricts the betting sizes, potentially preventing deep-pocketed institutions from operating on their platform. These obstacles could prevent prices from converging, similar to those on the cryptocurrency markets (Makarov and Schoar, 2020).

We start our exploration by plotting the daily prices of DJT and KH contracts in Figure 4. Panel (a) exhibits the prices of DJT on all four platforms, and Panel (b) exhibits the prices of KH on all platforms. In both plots, we observe substantial price differences across platforms. Importantly, these price dispersions seem one-sided. For example, we find that prices of KH on PredictIt are consistently above those on Polymarket in a large portion of our sample period. To quantify these observations, we report numerical measurements of price dispersion in Table 3.

Panel A exhibits the number of days when the contract price of one market is greater than that of the other market. In the lower triangle, we compare the contract prices of DJT, and the upper triangle exhibits the contract prices of KH. We report the number of days when the column market's price is higher than the row market's price. As a reference, we also report the number of overlapping days between the two markets in parentheses. Note that the prices of DJT and KH will not exactly mirror each other for two main reasons: First, for markets with longer sample periods, there are other dominant candidates in the race (i.e., President Joe Biden). Second, markets have different spreads, which can be substantial.

We find that Polymarket's DJT prices are consistently higher than those on Kalshi. Of the 33 overlapping days between these markets, Kalshi shows a higher price for DJT on only one day. Conversely, Polymarket's KH price is lower than Kalshi's in all 33 overlapping days. Like Kalshi, PredictIt also exhibits a bias toward the chance of Kamala Harris winning the election. The prices for KH on Polymarket are lower than those on PredictIt for 304 out of the 306 overlapping days. Robinhood's KH prices are higher than Kalshi's and Polymarket's on 9 out of 9 overlapping days, but are higher than PredictIt's on only 1 out of 9 overlapping days.

To gauge the economic magnitudes of these price spreads, we report the average price differences between market pairs in Panel B. Similar to Panel A, we report DJT prices in the lower triangle and KH in the upper triangle. We report the price differences in cents. These spreads can be substantial. Specifically, for DJT, the average absolute difference for Polymarket vs. Kalshi and Polymarket vs. PredictIt is 2.58 and 3.15 cents, respectively. On the KH contract sides, the largest average absolute difference involving is 6.08 cents between Polymarket and Robinhood, followed by 4.21 cents between Polymarket and PredictIt. The price spreads between Polymarket and Kalshi are also substantial, with Polymarket prices 3.26 cents lower than Kalshi for KH contracts. These price differences are also highly statistically significant.

We also report the absolute price difference across markets in Panel C. An arbitrager's profits come from price differences between two markets, regardless of the direction of the price differences. Thus, the numbers reported in Panel C are closer to an estimation of the arbitrage profits. Similar to Panel B, we find large and highly significant absolute price differences. In fact, the absolute price differences range from 1.24 to 9.66 cents.

3.2.1 Arbitrage Profits

Although we have already documented significant price dispersions across markets using daily data, these analyses have two limitations. First, daily prices are based on end-of-day prices and may be calculated at different times of the day. While Kalshi, Polymarket, and PredictIt report end-of-day prices based on 12 am UTC, the actual timing of the last trade is often unclear, particularly in thinner markets. This raises the risk that prices are stale and undermines our ability to assess whether arbitrage is feasible in real-time. Second, we do not know exactly how liquid those markets are, limiting our assessment of the potential dollar profits.

To mitigate these shortcomings, we collect high-frequency transaction-level data for Kalshi and Polymarket. Using the trade data, we thoroughly investigate the economic magnitudes of the trading profits. Our calculation follows Makarov and Schoar (2020) closely.[14] The platforms do not offer the ability to short contracts, essentially allowing arbitrageurs to buy four categories of contracts across the two trading platforms:

- Category 1: DJT-Win
- Category 2: KH-Lose
- Category 3: DJT-Lose
- Category 4: KH-Win

During this period of our analysis, there are only two candidates in consideration, as Joe Biden has dropped out at this point. We can thus collapse the four categories into two directions of trade. Categories 1 and 2 represent bullish on DJT (bearish on KH), while categories 3 and 4 are bearish on DJT (bullish on KH).

We consider an arbitrage opportunity to exist if the following condition has been met:

$$S(j,k) = P_j + P_k < 1 - \delta,$$

where $S(j,k)$ represents the sum of prices between two opposing categories j and k, and

[14]Appendix Table A2 reports the within-platform trading profits. We find little evidence that arbitrage opportunities exist within the same platform. More than 97% of the arbitrage also occurred in the last three hours of the sample. This is due to significant fluctuations in prices as election results are gradually released for each state.

δ represents the cost of arbitrage.[15] For example, if DJT-Win is 40 cents and KH-Win is 55 cents, a player can buy both positions and earn 5 cents when the market resolves. There are four pairs of these trades.[16] We consider values of δ at 0 and 0.02, representing the arbitrage opportunity without considering transaction costs and the scenario with a 2% cost of arbitrage. The 2% threshold follows the parameter set in Makarov and Schoar (2020).

When calculating arbitrage profits, we use prices and volumes of trades on both platforms. Thus, we assume that the maximum profits are restricted by the observed trading volume. This assumption may be too stringent, as it does not consider the possibility of trading on the potential of placing larger trades on the platforms without significantly impacting trading prices.

To quantify arbitrage profits, we implement an algorithm that matches opposing trades across platforms within a specified time interval (e.g., 1 second, 1 minute, 5 minutes, and 1 day). For each interval, we identify all pairs of trades where the sum of prices is less than $1 - \delta$. The profit for each matched pair is determined by the minimum volume of the two trades, ensuring that volume is not double-counted. This process is repeated iteratively until all profitable matches within the interval are exhausted.[17]

Table 4 shows the total possible cross-platform arbitrage profits. For each arbitrage frequency, we consider $\delta = 0$ and $\delta = 2\%$. We start by examining arbitrage profits with a second-by-second match. This allows the arbitrager to bear minimal inventory risks, as we require all trades to be matched within the next second. The first row displays the position when we take a long DJT position on Kalshi, matched by a short DJT position on Polymarket. This direction of arbitrage yields a significant total profit of

[15] In our context, the cost of arbitrage is defined broadly. First, it should capture the costs associated with trading and transferring money across platforms. Other costs of arbitrage may involve the arbitrager's risk-bearing. For example, Pontiff (2006) argues that volatility is the single most important cost of arbitrage. Makarov and Schoar (2020) also use the 2% wedge to prevent spurious results from volatility.

[16] Specifically, one can execute one of the following pairs of trades across two platforms (DJT-Win, DJT-Lose), (DJT-Win, KH-Win), (DJT-Lose, KH-Lose), (KH-Win, KH-Lose).

[17] To prevent overlap in calculations, we match DJT-Win on one platform trade with any DJT-Lose or KH-Win on trades on the other platform. After this, the remaining DJT-Lose and KH-Win trades that were in the unused trades are then arbitraged against the KH-Lose trades. We use this order because we previously determined that DJT-Win is the most liquid contract, and KH-Lose is the most illiquid trade consistently across the platforms.

$130,630. After imposing a 2% arbitrage cost, the total profit drops to $41,050. The second row illustrates the opposite direction of arbitrage, where we take a long position on DJT in Polymarket and a short position on DJT in Kalshi.[18] The total profit without arbitrage cost is $228.44, which is orders of magnitude lower than the other arbitrage arrangements. If we impose a 2% arbitrage cost, this profit drops to a mere $4. This is perhaps not surprising, as we have already shown in Table 3 that the price of the bullish DJT contracts is significantly higher on Polymarket than on Kalshi.

We next present the profits when we relax from second-by-second to one-minute and 5-minute intervals. As we increase the interval for arbitrage, naturally, the profits of arbitrage increase. However, arbitrageurs would also need to take on more risks to hold unhedged positions for longer periods of time. If we allow trades to be matched within 1-minute intervals, we find that the total arbitrage profits go up to $1.469 million by taking long DJT positions on Kalshi matched by long KH positions on Polymarket. The profits become $524,000 if we impose a 2% arbitrage cost. Lengthening the arbitrage interval also increases the arbitrage profits in the other direction to $1,300 (without arbitrage cost) and just $5 (with arbitrage cost). If we long DJT contracts on Kalshi and long KH position on Polymarket, the profits become $2.228 million (without arbitrage cost) and $853,000 (with arbitrage cost).[19]

We also visualize the average absolute price spread (blue bars) and the total daily dollar profit (red bars) in Figure 6. This plot suggests that the price spread between the two exchanges remains relatively stable over time and does not converge to zero, even up to the election day. However, the arbitrage profits are largely concentrated on the day before and the day of the election. This concentration is largely due to the higher volume on both exchanges.

Overall, while our estimates of arbitrage opportunities are highly conservative, we show that the violation of the law of one price is not negligible and can be exploited by sophisticated market participants.

[18]Short position means either buying DJT Lose or KH Win contracts.

[19]In Appendix Table A2, we repeat the same analysis within each platform. We find that the within-platform arbitrage profits are highly limited. Even at the 5-minute intervals, the profits of arbitrage are less than $30,000 after accounting for arbitrage costs.

3.3 Price Discovery and Lead-Lag Across Markets

Since prediction market prices do not fully align, the question arises as to where price discovery takes place. The two platforms we closely examine have distinct clientele and liquidity. On the one hand, Kalshi is based in the United States and attracts participants who may directly vote in the presidential election. This could potentially give a lead in capturing the real-time sentiment toward candidates. On the other hand, Polymarket has higher trading volume, indicating more active participation from investors. Chordia, Roll, and Subrahmanyam (2008) and Chordia et al. (2009) find that liquidity generally enhances price informativeness. Moreover, Karolyi (2006) observes that empirical evidence on cross-listed shares suggests that price discovery is related to where trading actually takes place, and Chordia and Swaminathan (2000) find that high volume portfolio returns tend to lead low volume portfolios. However, Tetlock (2008) finds opposite evidence based on a sports betting market. Given the past evidence, whether price discovery happens predominantly on more liquid markets remains an open question.

Since our analysis focuses on a single outcome, we cannot conduct cross-sectional tests on how liquidity affects price informativeness using this outcome. Instead, we will focus on the lead-lag patterns of returns in markets. This type of analysis has been widely used in the microstructure literature. The closest to our study is Eun and Sabherwal (2003), who study the price discovery of equities cross-listed in the US and Canadian markets.

Our empirical model follows a VAR(1) model in the following form:

$$\mathbf{R}_t = \mathbf{c} + \mathbf{\Phi}_1 \mathbf{R}_{t-1} + \mathbf{u}_t, \tag{1}$$

where $\mathbf{R}_t$ is the 2×1 vector of return variables at time t:

$$\mathbf{R}_t = \begin{pmatrix} R_t^K \\ R_t^P \end{pmatrix},$$

and R^K and R^P represent returns of Kalshi and Polymarket, respectively. We use returns at 30-minute intervals in calculating the returns. The choice of intraday interval enables us to detect potential relatively transient lead-lag patterns, and the one-lag model is also consistent with the suggestion from the Bayesian Information Criterion (BIC).

The results of this regression are reported in Table 5. We conduct separate regressions for DJT and KH contracts. The left panel presents the equations with R^K as the dependent variable. We find that both DJT and KH contracts exhibit similar patterns. First, we find that returns for both contracts on Kalshi exhibit strong and negative serial correlation. This pattern is consistent with a short-term reversal pattern that can be partially caused by a bid-ask bounce (e.g., Roll, 1984; Jegadeesh, 1990). Polymarket returns, in contrast, are strongly positively predictive of Kalshi returns. This lead-lag pattern potentially suggests that some information is first reflected on Polymarket before being impounded into Kalshi's prices.

The right panel presents the results with R^P as the dependent variable. Similar to the left panel, we find that lagged Polymarket prices are negatively related to R^P at time t. Nevertheless, this relationship is not statistically significant, potentially suggesting higher liquidity of Polymarket. We find that Kalshi returns do not significantly predict Polymarket returns for the DJT contract, and a smaller spillover for the KH contract compared to the spillover in the opposite direction. Overall, our results are consistent with Polymarket leading Kalshi in the price discovery process.[20]

To better understand the magnitude of the spillover, we plot the impulse response functions from our VAR systems in Figure 7. The upper panel exhibits the plot for the DJT contracts. The impulse response from Kalshi to Polymarket in blue and Polymarket to Kalshi in red. We find that price changes in Kalshi do not lead to significant changes in Polymarket returns in the following periods. A 1 percentage point return increase on Polymarket is followed by an approximately 0.26 percentage point increase in Kalshi's return on the same contract.

The bottom panel plots the KH contract. The overall patterns are similar. The shocks to Kalshi returns do not significantly transmit to Polymarket returns. In contrast, a 1 percentage point higher return in Polymarket is associated with a 0.25 percentage point

[20]In an unreported Granger causality test, we also find that Polymarket's return is the Granger cause of Kalshi's return, which is consistent with these arguments.

higher return in Kalshi.

3.3.1 Liquidity and Price Discovery

We next explore the role of market liquidity in the price discovery process. Drawing on insights from studies like Chordia, Roll, and Subrahmanyam (2008), which document how market liquidity improves efficiency by reducing return predictability associated with order flows, our key conjecture is that the lead-lag relationship across prediction markets should be largely concentrated in periods when Kalshi's liquidity is weak compared to Polymarket. To empirically investigate this, we conduct a VAR analysis in subsamples determined by the liquidity levels in both prediction markets. Given the absence of bid-ask spread data from either platform, we utilize the illiquidity measure by Amihud (2002), calculated over a 6-hour period. The relative liquidity ratio is then calculated as $\frac{1/I_P}{1/I_K}$, where I_P and I_K represent Amihud illiquidity measures of Polymarket and Kalshi, respectively. Our hypothesis is that when Kalshi is less liquid than Polymarket, trading frictions prevent the immediate absorption of new information on Kalshi, resulting in a stronger dependence of Kalshi returns on past Polymarket returns.

The results are reported in Table 6. In Panels A and B, we report results based on DJT and KH contracts, respectively. We find that during periods of high relative liquidity, Polymarket returns significantly predict Kalshi returns for both contracts. Meanwhile, during periods of low relative liquidity, this lead-lag relationship is statistically insignificant at 5% significance level. We also do not find significant evidence that Kalshi returns lead Polymarket in either high or low relative liquidity periods.

The impulse response function confirms these findings. As reported in Figure 8, we find that shocks to Kalshi returns do not lead to changes in returns on Polymarket. In contrast, Polymarket returns lead Kalshi returns by one period when Polymarket exhibits high relative liquidity. When the relative liquidity ratio is low, shocks to Polymarket returns do not predict changes in Kalshi returns in subsequent periods.

These results show the importance of market liquidity in facilitating price discovery. These results also differ significantly from Tetlock (2008), who find that prices are less efficient when liquidity is high. Our results suggest that there may be crucial differences between modern prediction markets with wide participation and those with more limited participation, as studied in prior literature.

3.3.2 Trading and Price Discovery

We previously hypothesized that one reason Polymarket could lead Kalshi in price discovery is that it appears to attract more active investor participation. Indeed, during our sample period, the mean daily volume on Polymarket is four times higher than that of Kalshi. To formally test the role of trading, we next split our sample based on the ratio of the six-hour cumulative trading volume on Polymarket to that on Kalshi, ahead of the return interval. We isolate high (low) relative volume ratios based on the median cutoff points and recompute the VAR(1) regression in (1) for the high and low relative volume ratio episodes, separately.

We report these results in Table 7. Panel A reports the results based on the DJT contracts, and Panel B reports the results of the KH contract. We find that during periods of high relative volume ratio, Polymarket returns significantly lead those of Kalshi for the DJT contract. However, such patterns are significantly weakened during periods of low relative volume ratio. For the KH contract, we find that Polymarket returns lead Kalshi returns in both high and low relative volume ratio periods, but the lead is both economically and statistically more significant in the high-volume ratio period.[21] In contrast, Kalshi returns do not lead Polymarket either in high or low relative volume ratio periods.

We also plot the corresponding impulse response functions for high and low relative volume ratio periods in Figure 9. These figures reveal a pattern similar to that in our VAR estimation. We find that Kalshi returns do not lead Polymarket in either period. Polymarket returns lead Kalshi only during periods of high relative volume. This lead, however, is not persistent and becomes statistically insignificant after the first 30-minute

[21] The contrast between high and low relative volume periods is less stark for KH contracts, possibly because they are traded much less heavily than DJT contracts, resulting in less variation in the volume ratio.

period. In terms of economic magnitudes, the high relative volume ratio period also exhibits a higher response compared to the overall sample period. For example, a 1 percentage point increase in the return of the DJT contract on Polymarket corresponds to a 0.37 percentage point increase in that on Kalshi in the following period.

Overall, these patterns support our conjecture that trading is crucial to price discovery in prediction markets.

3.4 Trading and Returns

As we documented in the previous section, trading plays an important role in facilitating price discovery in prediction markets. To delve into how trading affects price discovery, we further consider the direction of trades, measured by order imbalance on both markets. Motivated by Chordia, Roll, and Subrahmanyam (2002) and Chordia and Subrahmanyam (2004), who find that order imbalance impacts liquidity and returns, we examine how order imbalance relates to returns on both Kalshi and Polymarket.

Given that prediction markets are lightly regulated, sophisticated traders may be able to gain a significant information advantage over naive traders, leading to a stronger relationship between trading and returns. To investigate this relationship, we use regression specifications following Chordia and Subrahmanyam (2004):

$$R_t^p = \beta_0 + \beta_1 R_{t-1}^p + \beta_2 X_t^p + \beta_3 X_{t-1}^p + u_t, \quad p \in \{K, P\},$$

where the dependent variable is returns (in %) for a given contract on Kalshi (R_t^K) and Polymarket (R_t^P). X is the total order imbalance (OIB). The order imbalance is defined as the net buying volume normalized by the total volume.[22] In later analysis, we separate trades into whale and non-whale volume. A transaction is classified as a whale trade if its volume exceeds the 99th percentile of volume distribution on that trading day.

We report these results in Table 8. The top panel reports the results based on DJT contracts. As reported in column (1), we find that contemporaneous OIB is positively

[22] Please refer to Appendix B for details on how we identify the direction of trades for OIB calculation.

related to return on Kalshi. On Polymarket, the contemporaneous relationship is also positive, but statistically insignificant. Interestingly, the lagged OIB is significantly and positively related to future returns on both Kalshi and Polymarket. This result is a significant departure from Chordia and Subrahmanyam (2004), who find that lagged OIB is often negatively related to return. This difference may be due to the different horizon that we analyze, as Chordia and Subrahmanyam (2004) focus on daily OIB, rather than intraday patterns. Alternatively, it could also be a result of different participant clientele. In Chordia and Subrahmanyam (2004), the pattern is largely driven by the OIB of liquidity traders. On prediction markets, however, OIB may capture trades placed by informed traders.

To better understand this result, we decompose OIB into whale and non-whale ones, as previous literature documents the differential impact of large and small trades (e.g., Chan and Fong, 2000). Past research also uses trade sizes to differentiate institutional and retail order flows (Barber, Odean, and Zhu, 2008). Decomposing the OIB into whale and non-whale trades, we find that whale trades strongly predict returns both on Kalshi and Polymarket, while non-whale trades only predict returns on Kalshi.

In Panel B, we report the same analyses using the KH contract. For this contract, we find a weaker contemporaneous relationship between OIB and return. However, we continue to find a significant relationship between lagged OIB and returns on both Kalshi and Polymarket. Decomposing trades into whale and non-whale ones, we find that for the KH contract, whale order imbalance also exhibits stronger predictive power over the next period returns.

These findings highlight the importance of trading in shaping market returns. In particular, large trades on prediction markets tend to carry significant information, and such information does not fully impound into prices within a 30-minute window. Our results also indicate important differences between equity and prediction markets.

Motivated by these findings, we further investigate whether the lead-lag relationship between the two platforms is contingent on the relative activity of large trades. We hypothesize that price discovery is more pronounced on a platform experiencing greater directional large-trade order flow, which we refer to as the "conviction" of large trades. Intuitively, when large traders possess precise signals, we expect them to trade more aggressively, leading to a more significant order imbalance. We define relative whale-trade conviction as the ratio of the absolute whale OIB on Polymarket to that of the corresponding contract on Kalshi. We then segment the sample into "high" and "low" periods, corresponding to observations in which this ratio is above or below its sample-wide median, respectively. Table 9 reports the results of this subsample analysis.

The results provide strong evidence that the lead-lag relationship is related to the relative activity of whale trades on each platform. Specifically, during periods with a high relative whale trade conviction ratio, past returns on Polymarket are a highly significant positive predictor of returns on Kalshi, with significance at the 1% level for both contracts. Conversely, when this ratio is low, the leadership dynamic reverses: past Kalshi returns become a significant predictor of Polymarket returns. Collectively, these findings demonstrate that price discovery is concentrated in markets with strong one-sided whale trade activities, supporting the hypothesis that whale activity is a key driver of informational flow between the two markets.

3.5 Further Discussions

Our central empirical finding is that price discovery occurs at markedly different speeds across prediction markets, and that the more liquid venue systematically leads in incorporating new information. We discuss several institutional differences across platforms that may help explain these liquidity disparities and the differntial speeds of price discovery.

Polymarket operates as a blockchain-based platform settling on the Polygon network in USDC stablecoins. It is a globally accessible platform that, during our sample period, excluded U.S. residents but attracted a broad international user base trading in stablecoins. This architecture enables continuous 24/7 trading with near-instantaneous capital recycling, thereby reducing funding frictions and lowering the effective cost of market-making. Its transition from an automated market maker to a central limit order book with liquidity rewards (see Appendix B.1) further strengthened displayed depth. Crucially, Polymarket imposes no binding position limits, accommodating the large directional bets that, as we document, are particularly informative for price discovery. Taken together, rapid settlement, continuous trading, and unconstrained position sizes likely supported higher trading volume and greater participation by informed traders, resulting in faster price adjustment.

In contrast, Kalshi operates as a CFTC-regulated Designated Contract Market under the Commodity Exchange Act. While this regulatory status enhances legal clarity and investor protection, it also imposes stricter onboarding, KYC requirements, and compliance obligations that may slow capital flows and narrow participation. Kalshi settles through traditional financial infrastructure and operates a central limit order book. Relative to crypto-native settlement, conventional banking rails may limit the speed with which market makers recycle capital and respond to incoming order flow, particularly during periods of elevated trading activity. On Kalshi, although large trades are also informative, their impact on prices materializes with a delay when relative liquidity is lower.

The differences are starker still for PredictIt and Robinhood. PredictIt operated under a CFTC no-action letter with an \$850 position cap and charged a 10% fee on profits and a 5% withdrawal fee, mechanically capping trade size and raising the cost of active market-making. These constraints limited the presence of large, informed traders whose activity we show to be central to price discovery. Robinhood entered the election market only nine days before the vote and offered a simplified binary match engine rather than a limit order book, with no ability for users to post resting liquidity or execute paired cross-platform trades. By constraining depth and limiting active liquidity provision, these design features reduce the scope for informed trading to be rapidly impounded into prices, which helps explain why neither platform emerges as a leader in price discovery.

The persistent price dispersion across platforms highlights the role of regulatory and access frictions in limiting arbitrage. Because U.S. residents were barred from Polymarket (following its 2022 CFTC settlement) and non-U.S. users could not access Kalshi during our sample period, capital could not flow freely to eliminate cross-market mispricing. These institutional barriers resemble the location-of-trade effects studied by Froot and Dabora (1999) and the segmentation effects documented in cryptocurrency exchanges by Makarov and Schoar (2020). The one-sided nature of the price gaps, with DJT contracts persistently higher on Polymarket and KH contracts persistently higher on Kalshi and PredictIt, suggests systematic differences in beliefs across geographically segmented user bases. PredictIt's pricing reinforces this interpretation: despite its nominal international availability, its KH prices tracked above Polymarket's, likely because its position limits attract a participant base whose composition more closely resembles Kalshi's retail-oriented U.S. users.

In sum, the speed and location of price discovery across modern prediction markets reflect the interaction of platform design, regulatory status, and user composition. More permissive access and deeper liquidity accelerate information incorporation, whereas regulatory segmentation and participation constraints slow it. Importantly, the institutional features that facilitate liquidity and rapid price discovery, open access, large position limits, and pseudonymity, are often the same features that complicate oversight and raise concerns about market integrity and investor protection. Understanding this tension is essential for policymakers seeking to harness the informational benefits of prediction markets while maintaining adequate safeguards.

4 Conclusion

This paper studies price discovery across modern prediction markets using high-frequency, cross-platform data from the 2024 U.S. presidential election, the most actively traded outcome in prediction-market history.

We find that, while a liquid prediction market outperforms polls in predicting the election outcome, substantial and persistent price discrepancies remain across platforms. These discrepancies violate the law of one price and give rise to economically meaningful arbitrage opportunities. Using unique intraday, transaction-level data from Polymarket and Kalshi, we find that Polymarket dominates the price discovery process, particularly during periods when the platform is more liquid and accounts for a large share of trading activity. These patterns are consistent with microstructure theories in which liquidity facilitates price discovery and improves information efficiency.

We further examine the role of large trades in price discovery. We find that directional order flows of large trades, measured by the net order imbalances, strongly predict future returns, whereas order imbalance from smaller trades exhibits little predictive power. Crucially, the platform with greater large-trade order imbalance tends to lead in price discovery. To the extent that large-trade imbalance reflects episodes of more directional trading by large investors ("whales"), it provides a revealed measure of informed trading intensity. Accordingly, our evidence suggests that price discovery concentrates in venues where informed trading is more active.

Taken together, the evidence highlights that prediction markets are not frictionless aggregators of beliefs, but institutional markets in which design and regulation materially shape information revelation. Our findings provide new insights and raise open questions for policymakers, market designers, and regulators. For example, the dominance of Polymarket in price discovery underscores the critical role of liquidity in fostering efficient and informative markets. At the same time, although large, directional trades contribute to more rapid price discovery, the presence of such traders and the profitability of their trades also raise concerns about whether such traders enjoy an unfair advantage in markets that lack transparency and monitoring. As prediction markets continue to gain prominence, future research is needed to explore policy frameworks and platform design choices that harness the potential of these markets as valuable tools for aggregating public and private information, while ensuring fairness and transparency.

References

Amihud, Y., 2002. Illiquidity and stock returns: Cross-section and time-series effects. Journal of Financial Markets 5, 31–56.

Arrow, K. J., Forsythe, R., Gorham, M., Hahn, R., Hanson, R., Ledyard, J. O., Levmore, S., Litan, R., Milgrom, P., Nelson, F. D., et al., 2008. The promise of prediction markets. Science 320, 877–878.

Atanasov, P., Rescober, P., Stone, E., Swift, S. A., Servan-Schreiber, E., Tetlock, P., Ungar, L., Mellers, B., 2017. Distilling the wisdom of crowds: Prediction markets vs. prediction polls. Management Science 63, 691–706.

Barber, B. M., Odean, T., Zhu, N., 2008. Do retail trades move markets? Review of Financial Studies 22, 151–186.

Beyoud, L., Natarajan, S., White, N. M., 2025. Polymarket set for US return after deal to buy tiny exchange. The Wall Street Journal, July 21. https://www.bloomberg.com/news/articles/2025-07-21/crypto-betting-site-polymarket-set-for-us-return-after-deal-to-buy-qcx.

Boehmer, E., Kelley, E. K., 2009. Institutional investors and the informational efficiency of prices. Review of Financial Studies 22, 3563–3594.

Cabot, A., Cloward, G., 2021. Federal wire act should adjust to state-regulated sports wagering, not the other way around: A proposal for change. Gaming Law Review 25, 109–124.

Chan, K., Fong, W.-M., 2000. Trade size, order imbalance, and the volatility–volume relation. Journal of Financial Economics 57, 247–273.

Chernov, M., Elenev, V., Song, D., 2025. The comovement of voter preferences: Insights from US presidential election prediction markets beyond polls. NBER Working Paper.

Chordia, T., Goyal, A., Sadka, G., Sadka, R., Shivakumar, L., 2009. Liquidity and the post-earnings-announcement drift. Financial Analysts Journal 65, 18–32.

Chordia, T., Roll, R., Subrahmanyam, A., 2002. Order imbalance, liquidity, and market returns. Journal of Financial Economics 65, 111–130.

Chordia, T., Roll, R., Subrahmanyam, A., 2008. Liquidity and market efficiency. Journal of Financial Economics 87, 249–268.

Chordia, T., Subrahmanyam, A., 2004. Order imbalance and individual stock returns: Theory and evidence. Journal of Financial Economics 72, 485–518.

Chordia, T., Swaminathan, B., 2000. Trading volume and cross-autocorrelations in stock returns. Journal of Finance 55, 913–935.

Cowgill, B., Zitzewitz, E., 2015. Corporate prediction markets: Evidence from Google, Ford, and firm X. Review of Economic Studies 82, 1309–1341.

De Long, J. B., Shleifer, A., Summers, L. H., Waldmann, R. J., 1990. Noise trader risk in financial markets. Journal of Political Economy 98, 703–738.

Eichengreen, B., Viswanath-Natraj, G., Wang, J., Wang, Z., 2025. Under pressure? Central bank independence meets blockchain prediction markets. Working Paper.

Eun, C. S., Sabherwal, S., 2003. Cross-border listings and price discovery: Evidence from US-listed Canadian stocks. Journal of Finance 58, 549–575.

Fernandez, R., 2025. Prediction-market bettors on Polymarket, Kalshi back Mamdani in mayoral election. The Wall Street Journal, November 4. https://www.wsj.com/livecoverage/election-day-2025/card/prediction-market-bettors-back-mamdani-in-mayoral-election-0VWhCVW6ZTRqrocn72wP.

Foucault, T., Kadan, O., Kandel, E., 2005. Limit order book as a market for liquidity. Review of Financial Studies 18, 1171–1217.

Foucault, T., Sraer, D., Thesmar, D. J., 2011. Individual investors and volatility. Journal of Finance 66, 1369–1406.

Froot, K. A., Dabora, E. M., 1999. How are stock prices affected by the location of trade? Journal of Financial Economics 53, 189–216.

Glosten, L. R., Milgrom, P. R., 1985. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. Journal of Financial Economics 14, 71–100.

Gómez-Cram, R., Guo, Y., Jensen, T. I., Kung, H., 2025. Financial prediction markets: A new measure of earnings expectations. Working Paper.

Gromb, D., Vayanos, D., 2010. Limits of arbitrage. Annual Review of Financial Economics 2, 251–275.

Hasbrouck, J., 1995. One security, many markets: Determining the contributions to price discovery. Journal of Finance 50, 1175–1199.

Hochstein, M., 2024. Polymarket's probe highlights challenges of blocking U.S. users (and their VPNs). CoinDesk, November 14. https://www.coindesk.com/policy/2024/11/14/polymarkets-probe-highlights-challenges-of-blocking-us-users-and-their-vpns.

Jegadeesh, N., 1990. Evidence of predictable behavior of security returns. Journal of Finance 45, 881–898.

Kalshi, 2025. Kalshi hits $5 billion valuation amid international expansion. Kalshi News, October 10. https://news.kalshi.com/p/kalshi-hits-5-billion-valuation-amid-international-expansion.

KalshiEx LLC v. Commodity Futures Trading Commission, 2024. No. 24-5205. https://media.cadc.uscourts.gov/opinions/docs/2024/10/24-5205-2077790.pdf [Accessed: 2026-01-09].

Karolyi, G. A., 2006. The world of cross-listings and cross-listings of the world: Challenging conventional wisdom. Review of Finance 10, 99–152.

Kyle, A. S., 1985. Continuous auctions and insider trading. Econometrica 53, 1315–1335.

Lee, C., Radhakrishna, B., 2000. Inferring investor behavior: Evidence from torq data. Journal of Financial Markets 3, 83–111.

Makarov, I., Schoar, A., 2020. Trading and arbitrage in cryptocurrency markets. Journal of Financial Economics 135, 293–319.

O'Hara, M., 1995. Market Microstructure Theory. Cambridge, MA: Blackwell Publishers.

Osipovich, A., 2024. The mystery $30 million crypto bets on a Trump win. The Wall Street Journal, October 18. https://www.wsj.com/finance/betting-election-pro-trump-ad74aa71.

Osipovich, A., 2025. A mystery trader scored with prescient bets on the Nobel Peace Prize. The Wall Street Journal, October 10. https://www.wsj.com/finance/stocks/polymarket-nobel-peace-prize-bets-c34ee0c8.

Polymarket, 2025. Liquidity rewards. https://docs.polymarket.com/polymarket-learn/trading/liquidity-rewards [Accessed: 2025-06-28].

Polymarket Analytics, 2025a. Highest volume closed markets on Polymarket. https://polymarketanalytics.com/markets?sources=Polymarket&marketStatus=Closed [Accessed: 2025-11-17].

Polymarket Analytics, 2025b. Presidential election winner 2024. https://polymarketanalytics.com/markets/903193 [Accessed: 2025-06-30].

Pontiff, J., 2006. Costly arbitrage and the myth of idiosyncratic risk. Journal of Accounting and Economics 42, 35–52.

Pourpouneh, M., Nielsen, K., Ross, O., 2020. Automated market makers. IFRO Working Paper No. 2020/08.

Rhode, P. W., Strumpf, K. S., 2004. Historical presidential betting markets. Journal of Economic Perspectives 18, 127–141.

Roll, R., 1984. A simple implicit measure of the effective bid-ask spread in an efficient market. Journal of Finance 39, 1127–1139.

Sayre, K., 2020. West Virginia approves, then disapproves, betting on elections. The Wall Street Journal, April 8. https://www.wsj.com/articles/west-virginia-approves-then-disapproves-betting-on-elections-11586384497.

Shleifer, A., Vishny, R. W., 1997. The limits of arbitrage. Journal of Finance 52, 35–55.

Snowberg, E., Wolfers, J., Zitzewitz, E., 2007. Partisan impacts on the economy: Evidence from prediction markets and close elections. Quarterly Journal of Economics 122, 807–829.

Stewart, E., 2026. Prediction markets' insider advantage. Business Insider, January 8. https://www.businessinsider.com/prediction-markets-polymarket-kalshi-venezuela-raid-bet-insider-trading-2026-1.

Swanson, E. T., Renxuan, W., Wu, Y., 2025. The effects of monetary policy on macroeconomic expectations: High-frequency evidence from traded event contracts. Working Paper.

Taillard, J., Zeng, L., 2025. The market value of Fed independence: High-frequency evidence from a natural experiment. Working Paper.

Tetlock, P. C., 2008. Liquidity and prediction market efficiency. Working Paper.

White, H., 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica 48, 817–838.

Wolfers, J., Zitzewitz, E., 2004. Prediction markets. Journal of Economic Perspectives 18, 107–126.

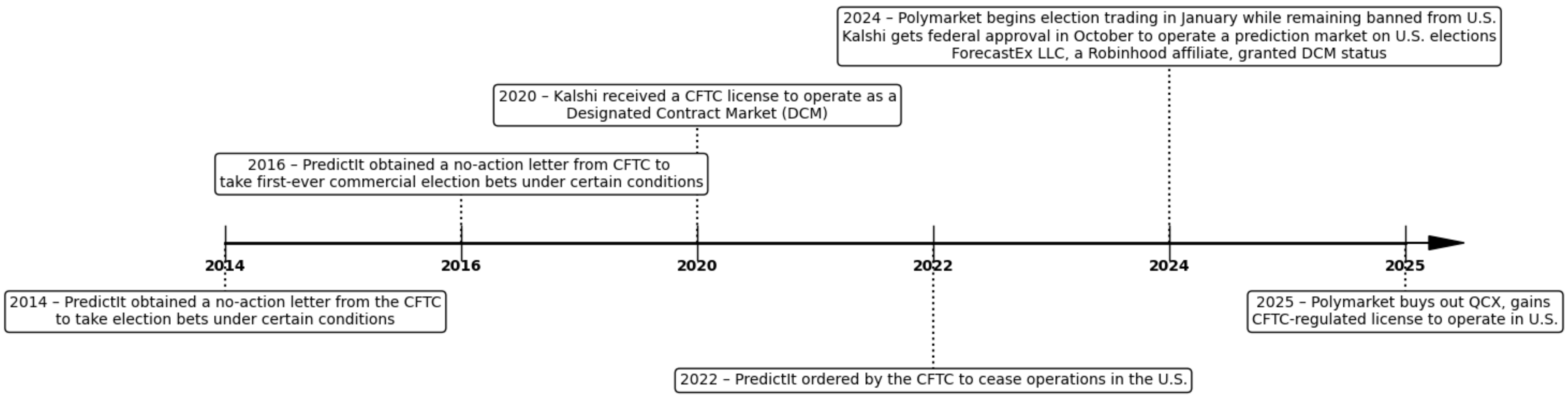


Figure 1. Election Prediction Markets Establishment and Regulatory Changes
This figure exhibits pivotal shifts in U.S. election prediction markets leading up to the 2024 presidential election, as well as developments after the election. This figure showcases key regulatory changes. A more complete timeline of significant events of prediction markets is given in Appendix A.

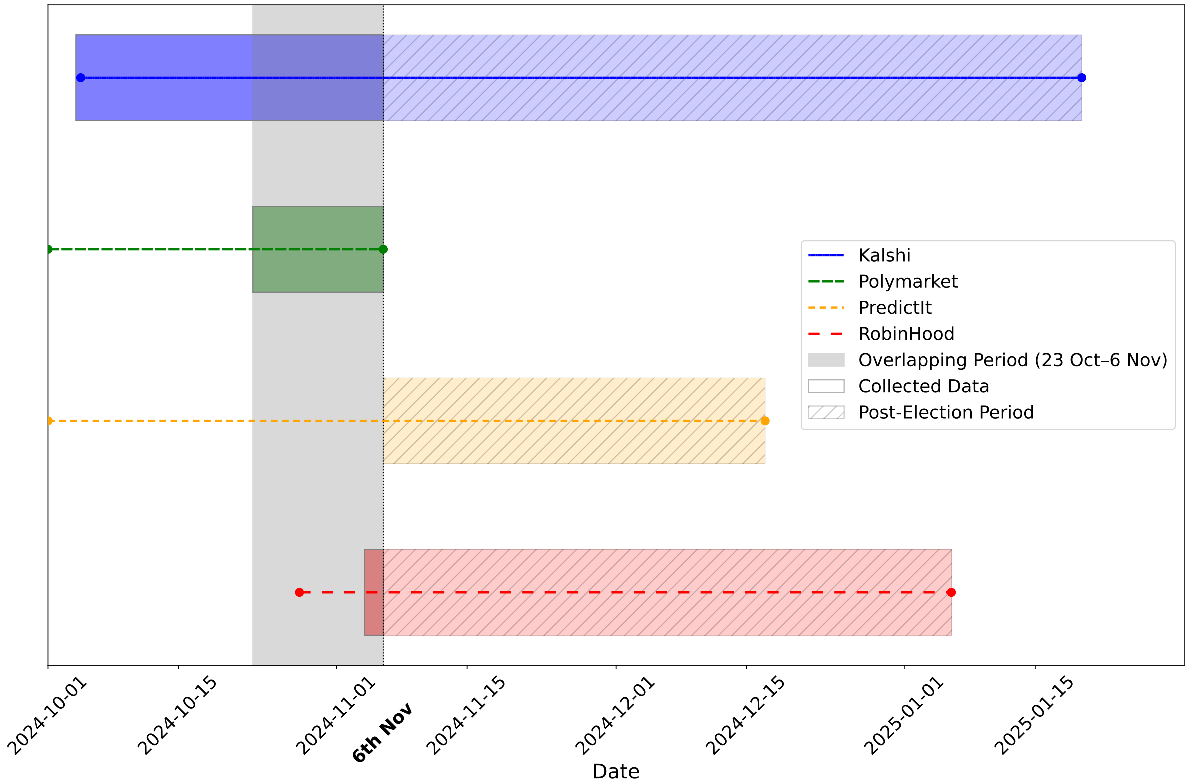


Figure 2. Active Periods of the Four Prediction Market Platforms
This figure maps the trading activity and our data availability for four election prediction markets. For each platform, the full horizontal line indicates the total trading window. The solid-colored bars show the specific periods for which we collected transaction-level data, while the hatched bars denote trading that occurred after the November 6th election. The vertical grey region highlights the key overlapping analysis period from October 23 to November 6. Transaction data for PredictIt was unavailable as it is not made public.

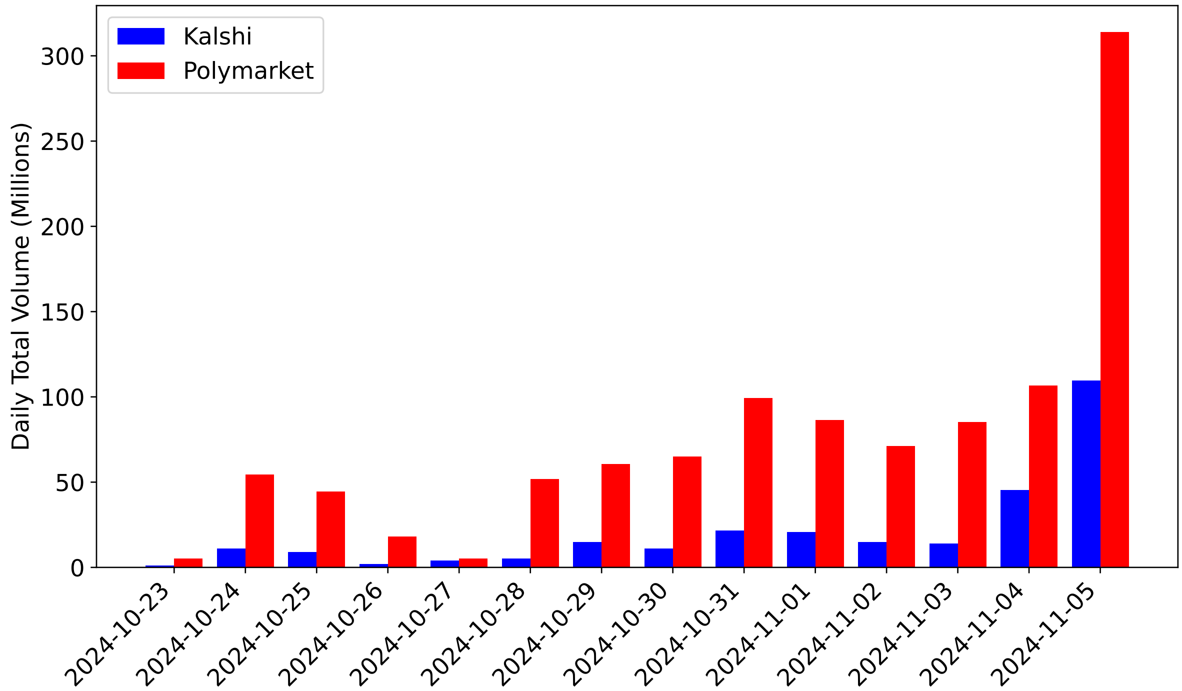
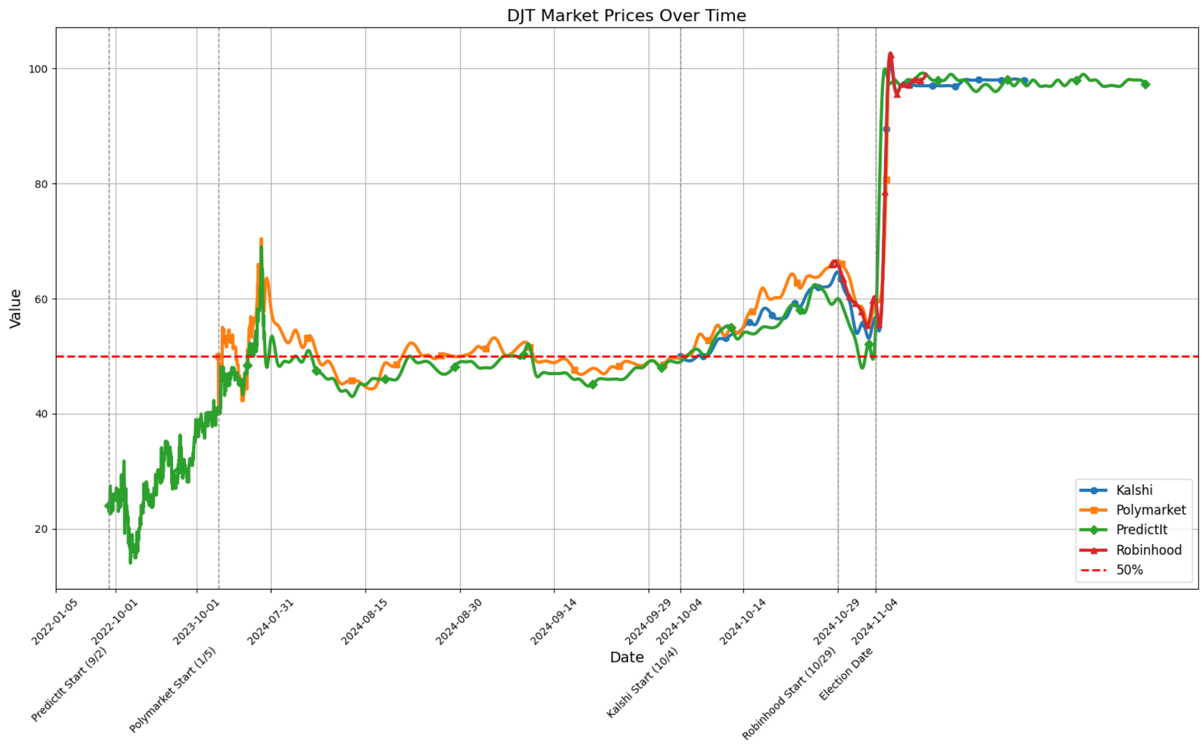
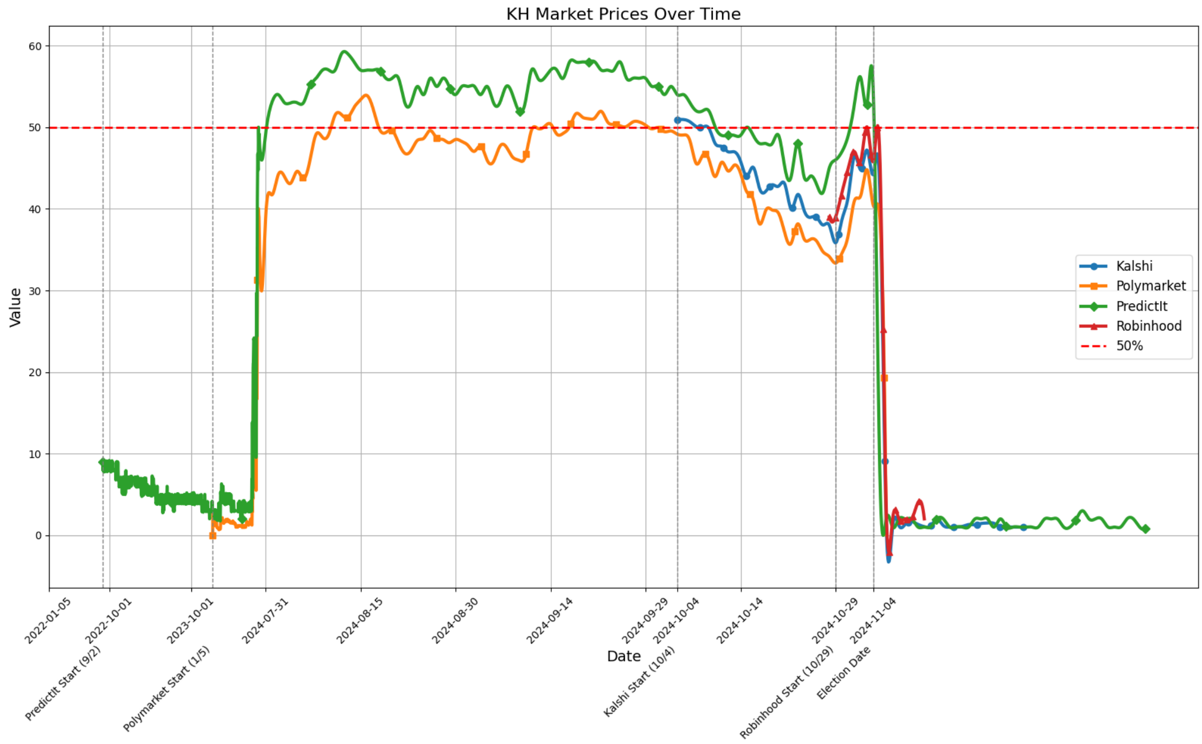


Figure 3. Daily Total Trading Volume on Kalshi and Polymarket.
This figure shows the daily total trading volume on Kalshi and Polymarket from October 23, 2024, to November 5, 2024. We calculate the total trading volume by summing up all the trading volumes of both DJT and KH contracts. The blue bar represents the volume on Kalshi, while the red bar represents the volume on Polymarket. The unit on the vertical axis is in millions of contracts.



(a) DJT-win Prices Across Markets



(b) KH-Win Prices Across Markets

Figure 4. Election Contract Prices Over Time
This figure shows the daily prices for contracts on the 2024 U.S. Presidential Election winner across four markets in cents. Panel (a) tracks the "DJT-win" contract, while Panel (b) tracks the "KH-win" contract. The time axis is compressed before mid-2023 to include early data from PredictIt. The dashed horizontal line in each panel represents 50 cents.

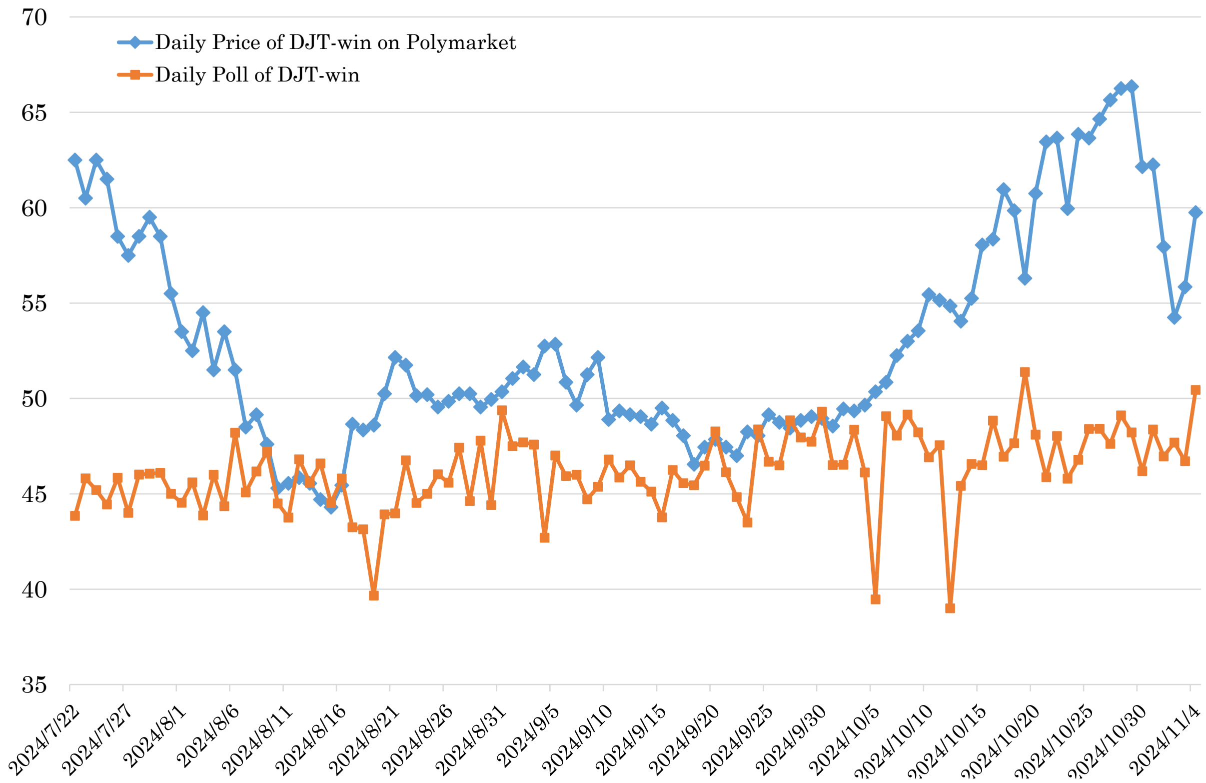


Figure 5. Daily Polls and Polymarket Returns on DJT-Win
This figure plots the average poll numbers for Donald Trump and the price of the DJT-Win contract on Polymarket.

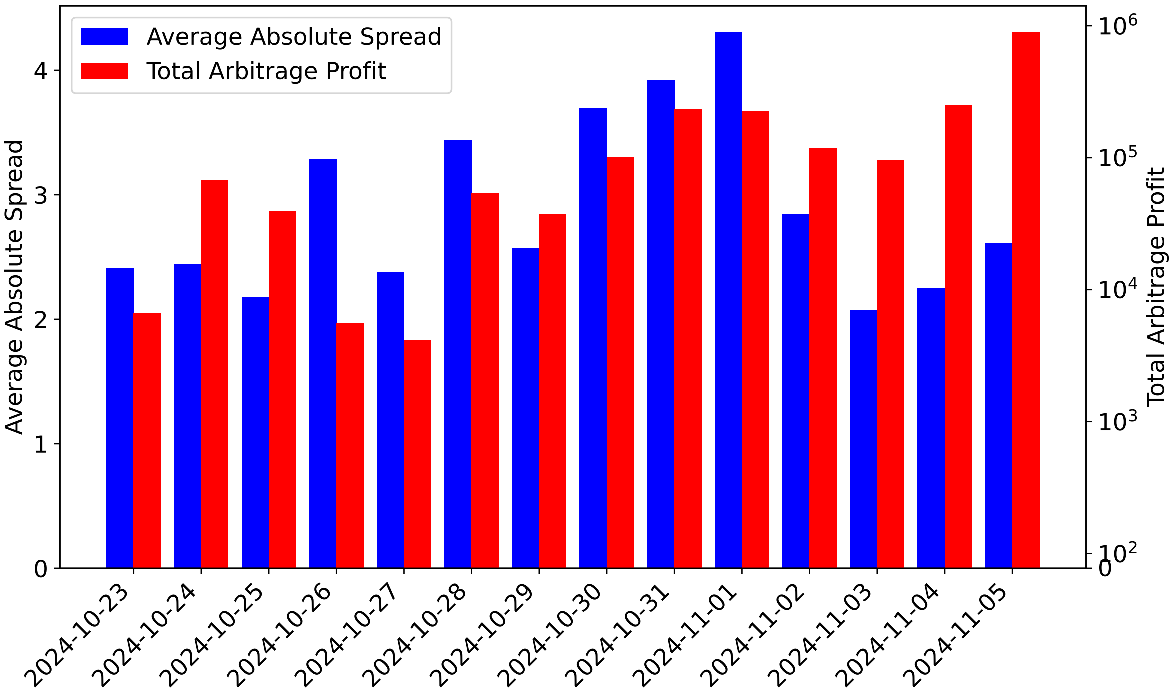
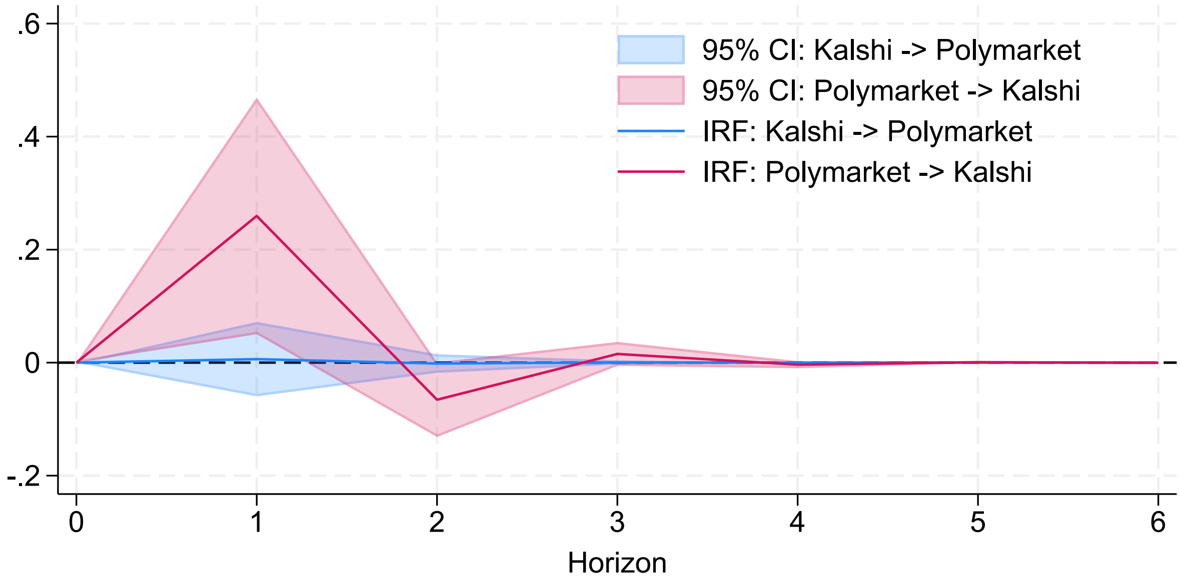
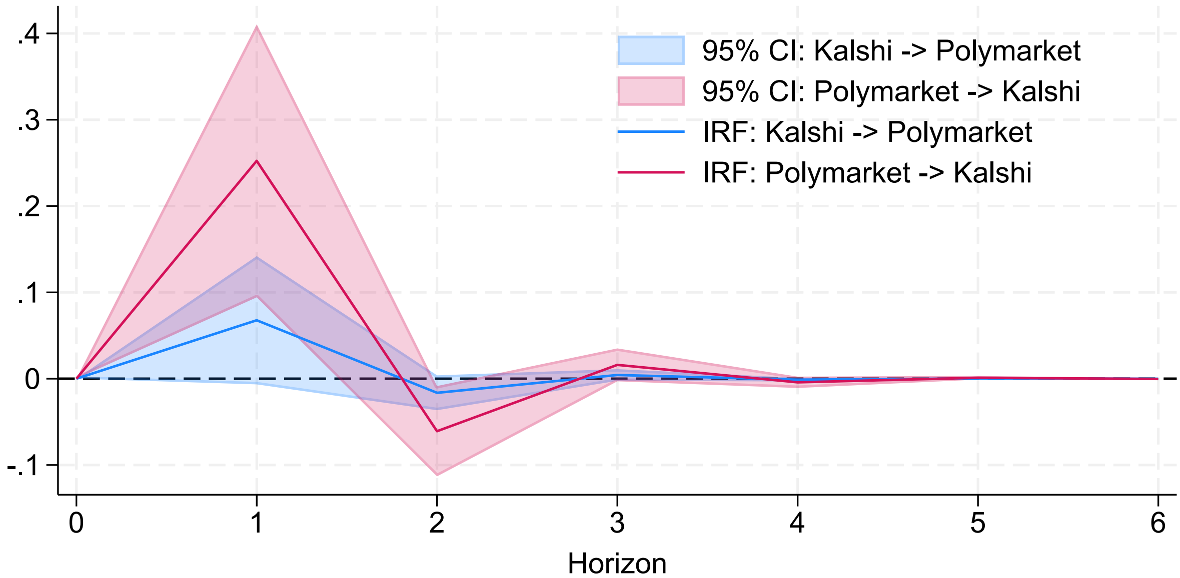


Figure 6. Daily Average Absolute Spread and Total Arbitrage Profit
This figure shows the average absolute spread between Polymarket and Kalshi and the daily total profit from October 23, 2024 to November 5, 2024. The blue bar is the average absolute spread in cents. We calculate this spread as an average of the DJT and KH contracts. The red bar represents the total daily dollar profit from trading both DJT and KH contracts, calculated using a 5-minute interval and assuming no arbitrage costs. The left vertical axis is in a linear scale, and the right vertical axis is in a symlog scale.

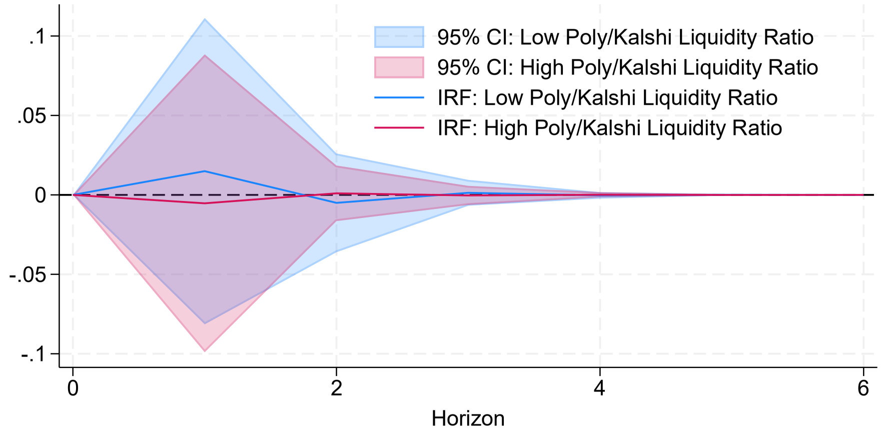


(a) DJT Contract

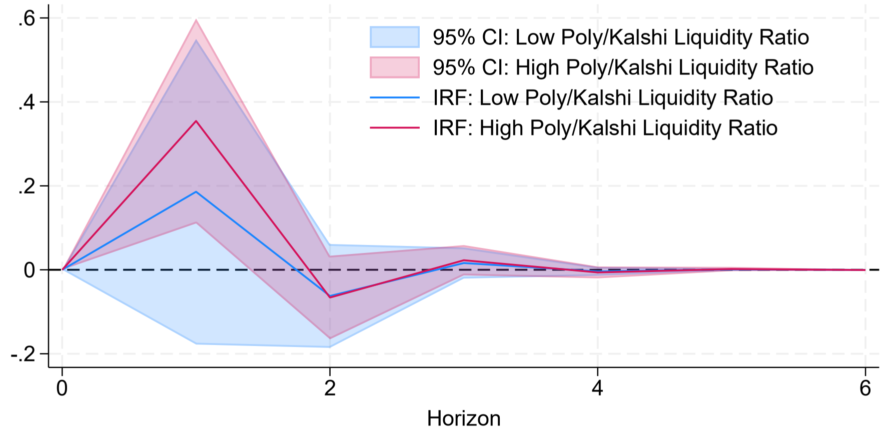


(b) KH Contract

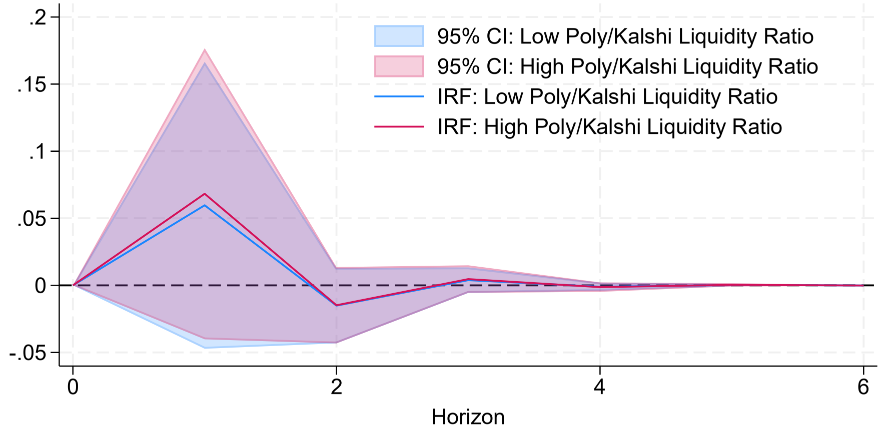
Figure 7. Cross-Market Impulse Response Functions
This figure presents impulse response functions (IRFs) from a VAR model of 30-minute returns for the DJT (Panel A) and KH (Panel B) contracts. The plots show how a return shock on one platform affects the other. The red line indicates Kalshi's response to a Polymarket shock, while the blue line indicates Polymarket's response to a Kalshi shock. Shaded bands represent 95% confidence intervals.



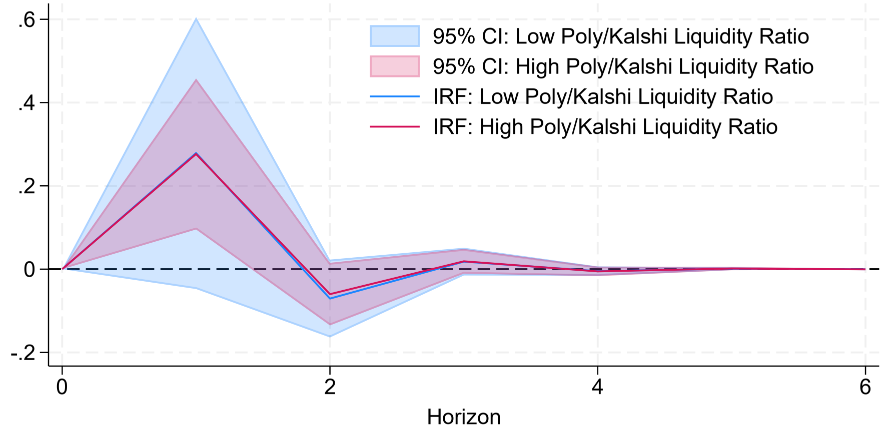
(a) DJT Contract: Kalshi → Polymarket



(b) DJT Contract: Polymarket → Kalshi

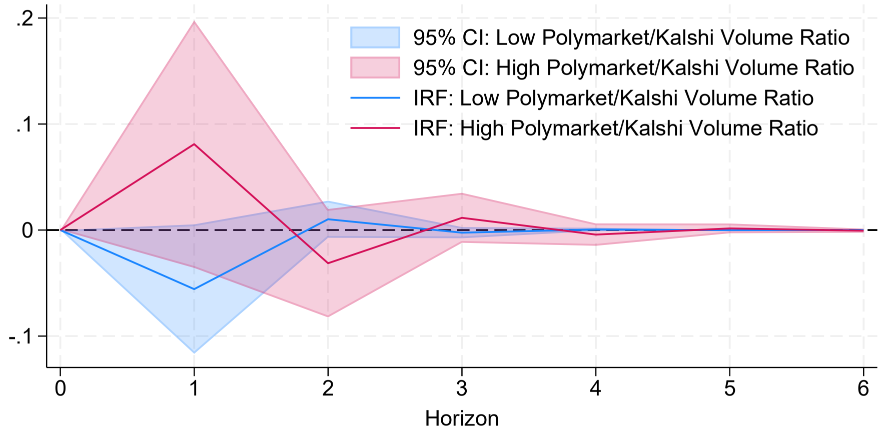


(c) KH Contract: Kalshi → Polymarket

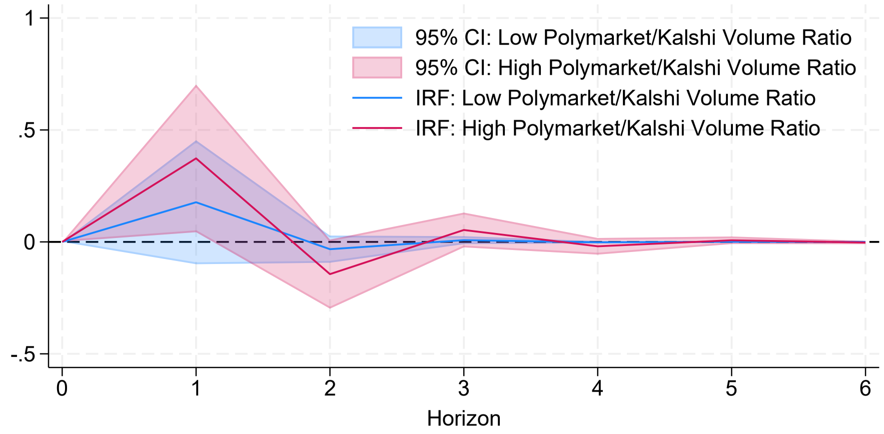


(d) KH Contract: Polymarket → Kalshi

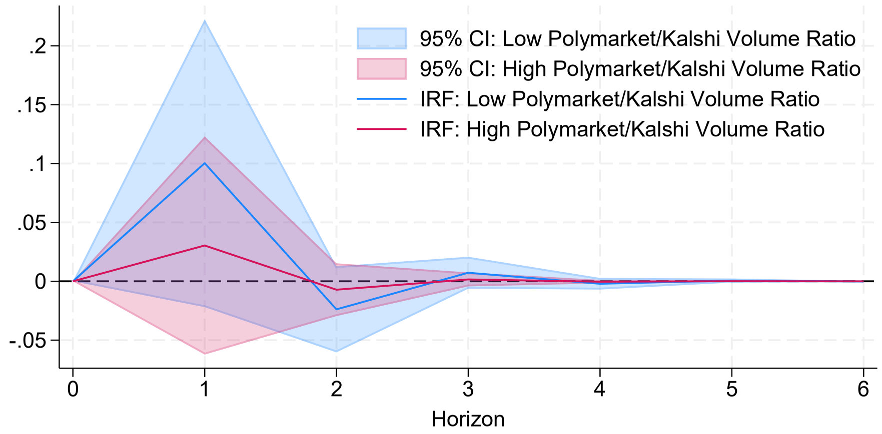
Figure 8. Impact of Illiquidity on Cross-Market Return Lead-Lag
This figure illustrates how relative liquidity conditions the flow of information between Polymarket and Kalshi. It plots impulse response functions (IRFs) from 30-minute returns for the DJT contract (top row) and the KH contract (bottom row). The left column (A, C) displays the impact of a Polymarket shock on Kalshi, while the right column (B, D) displays the impact of a Kalshi shock on Polymarket. Within each panel, we distinguish between periods when Polymarket's relative liquidity was high (red line) and low (blue line), based on a median split of the Polymarket-to-Kalshi liquidity ratio. Shaded areas are 95% confidence intervals.



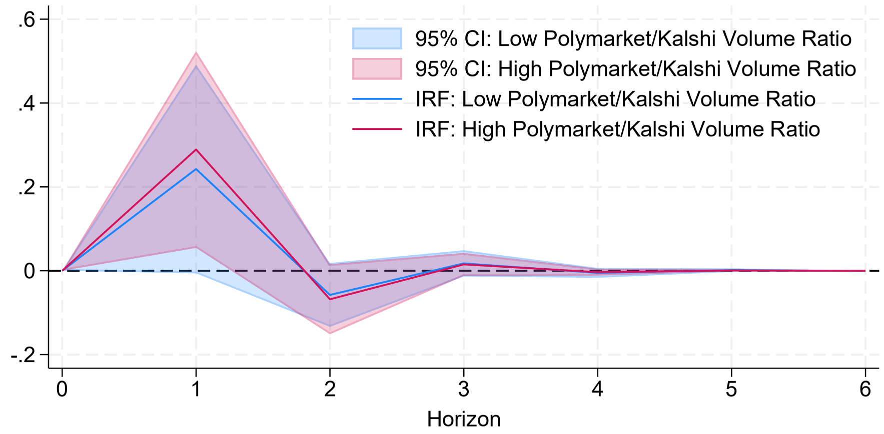
(a) DJT Contract: Kalshi → Polymarket



(b) DJT Contract: Polymarket → Kalshi



(c) KH Contract: Kalshi → Polymarket



(d) KH Contract: Polymarket → Kalshi

Figure 9. Impact of Trading Volume on Cross-Market Return Lead-Lag
This figure shows how the lead-lag relationship between Polymarket and Kalshi depends on trading activity. It plots impulse response functions (IRFs) from 30-minute returns for the DJT (top row) and KH (bottom row) contracts. The left column (A, C) displays the impact of a Polymarket shock on Kalshi's price, while the right column (B, D) displays the impact of a Kalshi shock on Polymarket's price. Each plot contrasts the response during periods of high relative Polymarket volume (blue line) with periods of low relative volume (red line). Shaded areas are 95% confidence intervals.

Table 1 Summary Statistics

This table reports descriptive statistics for prediction market contracts on the 2024 U.S. Presidential election for Donald J. Trump (DJT) and Kamala Harris (KH). Panels A and B present summary statistics for contract daily prices in cents, while Panels C and D detail dollar trading volume and log returns at 30-minute frequency. For consistency, daily pricing data are based on each day's final trade. The sample period for prices spans from each market's start date to November 5, 2024, while the period for dollar trading volume (in millions) and log returns (in %) spans from October 23, 2024, 18:00:00 EST to November 5, 2024, 03:00:00 EST.

Panel A: DJT-Win Daily Prices

Platform	Count	Mean	Std	25%	50%	75%
Kalshi	33	56.27	4.27	53.14	56.54	59.05
Polymarket	306	52.68	5.89	48.50	52.50	55.50
PredictIt	796	37.11	11.73	27.75	37.00	47.00
Robinhood	9	60.67	3.61	59.00	60.00	63.00

Panel B: KH-Win Daily Prices

Platform	Count	Mean	Std	25%	50%	75%
Kalshi	33	44.14	4.31	41.55	44.02	47.17
Polymarket	306	17.48	20.86	1.36	1.78	41.70
PredictIt	796	11.62	16.48	4.00	5.00	7.00
Robinhood	9	44.44	3.71	42.00	46.00	46.00

Panel C: DJT-Win Trading Volume and Log Returns

Variable	Platform	Count	Mean	Std	25%	50%	75%
Volume	Kalshi	475	0.110	0.165	0.015	0.042	0.131
Volume	Polymarket	412	0.430	0.346	0.214	0.346	0.526
Return	Kalshi	474	-0.024	1.234	0.000	0.000	0.000
Return	Polymarket	411	-0.023	0.809	-0.308	0.000	0.173

Panel D: KH-Win Trading Volume and Log Returns

Variable	Platform	Count	Mean	Std	25%	50%	75%
Volume	Kalshi	475	0.077	0.140	0.009	0.026	0.084
Volume	Polymarket	412	0.302	0.287	0.129	0.241	0.387
Return	Kalshi	474	0.066	1.603	0.000	0.000	0.000
Return	Polymarket	411	0.044	1.214	-0.440	0.000	0.585

Table 2 Lead-Lag Relationship between Polls and Polymarket Returns

This table reports the results from estimating the VAR(1) regressions for the vector of daily polls (or daily changes in poll) for DJT-win and the returns of the same contract on Polymarket over the sample period from July 22 to November 4, 2024. Reported are the regression coefficient estimates and their associated t-statistics, calculated using the White (1980) heteroskedasticity-robust standard errors. *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively.

	Dependent Variables			
	Poll_t	R_t^P	ΔPoll_t	R_t^P
	(1)	(2)	(3)	(4)
Poll_{t-1}	0.254**	0.001		
	[2.42]	[0.51]		
ΔPoll_{t-1}			-0.495***	0.001
			[-4.41]	[0.56]
R_{t-1}^P	13.987**	0.028	13.405**	0.03
	[2.45]	[0.24]	[2.24]	[0.25]
Intercept	34.533***	-0.036	0.073	0.000
	[7.08]	[-0.52]	[0.35]	[-0.04]
Observations	104	104	104	104
R^2	9.5%	0.3%	28.3%	0.3%
Granger Causality	6.02**	0.26	5.02**	0.31

Table 3 Cross-Platform Price Differences

This table presents pairwise price dispersion metrics for Donald J. Trump (DJT) and Kamala Harris (KH) contracts across four platforms: Kalshi, Polymarket, PredictIt, and Robinhood. In all panels, the lower triangle refers to the DJT contract and the upper triangle to the KH contract. Prices are quoted in cents. Panel A reports the number of days the column platform's price exceeded the row's, with the total number of overlapping days in parentheses. Panel B displays the signed average daily price difference (column price - row price), while Panel C shows the average absolute difference, which measures the magnitude of cross-platform disagreement. Brackets contain t-statistics calculated using White (1980) heteroskedasticity-robust standard errors.

Panel A: Number of Days (Price of Column > Price of Row)

	Kalshi	Polymarket	PredictIt	Robinhood
Kalshi	—	0 (33)	32 (33)	9 (9)
Polymarket	1 (33)	—	304 (306)	9 (9)
PredictIt	21 (33)	263 (306)	—	1 (9)
Robinhood	1 (9)	5 (9)	1 (9)	—

Panel B: Signed Price Dispersion (Column Price - Row Price)

	Kalshi	Polymarket	PredictIt	Robinhood
Kalshi	—	-3.2676***	3.5597***	2.3878***
	—	[-14.73]	[2.86]	[6.01]
Polymarket	-2.5885***	—	4.2137***	6.0889***
	[-8.29]	—	[16.02]	[12.90]
PredictIt	0.1161	3.1554***	—	-1.0000
	[0.10]	[12.43]	—	[-0.22]
Robinhood	-2.1644***	0.9889**	-1.5556	—
	[-5.67]	[2.18]	[-0.36]	—

Panel C: Absolute Price Difference

	Kalshi	Polymarket	PredictIt	Robinhood
Kalshi	—	3.2676***	5.6585***	2.3878***
	—	[14.73]	[5.16]	[6.01]
Polymarket	2.6030***	—	4.4209***	6.0889***
	[8.47]	—	[16.83]	[12.90]
PredictIt	2.8252***	3.7074***	—	9.6667***
	[2.70]	[17.18]	—	[2.74]
Robinhood	2.1756***	1.2444***	9.3333***	—
	[5.77]	[3.51]	[3.08]	—

Table 4 Cross-Platform Arbitrage on Kalshi and Polymarket

This table reports on cross-platform arbitrage opportunities between Kalshi and Polymarket from October 23, 2024, 18:00:00 EST to November 5, 2024, 23:59:59 EST. We list the arbitrage profits for different arbitrage intervals and by the platform where we take a bullish position on DJT. We consider two scenarios: one without a cost of arbitrage (i.e., $\delta = 0$) and another with a 2% arbitrage cost (i.e., $\delta = 2\%$). For each direction—Kalshi to Polymarket and Polymarket to Kalshi, we compute arbitrage profits by aggregating the difference between $1(or $0.98 when incorporating a 2% transaction cost) and the matched sum of the lowest available prices from both platforms within each interval, multiplied by the corresponding trade size. This produces realistic, volume-weighted arbitrage opportunities rather than theoretical bounds. To avoid double-counting within a given direction, we treat DJT-Win and KH-Lose as two payoff-equivalent ways of implementing the same "bullish DJT" leg. We first compute violations and associated profits using the DJT-Win implementation. We then exclude all intervals in which a DJT-Win-based violation is detected and compute any additional violations using the KH-Lose implementation only over the remaining intervals. This ensures that each interval contributes at most once to the total for that direction. The two directions (bullish DJT executed on Kalshi versus on Polymarket) are computed separately, and their totals can be summed to reflect the overall arbitrage potential available to a perfectly informed and unconstrained trader. The arbitrage profits are in dollars.

Interval	Platform with Long DJT Position	Total Violation Value (Without Fee)	Total Violation Value (2% Fee)
1-Second	Kalshi	130,629.55	41,053.78
1-Second	Polymarket	228.44	4.00
1-Minute	Kalshi	1,468,653.13	524,428.78
1-Minute	Polymarket	1,309.36	4.62
5-Minute	Kalshi	2,228,165.32	852,913.95
5-Minute	Polymarket	9,674.39	97.13

Table 5 Lead-Lag Relationship of Returns across Markets

This table reports the results from estimating the following 30-minute frequency VAR(1) regression over the sample period from October 23, 2024, 18:00:00 EST to November 5, 2024, 03:00:00 EST:

$$\mathbf{R}_t = \mathbf{c} + \mathbf{\Phi}_1 \mathbf{R}_{t-1} + \mathbf{u}_t,$$

where the dependent variable $\mathbf{R}_t$ is a 2×1 vector of returns (in %) for a given contract on Kalshi (R_t^K) and Polymarket (R_t^P) at the t-th 30-minute interval. Reported are the regression coefficient estimates and their associated t-statistics, calculated using the White (1980) heteroskedasticity-robust standard errors. *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively.

	Dependent Variables			
	R_t^K		R_t^P	
Contracts	DJT (1)	KH (2)	DJT (3)	KH (4)
R_{t-1}^K	-0.225***	-0.174***	0.006	0.068*
	[-3.47]	[-3.03]	[0.19]	[1.79]
R_{t-1}^P	0.260**	0.252***	-0.028	-0.067
	[2.43]	[3.13]	[-0.41]	[-1.07]
Intercept	-0.028	0.080	-0.031	0.056
	[-0.46]	[1.00]	[-0.78]	[0.93]
Observations	408	408	408	408
R^2	5.36%	4.95%	0.07%	1.04%

Table 6 Lead-Lag Relationship of Returns by Relative Liquidity

This table reports the results from estimating the following 30-minute frequency VAR(1) regressions over the high and low relative liquidity ratio periods, respectively:

$$\mathbf{R}_t = \mathbf{c} + \mathbf{\Phi}_1 \mathbf{R}_{t-1} + \mathbf{u}_t,$$

where the dependent variable $\mathbf{R}_t$ is a 2×1 vector of returns (in %) for a given contract on Kalshi (R_t^K) and Polymarket (R_t^P) at the t-th 30-minute interval. The relative liquidity measure for each contract is computed as the *inverse* of the Polymarket to Kalshi standardized Amihud (2002) illiquidity ratio. The Amihud (2002) illiquidity is computed as the absolute past 6-hour cumulative return divided by the corresponding 6-hour dollar trading volume. We define the high (low) liquidity ratio period using the median cut. Reported are the regression coefficient estimates and their associated t-statistics, calculated using the White (1980) heteroskedasticity-robust standard errors. *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively.

Panel A: DJT Contract

	Dependent Variables			
	R_t^K		R_t^P	
Polymarket/Kalshi Liquidity Ratio	High (1)	Low (2)	High (3)	Low (4)
R_{t-1}^K	-0.293***	-0.166*	-0.005	0.015
	[-3.10]	[-1.69]	[-0.11]	[0.30]
R_{t-1}^P	0.355***	0.186	0.107	-0.169
	[2.85]	[1.00]	[1.19]	[-1.64]
Intercept	0.025	-0.097	0.002	-0.066
	[0.29]	[-1.00]	[0.04]	[-1.06]
Observations	189	189	189	189
R^2	10.19%	2.53%	1.24%	2.34%

Panel B: KH Contract

	Dependent Variables			
	R_t^K		R_t^P	
Polymarket/Kalshi Liquidity Ratio	High (1)	Low (2)	High (3)	Low (4)
R_{t-1}^K	-0.227***	-0.115	0.068	0.06
	[-2.64]	[-1.46]	[1.23]	[1.09]
R_{t-1}^P	0.277***	0.279*	0.010	-0.138
	[2.99]	[1.67]	[0.13]	[-1.30]
Intercept	0.037	0.142	0.027	0.081
	[0.35]	[1.06]	[0.31]	[0.87]
Observations	189	189	189	189
R^2	8.42%	3.27%	1.00%	1.81%

Table 7 Lead-Lag Relationship of Returns by Relative Trading Volumes

This table reports the results from estimating the following 30-minute frequency VAR(1) regressions over the high and low relative volume ratio period, respectively:

$$\mathbf{R}_t = \mathbf{c} + \mathbf{\Phi}_1 \mathbf{R}_{t-1} + \mathbf{u}_t,$$

where the dependent variable $\mathbf{R}_t$ is a 2×1 vector of returns (in %) for a given contract on Kalshi (R_t^K) and Polymarket (R_t^P) at the t-th 30-minute interval. The relative volume ratio is defined as the past 6-hour standardized trading volume of the DJT or KH contract on the Polymarket platform divided by that of the corresponding contract on the Kalshi platform. We define the high (low) relative volume period as the period when the Polymarket to Kalshi volume ratio is above (below) the median of the whole sample period. Reported are the regression coefficient estimates and their associated t-statistics, calculated using the White (1980) heteroskedasticity-robust standard errors. *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively.

Panel A: DJT Contract

	Dependent Variables			
	R_t^K		R_t^P	
Polymarket/Kalshi Volume Ratio	High (1)	Low (2)	High (3)	Low (4)
R_{t-1}^K	-0.221**	-0.261***	0.081	-0.056*
	[-2.35]	[-2.87]	[1.36]	[-1.79]
R_{t-1}^P	0.373**	0.177	-0.164	0.077
	[2.22]	[1.25]	[-1.28]	[0.97]
Intercept	-0.073	0.000	-0.063	0.001
	[-0.77]	[0.01]	[-0.96]	[0.01]
Observations	190	192	190	192
R^2	5.64%	6.61%	2.06%	1.28%

Panel B: KH Contract

	Dependent Variables			
	R_t^K		R_t^P	
Polymarket/Kalshi Volume Ratio	High (1)	Low (2)	High (3)	Low (4)
R_{t-1}^K	-0.162**	-0.194**	0.03	0.1
	[-2.08]	[-2.21]	[0.64]	[1.61]
R_{t-1}^P	0.289**	0.243*	-0.073	-0.044
	[2.41]	[1.91]	[-0.78]	[-0.47]
Intercept	-0.001	0.186	-0.024	0.110
	[-0.01]	[1.56]	[-0.26]	[1.24]
Observations	190	192	190	192
R^2	5.91%	4.56%	0.59%	1.87%

Table 8 Trading Order Imbalance, Whale Trades and Returns

This table presents the results from regressing the returns on the lagged returns and order imbalances from October 23, 2024, 18:00:00 EST to November 5, 2024, 03:00:00 EST:

$$R_t^p = \beta_0 + \beta_1 R_{t-1}^p + \beta_2 X_t^p + \beta_3 X_{t-1}^p + u_t, \quad p \in \{K, P\},$$

where the dependent variable is returns (in %) for a given contract on Kalshi (R_t^K) and Polymarket (R_t^P) at the t-th 30-minute interval. X is the total order imbalance (OIB) or the order imbalance for whale (OIB^w) and non-whale (OIB^{nw}) trades on the respective platform. The order imbalance is defined as the net buying volume normalized by the total volume, i.e., $(V_{Buy} - V_{Sell})/(V_{Buy} + V_{Sell})$. A transaction is classified as a whale trade if its volume exceeds the 99th percentile of volume distribution on that trading day. All OIB variables are standardized to have zero means and unit variances, allowing the regression coefficients to be interpreted as the impact of a one-standard-deviation change. The t-statistics in brackets are calculated using the White (1980) heteroskedasticity-robust standard errors. *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively.

Panel A: DJT Contract

	Dependent Variables			
	R_t^K		R_t^P	
Independent Variables	(1)	(2)	(3)	(4)
R_{t-1}	-0.219***	-0.214***	-0.056	-0.056
	[-3.43]	[-3.37]	[-0.82]	[-0.82]
OIB_t	0.125***		0.053	
	[2.73]		[1.22]	
OIB_{t-1}	0.239***		0.209***	
	[3.93]		[3.33]	
OIB_t^w		0.048		0.059
		[0.90]		[1.46]
OIB_{t-1}^w		0.235***		0.198***
		[3.63]		[3.59]
OIB_t^{nw}		0.153***		-0.007
		[3.08]		[-0.13]
OIB_{t-1}^{nw}		0.172***		0.064
		[2.81]		[1.02]
Intercept	-0.024	-0.023	-0.031	-0.031
	[-0.42]	[-0.42]	[-0.80]	[-0.79]
Observations	455	455	404	404
R^2	8.95%	9.64%	7.76%	7.89%

Table 8 (Cont'd) Trading Order Imbalance, Whale Trades and Returns

Panel B: KH Contract

	Dependent Variables			
	R_t^K		R_t^P	
Independent Variables	(1)	(2)	(3)	(4)
R_{t-1}	-0.130**	-0.136**	-0.065	-0.066
	[-2.18]	[-2.30]	[-1.04]	[-1.04]
OIB_t	-0.033		0.086	
	[-0.40]		[1.65]	
OIB_{t-1}	0.302***		0.249***	
	[3.14]		[3.70]	
OIB_t^w		0.06		0.078
		[0.69]		[1.47]
OIB_{t-1}^w		0.328***		0.203***
		[3.22]		[3.24]
OIB_t^{nw}		-0.059		0.026
		[-0.76]		[0.37]
OIB_{t-1}^{nw}		0.194**		0.145*
		[2.07]		[1.76]
Intercept	0.074	0.070	0.063	0.063
	[0.92]	[0.88]	[1.07]	[1.07]
Observations	420	420	404	404
R^2	4.87%	6.52%	5.24%	5.44%

Table 9 Lead-Lag Relationship of Returns by Relative Whale Trades

This table reports the results from estimating the following 30-minute frequency VAR(1) regressions over the high and low relative whale trades ratio period, respectively:

$$\mathbf{R}_t = \mathbf{c} + \mathbf{\Phi}_1 \mathbf{R}_{t-1} + \mathbf{u}_t,$$

where the dependent variable $\mathbf{R}_t$ is a 2×1 vector of returns (in %) for a given contract on Kalshi (R_t^K) and Polymarket (R_t^P) at the t-th 30-minute interval. The relative whale trades ratio is defined as the absolute whale order imbalance of the DJT or KH contract on the Polymarket platform divided by that of the corresponding contract on the Kalshi platform. A transaction is classified as a whale trade if its volume exceeds the 99th percentile of volume distribution on that trading day. We define the high (low) relative whale trades period as the period when the Polymarket to Kalshi whale trades ratio is above (below) the median of the whole sample period. The regression is estimated only over intervals with non-zero whale trading activity on both platforms (so the whale OIB ratio is defined). Reported are the regression coefficient estimates and their associated t-statistics, calculated using the White (1980) heteroskedasticity-robust standard errors. *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively.

Panel A: DJT Contract

	Dependent Variables			
	R_t^K		R_t^P	
Polymarket/Kalshi Whale Trades Ratio	High (1)	Low (2)	High (3)	Low (4)
R_{t-1}^K	-0.263**	-0.182	-0.093	0.098*
	[-2.32]	[-1.59]	[-1.46]	[1.92]
R_{t-1}^P	0.480***	0.073	0.162	-0.111
	[3.18]	[0.37]	[1.41]	[-1.29]
Intercept	0.083	-0.069	0.005	-0.039
	[0.67]	[-0.58]	[0.06]	[-0.61]
Observations	129	130	129	130
R^2	8.69%	2.38%	2.05%	3.02%

Panel B: KH Contract

	Dependent Variables			
	R_t^K		R_t^P	
Polymarket/Kalshi Whale Trades Ratio	High (1)	Low (2)	High (3)	Low (4)
R_{t-1}^K	-0.264***	-0.216**	0.011	0.122*
	[-2.95]	[-2.04]	[0.13]	[1.71]
R_{t-1}^P	0.425***	0.305	-0.042	-0.104
	[3.88]	[1.60]	[-0.38]	[-0.75]
Intercept	0.054	0.222	0.125	0.085
	[0.34]	[1.13]	[0.97]	[0.58]
Observations	115	114	115	114
R^2	11.10%	5.22%	0.13%	2.43%

Appendix

A Significant Events in the Prediction Market Space

Year / Period	Event Description
1880s	- Start of curbside election betting.
1936	- Start of Gallup scientific polls. - Commodity Exchange Act (CEA) set up to regulate commodity futures trading; Chicago Board of Trade designated a DCM under the CEA.
1974	- Commodity Futures Trading Commission (CFTC) established to replace the CEA, creating an independent regulatory agency rather than one under the U.S. Department of Agriculture.
1988	- Iowa Electronic Markets (IEM) launched as a small-scale political prediction market by the University of Iowa.
1996	- Hollywood Stock Exchange (HSX) launched, allowing trading of "virtual" movie shares.
2014	- PredictIt received a CFTC no-action letter permitting a small-scale political prediction market under strict conditions.
2016	- PredictIt opened markets for the 2016 U.S. elections.
2018	- Augur, built on the Ethereum blockchain, launched publicly in July 2018; world's first decentralized prediction-market platform.
2020	- Polymarket launched as a blockchain-based platform. - Kalshi was granted CFTC designation as a contract market in November 2020.
2022	- CFTC ordered Polymarket (Blockratize Inc.) to pay a $1.4 million civil penalty and to wind down non-compliant markets, banned in the U.S.
2023	- CFTC disapproved KalshiEX's congressional control event contracts, determining that the political event contracts constituted gaming.
2024	- Polymarket begins election trading in January 2024. - Kalshi granted federal approval to operate prediction markets on elections. - Robinhood begins offering prediction markets following Kalshi's approval.
2025	- Coinbase CEO triggers prediction markets by saying key crypto terms, highlights manipulability of such markets - Kalshi valued at $5B after August funding round. - Google integrates predictions from Polymarket and Kalshi into its engine. - Polymarket acquires QCX for CFTC-regulated U.S. access. - Polymarket Open Interest hits $200M in November; volume $2.4B (vs. Kalshi $42M / $72.8M). - Polymarket receives $2B investment from Intercontinental Exchange, parent company of NYSE.

B Data Collection

Table A1 Data Sources Information

This table outlines the data collected from four retail prediction market platforms: Kalshi, Polymarket, PredictIt, and Robinhood. It specifies the market and data collection periods (in GMT), data sources, availability of pricing information, and the granularity of trade-level data.

	Kalshi	Polymarket	PredictIt	Robinhood
Start Date of Prediction Market	10/4/2024 12:15	1/5/2024 0:00	9/2/2022 0:00	10/28/2024 0:00
End Date of Prediction Market (Market Resolves)	1/20/2025 0:00	11/6/2024 0:00	12/17/2024 0:00	1/6/2025 0:00
Start Date of Data Collection	10/4/2024 12:15	10/23/2024 0:00	9/2/2022 0:00	11/4/2024 0:00
End Date of Data Collection	11/6/2024 20:00	11/6/2024 20:00	11/6/2024 20:00	11/6/2024 20:00
Data Source	Transaction Data Collected from Backend API, Daily Prices from Kalshi's website	Transaction Data Collected using server that tracked live transactions posted to website, Daily Prices from Polymarket's website	Daily Prices from PredictIt's Website and Ballotpedia backend API	Daily Prices from Robinhood's website and limited transaction data from manual reconciliation of screen recordings of the market
Country/Region (At Time of Data Collection)	US residents only	Global except US residents	US residents only	US, UK, EU, and Australia residents only
Market Mechanism	Listed Order Book	Listed Order Book	Listed Order Book	Binary match engine
Daily Price Availability	Y	Y	Y	Y
Trade-level Information Availability	Y	Y	N	Y

B.1 Market Mechanisms

In this section, we first provide an overview of the various market mechanisms before going into the data collection process. The prediction markets are largely similar in their market microstructure design. Kalshi, Polymarket, and PredictIt operate on traditional Limit Order Books (LOBs), and Robinhood employs a binary match engine that only allows directional "yes" purchases on each candidate.

LOBs, the dominant architecture in traditional equity markets, facilitate liquidity through the interaction of patient and impatient traders. As shown by Foucault, Kadan, and Kandel (2005), patient traders post limit orders and supply liquidity by accepting execution risk and delay, while impatient traders submit market orders to demand immediacy. Liquidity in LOBs is thus endogenous and varies with the composition of trader types and the frequency of order arrivals. When patient traders dominate, limit orders become more aggressive, spreads narrow, and the market becomes more resilient. Conversely, in periods of low participation or high impatience, liquidity deteriorates as spreads widen and execution delays lengthen.

A key point to note is that Polymarket had utilized another market-making mechanism, known as the automated market maker (AMM), from its inception in 2020 to early 2023. AMMs allow traders to buy or sell directly against a smart contract governed by a deterministic pricing formula. This design guarantees liquidity and immediate execution, regardless of market depth, and eliminates the need for a matching counterparty. However, AMMs are vulnerable to price inefficiencies, especially during periods of rapid price movement, and liquidity providers may incur arbitrage-driven losses unless compensated by transaction fees (Pourpouneh, Nielsen, and Ross, 2020). Based on our reading of Polymarket's Discord server log, it appears that there was rampant "AMM stalking" activity, which led to the transition towards LOB. "AMM stalking" refers to traders exploiting latency by monitoring the memory pool or refreshing contract states to front-run large or slow retail trades.[23] By arbitraging against the AMM's pricing curve before it could adjust, these actors profited systematically, undermining market integrity and deterring honest liquidity provision. In May 2023, Polymarket had largely deprecated its AMM (See Appendix Figure A3), instead choosing to use a "rewards for liquidity providers" for users who made resting limited orders (Polymarket, 2025).

B.2 Polymarket Data Collection

We collected the transaction data for the following dates (All timings in GMT). We began our data collection on October 23, 2024, at 21:41:19 and continued until November 6, 2024, at 19:45:36. We also performed nominal trades and confirmed that the activities shown in Figure A2 are trades that close the deal, and not activity of listed trades. In other words, only the completing party's trade is shown, and not the originating counterparty. During this period, we encountered issues with the API server. Thus, we caveat our results with the following gaps in the data.

- Thursday, October 24, 2024 01:12:10 to Thursday, October 24, 2024 04:29:44

[23]For example, suppose Alice submits a large order to buy Yes shares on an AMM-based market, which would raise the price from 60 to 70 cents. While her trade is waiting in the mempool, a faster trader, Bob, sees it and buys shares at 60 cents before Alice's order is confirmed. Once Alice's trade executes and pushes the price to 70 cents, Bob immediately sells his shares at the higher price, locking in a risk-free 10-cent profit per share by exploiting her trade. This is also known as a sandwich attack.

- Saturday, October 26, 2024 16:16:51 to Monday, October 28, 2024 02:30:21

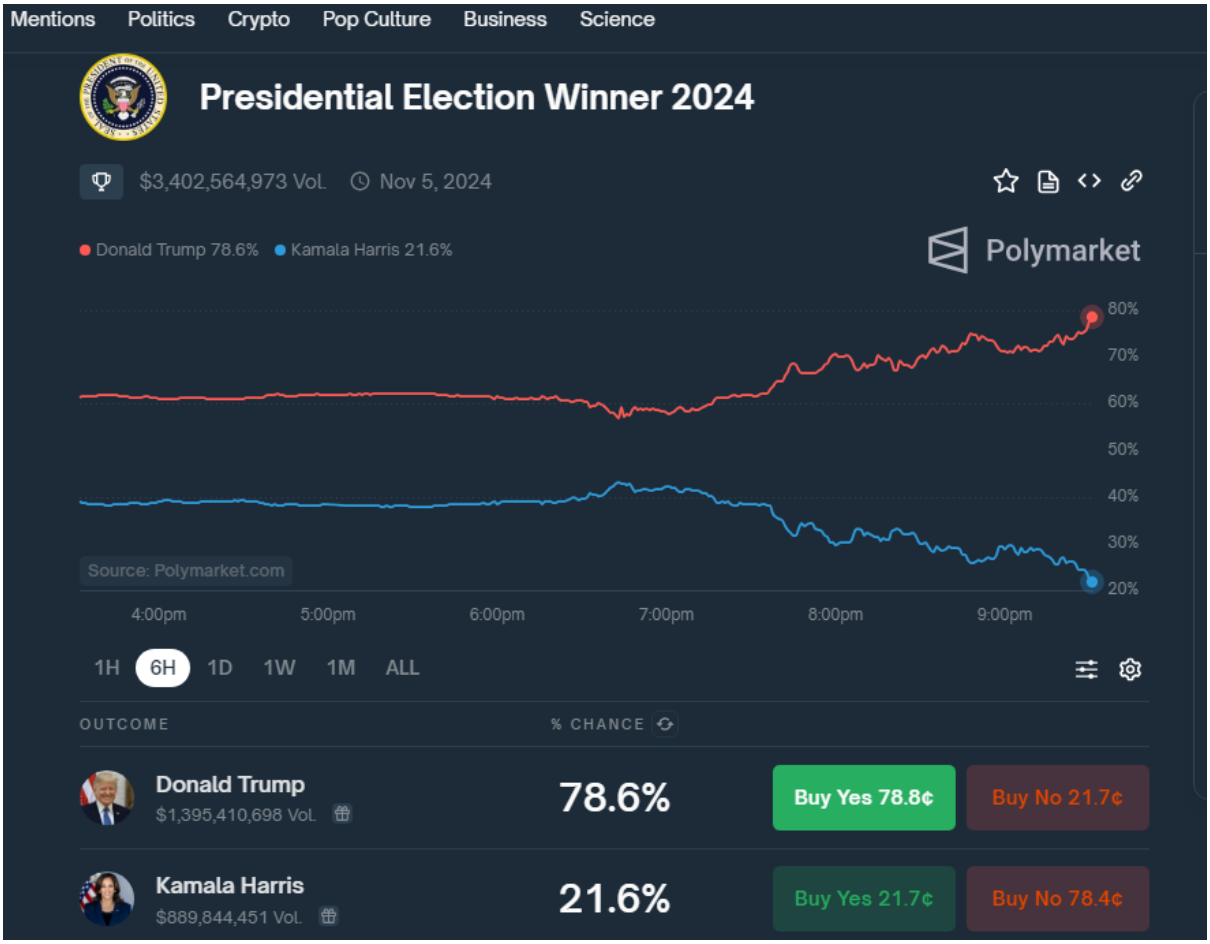
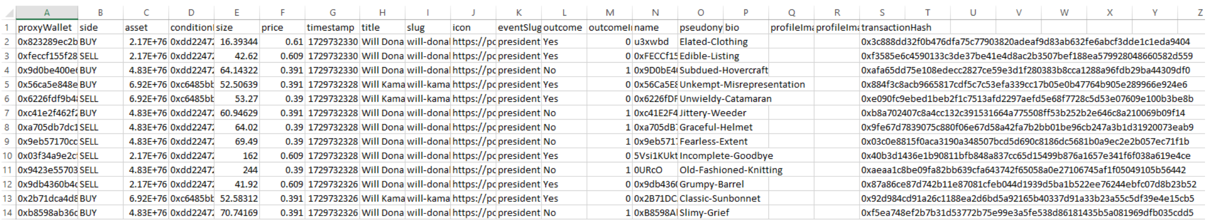


Figure A1. Screenshot of Polymarket Interface



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	proxyWallet	side	asset	conditionI	size	price	timestamp	title	slug	icon	eventSlug	outcome	outcomeI	name	pseudony	bio	profileIma	profileIma	transactionHash
2	0x823289ec2b	BUY	2.17E+76	0xdd22472	16.39344	0.61	1729732330	Will Dona	will-donal	https://pc	president	Yes	0	u3xwbd	Elated-Clothing				0x3c888dd32f0b476dfa75c77903820adeaf9d83ab632fe6abcf3dde1c1eda9404
3	0xfeccf155f28	SELL	2.17E+76	0xdd22472	42.62	0.609	1729732330	Will Dona	will-donal	https://pc	president	Yes	0	0xFECCf15	Edible-Listing				0xf3585e6c4590133c3de37be41e4d8ac2b3507bef188ea579928048660582d559
4	0x9d0be400e	BUY	4.83E+76	0xdd22472	64.14322	0.391	1729732330	Will Dona	will-donal	https://pc	president	No	1	0x9D0bE4	Subdued-Hovercraft				0xafa65dd75e108edecc2827ce59e3d1f280383b8cca1288a96fdb29ba44309df0
5	0x56ca5e848e	BUY	6.92E+76	0xc6485bb	52.50639	0.391	1729732328	Will Kama	will-kama	https://pc	president	Yes	0	0x56Ca5E8	Unkempt-Misrepresentation				0x884f3c8acb9665817cdf5c7c53efa339cc17b05e0b47764b905e289966e924e6
6	0x6226fdf9b4	SELL	6.92E+76	0xc6485bb	53.27	0.39	1729732328	Will Kama	will-kama	https://pc	president	Yes	0	0x6226fDF	Unwieldy-Catamaran				0xe090fc9ebed1beb2f1c7513afd2297aefd5e68f7728c5d53e07609e100b3be8b
7	0xc41e2f462f2	BUY	4.83E+76	0xdd22472	60.94629	0.391	1729732328	Will Dona	will-donal	https://pc	president	No	1	0xc41E2F4	Jittery-Weeder				0xb8a702407c8a4cc132c391531664a775508ff53b252b2e646c8a210069b09f14
8	0xa705db7dc1	SELL	4.83E+76	0xdd22472	64.02	0.39	1729732328	Will Dona	will-donal	https://pc	president	No	1	0xa705dB7	Graceful-Helmet				0x9fe67d7839075c880f06e67d58a42fa7b2bb01be96cb247a3b1d31920073eab9
9	0x9eb57170cc	SELL	4.83E+76	0xdd22472	69.49	0.39	1729732328	Will Dona	will-donal	https://pc	president	No	1	0x9eb5717	Fearless-Extent				0x03c0e8815f0aca3190a348507bcd5d690c8186dc5681b0a9ec2e2b057ec71f1b
10	0x03f34a9e2c	SELL	2.17E+76	0xdd22472	162	0.609	1729732328	Will Dona	will-donal	https://pc	president	Yes	0	5Vsi1KUkt	Incomplete-Goodbye				0x40b3d1436e1b90811bfb848a837cc65d15499b876a1657e341f6f038a619e4ce
11	0x9423e55703	SELL	4.83E+76	0xdd22472	244	0.39	1729732328	Will Dona	will-donal	https://pc	president	No	1	0URcO	Old-Fashioned-Knitting				0xaeaa1c8be09fa82bb639cfa643742f65058a0e27106745af1f05049105b56442
12	0x9db4360b4c	SELL	2.17E+76	0xdd22472	41.92	0.609	1729732326	Will Dona	will-donal	https://pc	president	Yes	0	0x9db436	Grumpy-Barrel				0x87a86ce87d742b11e87081cfeb044d1939d5ba1b522ee76244ebfc07d8b23b52
13	0x2b71dca4d8	BUY	6.92E+76	0xc6485bb	52.58312	0.391	1729732326	Will Kama	will-kama	https://pc	president	Yes	0	0x2B71DC	Classic-Sunbonnet				0x92d984cd91a26c1188ea2d6bd5a92165b40337d91a33b23a55c5df39e4e15cb5
14	0xb8598ab36c	BUY	4.83E+76	0xdd22472	70.74169	0.391	1729732326	Will Dona	will-donal	https://pc	president	No	1	0xB8598A	Slimy-Grief				0xf5ea748ef2b7b31d53772b75e99e3a5fe538d86181435b5a081969dfb035cdd5

Figure A2. Screenshot of Polymarket Transaction Data

We also provide screenshots of the trail of Polymarket's changes from AMM to the Central LOB. Polymarket initially employed an Automated Market Maker (AMM) model for all trading activities. On December 12, 2022, a Polymarket admin announced that an Order Book was live for public beta testing on a small number of markets, marking the beginning of a phased shift in trading infrastructure.

By January 2023, support clarified that some markets now used order books, while others still relied on AMMs. On January 25, the admin update noted backend improvements: liquidity displays were updated to reflect both sources, and bugs related to limit orders and UI pricing were resolved.

On April 3, 2023, an admin stated that Polymarket had “mostly transitioned to order book” and that the vast majority of liquidity was now there.

By May 30, 2023, the transition was effectively complete. An admin confirmed that “AMM is a legacy product at this point”, with most markets having no AMM liquidity and liquidity fully migrated to the order book.

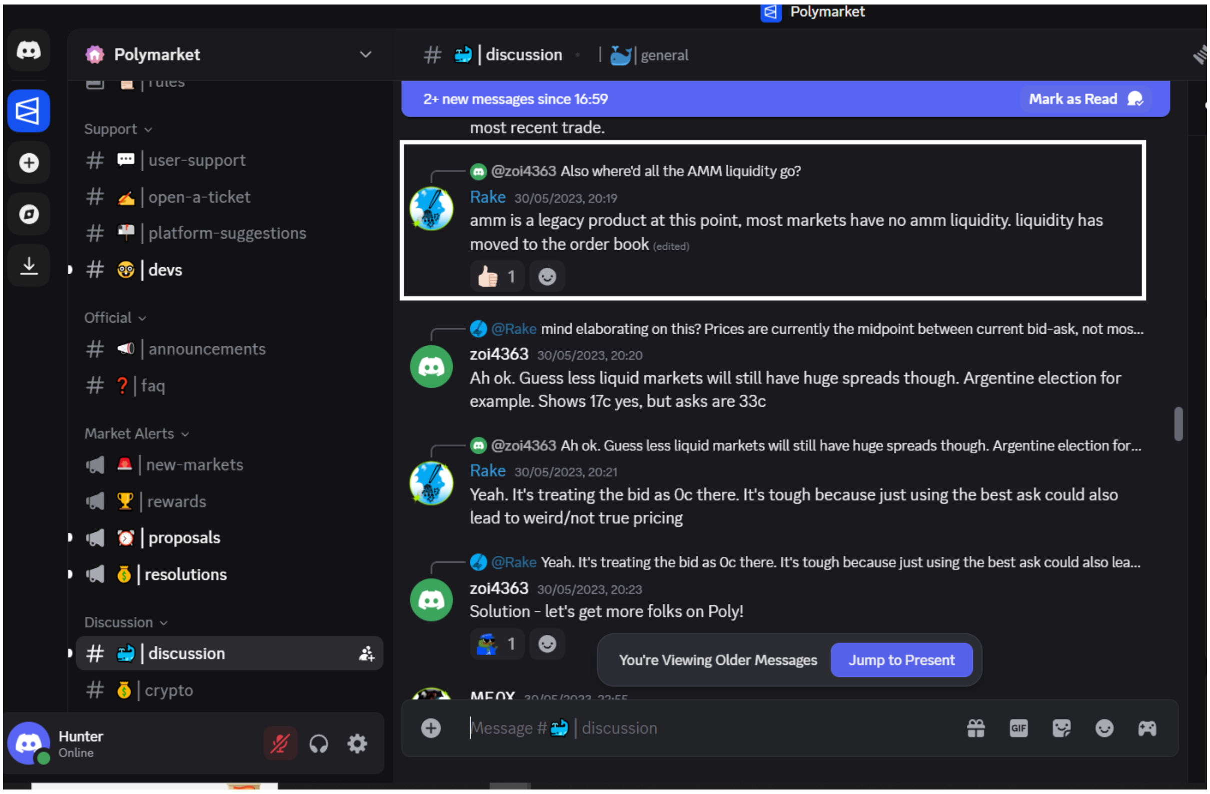


Figure A3. Screenshot of Polymarket Historical Discord Server Log

B.3 Kalshi Data Collection

We collected the transaction data for the following dates (All timings in GMT): Friday, October 4, 2024, 12:15:05 to Wednesday, November 6, 2024, 18:59:50

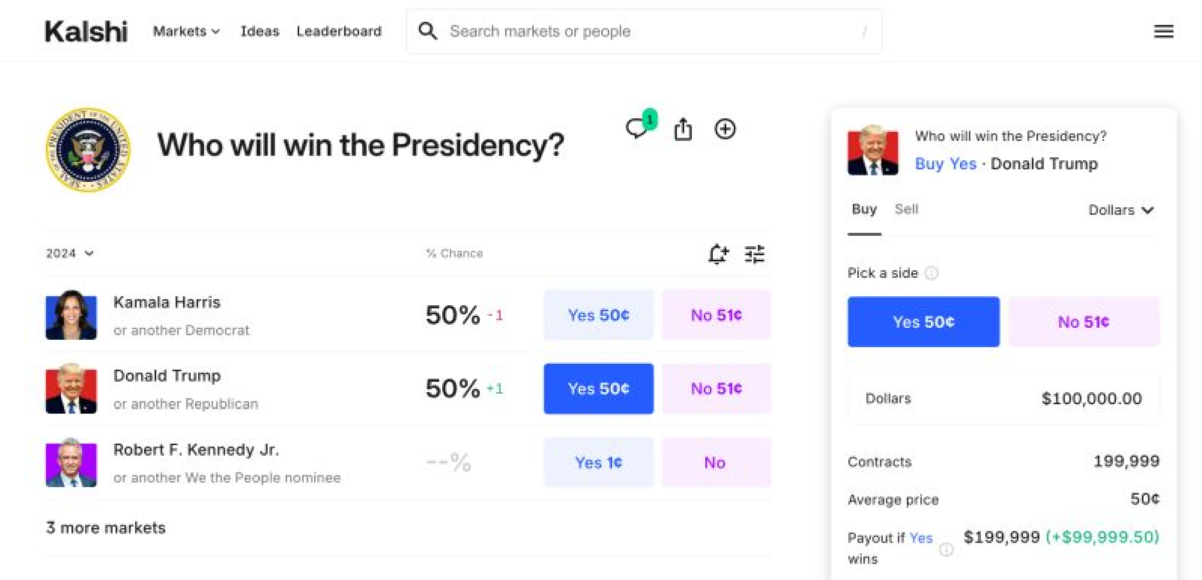


Figure A4. Screenshot of Kalshi Interface

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	trade_id	market_ic	ticker	price	count	taker_side	maker_ac	taker_acti	maker_nic	taker_nick	create_date		
2	42dc79dd-	ac7c5a09-!	PRES-2024	97	16	no	buy	sell			2024-11-06T18:59:50.160938Z		
3	b76fa492-	0e332816-	PRES-2024	1	2016	no	buy	sell			2024-11-06T18:59:49.771754Z		
4	9d488d3e-	0e332816-	PRES-2024	1	483	no	buy	sell			2024-11-06T18:59:49.771754Z		
5	d339f711-	0e332816-	PRES-2024	1	263	no	buy	sell			2024-11-06T18:59:49.007171Z		
6	f40d16af-(	ac7c5a09-!	PRES-2024	98	111	yes	sell	sell			2024-11-06T18:59:48.768853Z		
7	3435c497-	ac7c5a09-!	PRES-2024	97	166	no	buy	sell			2024-11-06T18:59:48.646955Z		
8	5454423a-	ac7c5a09-!	PRES-2024	97	112	no	buy	sell			2024-11-06T18:59:48.479764Z		

Figure A5. Screenshot of Kalshi Transaction Data

B.4 PredictIt and Robinhood Data Collection

For PredictIt, no transaction data is made public. The interface only shows the current prices. For Robinhood, the market is only viewable on mobile, and only the incoming new transaction amounts are shown. We also could not monitor the network or reverse-engineer the incoming transactions, which is probably due to Robinhood having enhanced protection methods, such as SSL pinning and obfuscation techniques. We used a proxy server to monitor the API calls, but this approach was also ineffective, as the server was highly secure. Finally, we recorded the entire market transactions as a video and hand-collected data.

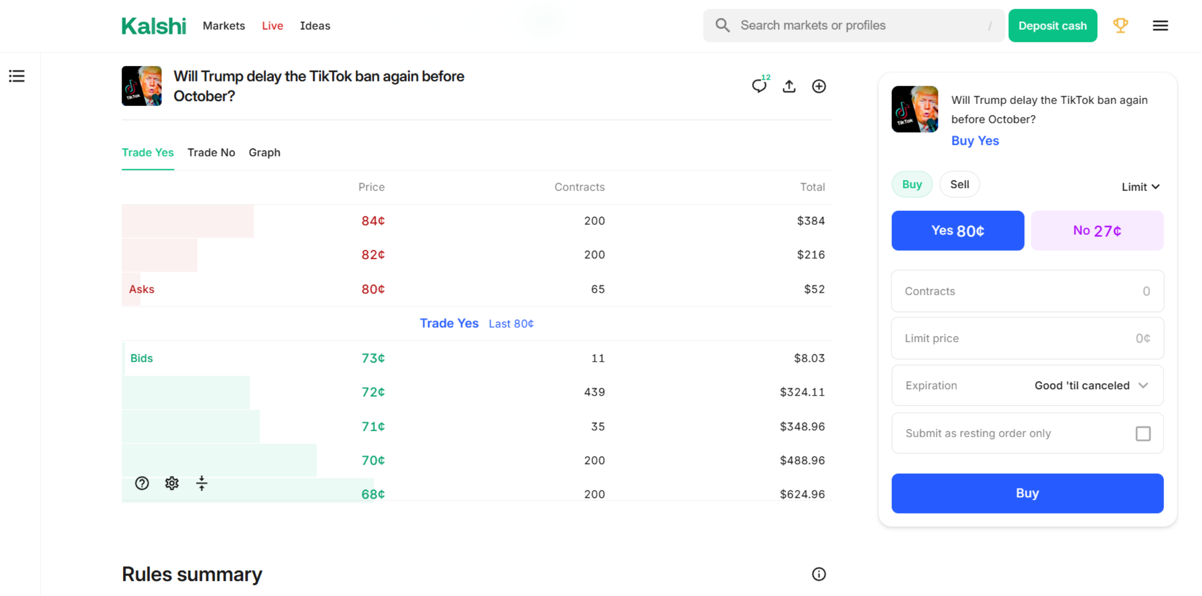


Figure A6. Screenshot of Kalshi Listed Order Book for the market on whether President Donald Trump will delay the TikTok ban before October 2025 — Retrieved on 28th June 2025

Figure A7. Screenshot of PredictIt Interface

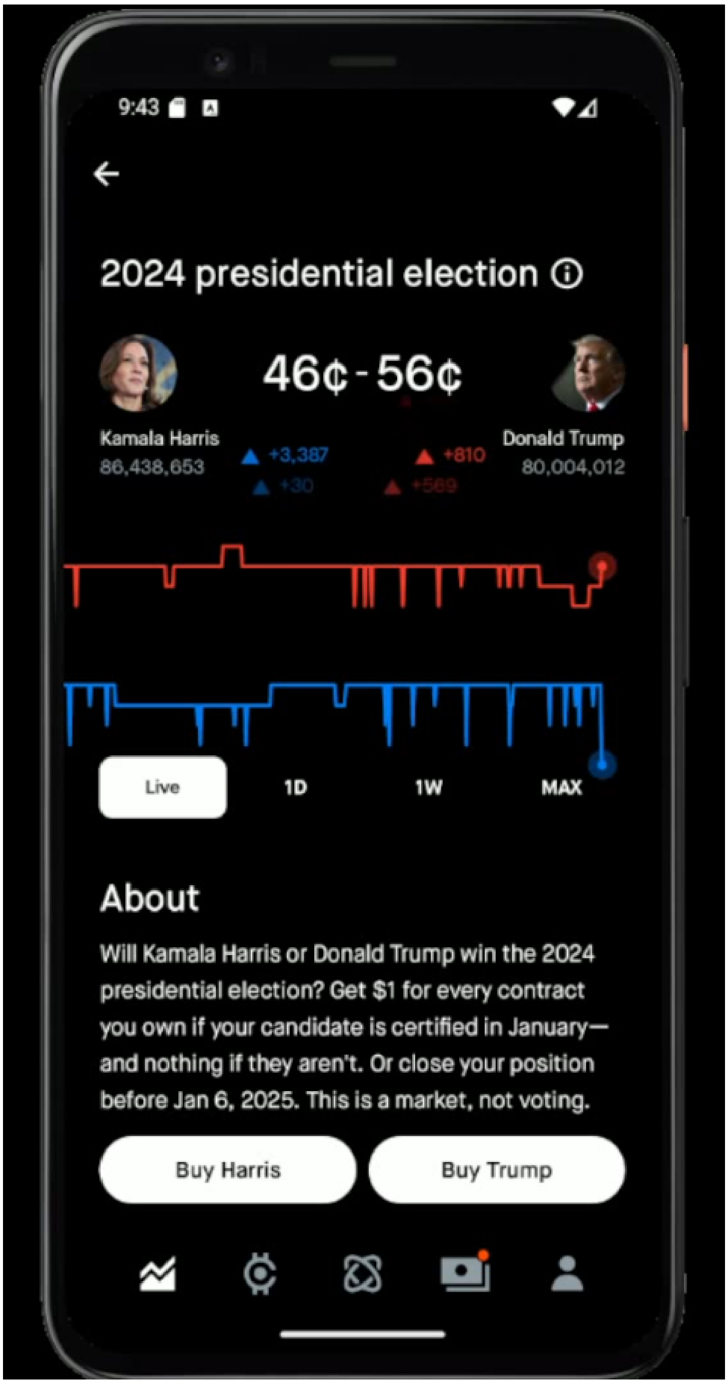


Figure A8. Screenshot of Robinhood Interface

C Additional Results

Table A2 Within-Platform Arbitrage on Kalshi and Polymarket

This table reports on within-platform arbitrage opportunities on Kalshi and on Polymarket from October 23, 2024, 18:00:00 EST to November 5, 2024, 23:59:59 EST, respectively. We list the arbitrage profits for different arbitrage intervals and by the platform where we take a bullish position on DJT. We consider two scenarios: one without a cost of arbitrage (i.e., $\delta = 0$) and another with a 2% arbitrage cost (i.e., $\delta = 2\%$). For each platform, Kalshi or Polymarket, we compute arbitrage profits by aggregating the difference between \$1 (or \$0.98 when incorporating a 2% transaction cost) and the matched sum of the lowest available prices from the same platform within each interval, multiplied by the corresponding trade size. The arbitrage profits are in dollars.

Interval	Platform	Total Violation Value (Without Fee)	Total Violation Value (2% Fee)
1-Second	Kalshi	2,955.34	256.06
1-Second	Polymarket	27,161.04	1,276.43
1-Minute	Kalshi	14,448.40	1,902.62
1-Minute	Polymarket	101,441.20	5,436.63
5-Minute	Kalshi	44,377.71	6,627.88
5-Minute	Polymarket	281,140.14	21,677.82