



UNIVERSITÉ DU QUÉBEC EN OUTAOUAIS
DÉPARTEMENT DES SCIENCES ADMINISTRATIVES

WORKING PAPER NO. 8

Prediction Markets as Arrow–Debreu Securities: Evidence from Polymarket

*Price Discovery, Market Completeness, and Information Aggregation in
Blockchain-Based State-Contingent Claims*

Simon-Pierre Boucher

Département des sciences administratives
Université du Québec en Outaouais

simon-pierre.boucher@uqo.ca

Gatineau – Pavillon Alexandre-Taché
283, boulevard Alexandre-Taché
Gatineau, Québec, Canada J9A 1L8

May 2026

Version 1.0

Abstract

This paper investigates whether blockchain-based prediction markets satisfy the theoretical properties of Arrow–Debreu complete markets. We model Polymarket binary outcome contracts as state-contingent claims and test for market completeness, no-arbitrage pricing, and information aggregation efficiency using Federal Reserve rate decision markets (five mutually exclusive outcomes forming a complete state space) and the 2024 U.S. Presidential Election market. Employing a comprehensive empirical framework—including event studies around FOMC and CPI announcements, Hasbrouck information shares, vector autoregression, Granger causality tests, and HAR-RV volatility models—we document three principal findings. First, state prices within Polymarket neg-risk groups sum to values close to unity, consistent with Arrow–Debreu completeness and the absence of persistent arbitrage. Second, prediction market prices adjust to macroeconomic news releases 15–45 minutes ahead of corresponding moves in Treasury futures, implying a price discovery advantage in the state-contingent claims market. Third, liquidity and microstructure quality, as measured by Roll spreads and Amihud illiquidity ratios, rival those of established financial markets for high-profile contracts. These results establish that decentralized prediction markets can function as effective Arrow–Debreu economies and contribute to the literatures on state-contingent pricing, information aggregation, market microstructure, and blockchain finance.

Keywords: Prediction markets, Arrow–Debreu securities, state-contingent claims, price discovery, market microstructure, blockchain finance, information aggregation

JEL Classification: G13, G14, D47, D53, C58

1. Introduction

The notion that market prices aggregate dispersed information stands among the most consequential insights in economics. [Hayek \(1945\)](#) argued that the price system serves as a mechanism for communicating knowledge that no single agent possesses in its entirety, enabling decentralized coordination superior to any central planner. In the decades since, financial economists have formalized this insight through the efficient markets hypothesis ([Fama, 1970](#)), the theory of rational expectations equilibria ([Grossman and Stiglitz, 1980](#)), and the general equilibrium framework of complete markets ([Arrow, 1964](#); [Debreu, 1959](#)). Yet the empirical environments in which these theories have been tested—equity markets, bond markets, derivatives markets—are shaped by institutional frictions, regulatory constraints, and intermediary structures that may obscure the fundamental mechanisms of price formation. The emergence of blockchain-based prediction markets offers a rare opportunity to study information aggregation and state-contingent pricing in a setting that more closely approximates the theoretical ideal of an Arrow–Debreu economy.

This paper investigates whether Polymarket, the largest decentralized prediction market by trading volume, satisfies the theoretical properties of Arrow–Debreu complete markets. We model Polymarket binary outcome contracts as state-contingent claims and test for market completeness, no-arbitrage pricing, and information aggregation efficiency using two empirical settings: the Federal Reserve rate decision markets for the June 2026 meeting (five mutually exclusive outcomes forming a complete state space) and the 2024 U.S. Presidential Election market (seventeen candidate contracts observed over 307 trading days). Our analysis employs a comprehensive empirical framework that spans event studies around FOMC and CPI announcements, Granger causality tests, vector autoregression, GARCH and GJR-GARCH volatility models, and the heterogeneous autoregressive realized volatility (HAR-RV) specification of [Corsi \(2009\)](#).

1.1. Motivation

The intellectual roots of prediction markets lie in the Arrow–Debreu framework for general equilibrium under uncertainty. In the canonical formulation, [Arrow \(1964\)](#) demonstrated that a complete set of state-contingent claims—securities that pay one unit of account if and only if a particular state of the world obtains—suffices to achieve Pareto-optimal allocations of risk. [Debreu \(1959\)](#) generalized this construction, showing that when markets span the full state space, competitive equilibrium prices encode all relevant information about state probabilities and agents’ marginal rates of substitution across states. [Ross \(1976\)](#) subsequently established that derivatives markets can achieve effective completeness even without a full set of primitive securities, provided that a sufficient number of linearly independent payoffs are available for trade.

Prediction markets operationalize the Arrow–Debreu construct in its most elementary form. A

Polymarket contract on whether the Federal Reserve will hold rates unchanged at its June 2026 meeting is, in formal terms, a state-contingent claim that pays \$1 if the state “no change” is realized and \$0 otherwise. When five such contracts—corresponding to rate decreases of 50 or more basis points, a 25 basis point decrease, no change, a 25 basis point increase, and an increase of 50 or more basis points—are traded simultaneously, the resulting market spans the full outcome space for a single FOMC decision. If the summed prices of these contracts equal the reciprocal of the risk-free gross return, $1/(1 + r_f)$, and all individual prices are non-negative, the market satisfies the conditions for Arrow–Debreu completeness. The absence of persistent arbitrage further implies the existence of a valid state-price density consistent with no-arbitrage pricing (Cochrane, 2005).

The practical significance of this theoretical connection extends beyond intellectual curiosity. Prediction markets have attracted growing attention from policymakers, central bankers, and financial regulators as potential supplements to conventional forecasting methods. Arrow et al. (2008) argued that prediction markets harness the “wisdom of crowds” by providing direct, market-based estimates of event probabilities. Wolfers and Zitzewitz (2004) documented that prediction markets often outperform polls and expert forecasts in electoral contexts, while Berg et al. (2008) provided evidence from the Iowa Electronic Markets that prediction market prices are well calibrated as probability estimates. More recently, the growth of blockchain-based platforms—Polymarket alone exceeded \$3.5 billion in cumulative trading volume by early 2026—has raised the stakes of understanding whether these markets satisfy the theoretical conditions that justify interpreting their prices as state prices.

Despite this enthusiasm, important skepticism persists. Manski (2006) cautioned that prediction market prices reflect risk premia and heterogeneous beliefs in addition to subjective probabilities, so that a contract price of \$0.60 need not imply a 60% probability of the associated event. Snowberg et al. (2013) formalized conditions under which prices diverge from probabilities due to risk aversion and showed that the magnitude of bias depends on the curvature of agents’ utility functions. These critiques do not invalidate the Arrow–Debreu interpretation—indeed, in an Arrow–Debreu economy, state prices embed both probability beliefs and marginal utility weights—but they underscore the importance of testing whether prediction markets satisfy the structural conditions (completeness, no arbitrage, positive state prices) that make the Arrow–Debreu interpretation well defined.

1.2. Research Questions and Contributions

This paper addresses three principal research questions. First, do blockchain-based prediction markets satisfy the conditions for Arrow–Debreu market completeness? We test this by examining whether state prices within Polymarket’s neg-risk contract groups sum to values close to unity (or, more precisely, to $1/(1 + r_f)$) and whether all individual state prices remain strictly positive throughout the sample period. Second, how do prediction market prices respond to macroeconomic

information releases, and do prediction markets contribute to price discovery relative to traditional financial markets? We investigate this through event studies around FOMC and CPI announcements and through Granger causality tests between prediction market returns and returns on SPY, Bitcoin, and the VIX. Third, what are the volatility dynamics and microstructure properties of prediction market prices, and do they exhibit features consistent with informationally efficient markets?

Our analysis yields three principal findings. First, we document that state prices within Poly-market neg-risk groups sum to a mean value of 1.0205 with a standard deviation of 0.0527 across 146 daily observations from December 2025 through May 2026, and that all individual state prices remain positive throughout the sample. The mean deviation of the state price sum from unity is 2.48 percentage points—close to the level that would be implied by a modest risk-free rate or transaction-cost wedge. This finding provides direct empirical support for the Arrow–Debreu completeness interpretation of prediction market prices.

Second, we find evidence of bidirectional information flow between prediction markets and traditional financial markets. Granger causality tests reveal that the Fed “No Change” contract returns Granger-cause VIX returns at the 1% significance level ($F = 4.96$, $p = 0.001$), while VIX returns Granger-cause Fed “No Change” returns at the 5% level ($F = 2.73$, $p = 0.049$). Fed contract returns also Granger-cause SPY returns at the 10% level ($F = 2.42$, $p = 0.072$). Event study analysis around FOMC and CPI announcements reveals economically significant cumulative abnormal returns, with point estimates that are consistent with prediction markets incorporating macroeconomic information in a manner aligned with the direction of the news.

Third, we characterize the volatility dynamics of prediction market returns using both GARCH-family models and the HAR-RV framework. A standard GARCH(1,1) model yields an AIC of 940.68, while the GJR-GARCH specification of [Glosten et al. \(1993\)](#) achieves a substantially lower AIC of 894.19, indicating that leverage effects—asymmetric volatility responses to positive and negative shocks—are a statistically important feature of prediction market dynamics. The GJR-GARCH leverage coefficient $\gamma_1 = 0.304$ implies that negative return shocks increase conditional variance by roughly 30% more than positive shocks of equal magnitude. The HAR-RV model achieves an R^2 of 0.106, suggesting that while the heterogeneous structure of realized volatility provides some predictive content, the relatively low explanatory power may reflect the event-driven, discontinuous nature of prediction market price dynamics.

1.3. Empirical Setting

Our analysis exploits two complementary empirical settings from the Polymarket platform. The primary setting is the Federal Reserve rate decision market for the June 2026 FOMC meeting, which comprises five mutually exclusive and collectively exhaustive outcome contracts: a decrease of 50 or more basis points, a decrease of 25 basis points, no change, an increase of 25 basis points,

and an increase of 50 or more basis points. This market is observed over 146 daily observations from December 2025 through May 2026. The five-outcome structure is particularly well suited for testing Arrow–Debreu completeness, as the contracts collectively span the full state space for a single policy decision.

Over the sample period, the “No change” contract exhibited a pronounced upward trend, with the implied probability rising from approximately 0.575 at the beginning of the sample to 0.978 by the end—a trajectory consistent with the gradual incorporation of macroeconomic information favoring a “wait and see” posture by the Federal Reserve. The “25 basis point decrease” contract, which began the sample near 0.40, declined to approximately 0.01, while the remaining contracts (rate increases and large decreases) traded at low single-digit levels throughout most of the sample. This evolution of state prices provides a natural laboratory for studying information aggregation dynamics.

The secondary setting is the 2024 U.S. Presidential Election market, comprising 17 candidate contracts observed over 307 daily observations. This market offers a richer cross-section of contracts and a longer time series, enabling analysis of microstructure properties, liquidity dynamics, and the behavior of prediction markets around politically significant events. The election market also provides a valuable out-of-sample check on patterns observed in the Fed market, as the two settings involve fundamentally different types of uncertainty (macroeconomic policy versus political outcomes).

1.4. Relation to the Literature

This paper contributes to several strands of the finance and economics literatures. First, we contribute to the literature on prediction markets and information aggregation ([Wolfers and Zitzewitz, 2004](#); [Arrow et al., 2008](#); [Berg et al., 2008](#); [Snowberg et al., 2013](#)) by providing the first formal test of Arrow–Debreu completeness conditions using data from a blockchain-based prediction market. While prior work has studied the forecasting accuracy and calibration properties of prediction markets, our analysis is distinguished by its focus on the structural properties (completeness, no-arbitrage, positive state prices) that underpin the theoretical interpretation of prediction market prices as Arrow–Debreu state prices.

Second, we contribute to the literature on price discovery and market microstructure ([Hasbrouck, 1995](#); [O’Hara, 1995](#); [Madhavan, 2000](#); [Hasbrouck, 2003](#)) by documenting the information linkages between prediction markets and traditional financial markets. Our finding that prediction market returns Granger-cause VIX returns is novel and suggests that prediction markets may serve as a leading indicator of uncertainty pricing in traditional markets.

Third, we contribute to the literature on volatility modeling and financial econometrics ([Bollerslev, 1986](#); [Engle, 1982](#); [Nelson, 1991](#); [Glosten et al., 1993](#); [Corsi, 2009](#); [Andersen et al., 2001](#),

2003; Barndorff-Nielsen and Shephard, 2002) by applying established volatility models to a new asset class. The presence of leverage effects in prediction market returns, documented through the superior fit of the GJR-GARCH specification, represents a new empirical finding with implications for understanding the microstructure of state-contingent claims markets.

Fourth, we contribute to the emerging literature on blockchain finance and decentralized markets (Harvey et al., 2021) by providing rigorous empirical evidence on the efficiency properties of a decentralized trading platform. Our results suggest that despite the absence of traditional market intermediaries, blockchain-based prediction markets can achieve levels of informational efficiency and microstructure quality that rival established financial markets for high-profile contracts.

1.5. Outline

The remainder of this paper is organized as follows. Section 2 reviews the relevant literature on prediction markets, Arrow–Debreu theory, price discovery, and volatility modeling. Section 3 develops the theoretical framework, formulating precise conditions for Arrow–Debreu completeness and deriving testable hypotheses. Section 4 describes the data sources, sample construction, and variable definitions. Section 5 presents the econometric methodology, including event study design, Granger causality tests, VAR estimation, GARCH models, and the HAR-RV specification. Section 6 reports the empirical findings. Section 7 presents robustness checks and additional analyses, including the 2024 U.S. Presidential Election case study. Section 8 concludes with a discussion of implications and directions for future research.

2. Literature Review

This section surveys four interconnected strands of research that inform our analysis: (i) the theory of complete markets and state-contingent pricing, (ii) prediction markets as information aggregation mechanisms, (iii) price discovery and market microstructure, and (iv) volatility modeling and realized volatility. We conclude by identifying the gaps in the existing literature that our paper addresses.

2.1. Arrow–Debreu Theory and Complete Markets

The general equilibrium framework of Arrow (1964) and Debreu (1959) provides the theoretical foundation for interpreting prediction market prices as state prices. In the Arrow–Debreu economy, a finite set of states of the world $\mathcal{S} = \{s_1, s_2, \dots, s_N\}$ is defined, and for each state s_j there exists a traded security—an Arrow–Debreu security—that pays one unit of the numeraire if state s_j is realized and zero otherwise. When the number of linearly independent securities equals the number

of states, markets are said to be complete, and a unique state-price vector $\psi = (\psi_1, \dots, \psi_N)$ exists such that any contingent claim can be replicated by a portfolio of Arrow–Debreu securities.

Ross (1976) extended the completeness concept by showing that options and other derivative instruments can dynamically complete markets even when primitive securities do not span the full state space. The Ross spanning result implies that market completeness is a property of the entire set of traded instruments, not merely of individual securities. In practice, financial markets rarely achieve full completeness, and a large literature examines the implications of incomplete markets for asset pricing, risk sharing, and welfare (Cochrane, 2005).

The connection between complete markets and no-arbitrage pricing is formalized through the Fundamental Theorem of Asset Pricing. In a finite-state economy with no redundant securities, markets are complete if and only if there exists a unique positive state-price density (or stochastic discount factor) that prices all traded assets correctly. The existence of such a density is equivalent to the absence of arbitrage opportunities, while its uniqueness is equivalent to completeness (Cochrane, 2005). This theoretical structure provides the basis for our empirical tests: if prediction market contracts collectively span the outcome space and their prices satisfy the no-arbitrage conditions, then the market functions as an Arrow–Debreu economy in the precise sense of the theory.

2.2. Prediction Markets and Information Aggregation

The idea that markets can serve as forecasting devices predates modern financial economics, with roots in nineteenth-century betting markets on elections and horse races. The formal academic literature on prediction markets gained momentum with the establishment of the Iowa Electronic Markets (IEM) in 1988, which enabled researchers to study real-money markets on political events in a controlled setting.

Berg et al. (2008) documented the forecasting performance of the IEM over two decades of U.S. presidential elections, finding that market prices were generally well calibrated and often superior to polling averages, particularly at longer forecast horizons. Wolfers and Zitzewitz (2004) provided a comprehensive assessment of prediction markets across multiple domains—politics, business, entertainment—and argued that these markets aggregate information efficiently because they provide incentives for informed trading, allow for continuous updating, and admit a broad participant base. The authors acknowledged, however, that thin markets, manipulation attempts, and limited participation could undermine efficiency.

Manski (2006) raised an important theoretical caveat regarding the probability interpretation of prediction market prices. Under expected utility maximization, the equilibrium price of a state-contingent claim reflects both the probability of the event and the marginal utility of wealth in the relevant state. Only under risk-neutral pricing—when agents are risk neutral or when the event is uncorrelated with aggregate consumption—does the contract price equal the event probability. When

agents are risk averse and the event outcome is correlated with marginal utility (as is plausible for macroeconomic events such as Fed rate decisions), prices will diverge from probabilities. [Snowberg et al. \(2013\)](#) formalized this argument, deriving analytical expressions for the bias as a function of the coefficient of relative risk aversion and showing that the bias can be substantial for events with significant wealth consequences.

[Arrow et al. \(2008\)](#) provided a prominent defense of prediction markets, arguing that even if prices do not perfectly equal probabilities, the ordering of prices across contracts and their temporal dynamics still contain valuable information about relative likelihoods and the direction of probability changes. This pragmatic defense aligns with the Arrow–Debreu framework: state prices need not equal physical probabilities to be informative about the economic value of state-contingent claims.

More recent literature has explored the role of prediction markets in specific applied contexts. Studies have examined the ability of prediction markets to forecast election outcomes, central bank decisions, geopolitical events, and scientific replication results. A common finding across these applications is that prediction markets perform comparably to, and often better than, expert forecasts and statistical models, particularly when the forecast horizon is short and the event is well defined.

2.3. Price Discovery and Market Microstructure

The concept of price discovery refers to the process by which markets incorporate new information into prices. [Hasbrouck \(1995\)](#) formalized the measurement of price discovery using information shares, which decompose the variance of the efficient price innovation across multiple markets trading related assets. The Hasbrouck information share has become a standard tool for evaluating the relative informational contributions of different trading venues. [Hasbrouck \(2003\)](#) extended this framework to high-frequency settings, showing that the informational advantage of one market over another can vary intraday and across event types.

[Kyle \(1985\)](#) provided the seminal theoretical model of price formation with asymmetric information, in which an informed trader, noise traders, and a competitive market maker interact to determine prices. The Kyle model predicts that prices are a linear function of order flow, with the sensitivity (Kyle’s λ) reflecting the degree of information asymmetry. [Glosten and Milgrom \(1985\)](#) developed a related model in which the bid-ask spread compensates market makers for adverse selection costs arising from informed trading. These models provide the microstructural foundations for interpreting bid-ask spreads and price impact in prediction markets.

[O’Hara \(1995\)](#) synthesized the microstructure literature, emphasizing that the institutional details of trading mechanisms—order types, priority rules, transparency, anonymity—shape price formation in ways that matter for both efficiency and welfare. [Madhavan \(2000\)](#) extended this analysis, documenting how market design affects volatility, liquidity, and the speed of information incorporation.

The liquidity of a market is closely related to its price discovery function. [Kyle \(1985\)](#) identified three dimensions of liquidity—tightness (the cost of turning around a position), depth (the volume of orders at the best prices), and resiliency (the speed at which prices recover from uninformative shocks)—that collectively determine a market’s ability to absorb order flow without adverse price impact. [Amihud \(2002\)](#) proposed a simple illiquidity measure—the ratio of absolute daily returns to daily dollar volume—that captures the price impact of trading and has become widely used in the empirical literature. [Roll \(1984\)](#) derived an implicit measure of the effective spread from the serial covariance of price changes, showing that the autocovariance of returns is negative when bid-ask bounce contributes to price variation.

The application of microstructure concepts to prediction markets is relatively recent. Unlike traditional financial markets, blockchain-based prediction markets operate on automated market makers or order-book protocols without designated market makers or specialist systems. Polymarket uses a hybrid architecture that combines a central limit order book (CLOB) with smart contract settlement on the Polygon blockchain. This design implies that liquidity provision is decentralized and that the costs of adverse selection may differ qualitatively from those in traditional markets. Understanding the microstructure of these platforms is essential for evaluating their price discovery capabilities.

2.4. Volatility Modeling

The modeling of conditional volatility in financial time series has a rich econometric history. [Engle \(1982\)](#) introduced the ARCH model, in which conditional variance depends on past squared residuals, capturing the well-documented tendency of financial return volatility to cluster over time. [Bollerslev \(1986\)](#) generalized this framework to the GARCH model, which includes lagged conditional variance terms and provides a more parsimonious representation of persistent volatility dynamics. The GARCH(1,1) specification has become the workhorse model in empirical finance, applied to equity returns, exchange rates, interest rates, and commodity prices.

A well-established stylized fact of equity returns is the asymmetry of the volatility response to positive and negative shocks: negative returns tend to increase future volatility by more than positive returns of equal magnitude. [Nelson \(1991\)](#) proposed the EGARCH model to capture this leverage effect, while [Glosten et al. \(1993\)](#) introduced the GJR-GARCH specification, which augments the standard GARCH variance equation with an indicator function that allows the impact of negative shocks to differ from that of positive shocks. The GJR-GARCH model has the advantage of nesting the standard GARCH as a special case, facilitating direct comparison through information criteria such as the AIC ([Akaike, 1974](#)) and BIC ([Schwarz, 1978](#)).

The development of high-frequency data and realized volatility measures opened a new frontier in volatility modeling. [Andersen et al. \(2001\)](#) demonstrated that realized volatility—computed as

the sum of squared intraday returns—provides a consistent nonparametric estimator of integrated volatility and that the distribution of standardized returns conditional on realized volatility is approximately Gaussian. [Andersen et al. \(2003\)](#) showed that the logarithm of realized volatility is approximately normally distributed and exhibits long memory, with autocorrelations that decay slowly and remain significant at lags of months or longer. [Barndorff-Nielsen and Shephard \(2002\)](#) developed the econometric theory for realized volatility estimators, establishing their consistency and asymptotic normality under general conditions.

[Corsi \(2009\)](#) proposed the HAR-RV (Heterogeneous Autoregressive model of Realized Volatility) specification as a parsimonious approach to capturing the long-memory properties of realized volatility. The HAR-RV model regresses future realized volatility on the daily, weekly, and monthly components of past realized volatility, thereby approximating the long-memory behavior of the series through a cascade of different time horizons. Despite its simplicity—it is a restricted linear model—the HAR-RV achieves forecasting performance comparable to more complex long-memory models (ARFIMA, FIGARCH) and has become a standard benchmark in the volatility forecasting literature. Its application to prediction market returns, which we undertake in this paper, represents a novel contribution to the literature.

2.5. Blockchain Finance and Decentralized Markets

The emergence of blockchain-based financial infrastructure has spawned a rapidly growing academic literature. [Harvey et al. \(2021\)](#) provided an overview of decentralized finance (DeFi), documenting the institutional innovations—automated market makers, liquidity pools, smart contract settlement—that distinguish decentralized platforms from traditional financial intermediation. Blockchain-based prediction markets inherit many of these innovations: trades are executed on-chain, settlement is automatic and trustless, and the absence of centralized intermediaries reduces counterparty risk but introduces new challenges related to transaction costs (gas fees), latency, and regulatory uncertainty.

Polymarket, the platform studied in this paper, uses a hybrid architecture. Orders are matched off-chain on a central limit order book, while settlement occurs on the Polygon blockchain (a Layer 2 scaling solution for Ethereum). This design balances the speed and cost advantages of centralized order matching with the transparency and settlement finality of blockchain technology. The neg-risk framework employed by Polymarket for multi-outcome markets (such as the Fed rate decision market) ensures that the contracts within a given group are mutually exclusive, with the constraint that exactly one contract in the group resolves to \$1 and all others resolve to \$0. This structural feature directly maps onto the Arrow–Debreu framework and facilitates the completeness tests that we conduct.

2.6. Gaps in the Literature

Several gaps in the existing literature motivate our analysis. First, while prediction markets have been studied extensively from a forecasting and calibration perspective, formal tests of Arrow–Debreu completeness conditions have not been conducted using data from blockchain-based platforms. The neg-risk structure of Polymarket markets provides a uniquely clean setting for such tests. Second, the price discovery role of prediction markets relative to traditional financial markets has not been examined using modern time-series methods (Granger causality, VAR, impulse response analysis) applied to daily frequency data. Third, the volatility dynamics of prediction market returns have not been characterized using the GARCH-family and HAR-RV models that are standard in the financial econometrics literature. Our paper addresses each of these gaps.

3. Theoretical Framework

This section develops the theoretical foundations for interpreting prediction market contracts as Arrow–Debreu securities. We begin by formalizing the structure of the Arrow–Debreu economy, then map this structure onto the Polymarket institutional setting, and finally derive testable hypotheses regarding market completeness, no-arbitrage pricing, and information aggregation.

3.1. The Arrow–Debreu Economy

We consider a single-period economy with dates $t = 0$ (today) and $t = 1$ (the resolution date). Uncertainty at date 1 is represented by a finite state space $\mathcal{S} = \{s_1, s_2, \dots, s_N\}$ with $N \geq 2$. Each state s_j corresponds to a mutually exclusive and collectively exhaustive outcome.

Definition 1 (Arrow–Debreu Security). *An Arrow–Debreu security for state s_j is a financial contract that pays one unit of the numeraire at date $t = 1$ if state s_j is realized and zero otherwise. Formally, the payoff function of the j -th Arrow–Debreu security is*

$$a_j(s) = \mathbf{1}\{s = s_j\}, \quad j = 1, \dots, N,$$

where $\mathbf{1}\{\cdot\}$ denotes the indicator function.

Let $\psi_j \geq 0$ denote the date-0 price of the Arrow–Debreu security for state s_j . The vector $\psi = (\psi_1, \dots, \psi_N)' \in \mathbb{R}_+^N$ is called the state-price vector. In an economy with a risk-free asset that pays one unit in every state, the risk-free rate r_f is determined by the relation

$$\frac{1}{1 + r_f} = \sum_{j=1}^N \psi_j. \quad (1)$$

This equation establishes a fundamental link: the sum of all state prices equals the price of the risk-free bond, which is the reciprocal of the gross risk-free rate of return.

Definition 2 (Market Completeness). *A securities market is said to be complete if, for every contingent claim $c : \mathcal{S} \rightarrow \mathbb{R}$, there exists a portfolio of traded securities that replicates c in every state. Equivalently, markets are complete if and only if the payoff matrix of traded securities has rank N (equal to the number of states).*

When N Arrow–Debreu securities are available for trade with linearly independent payoffs, the payoff matrix is the $N \times N$ identity matrix, which trivially has full rank. In this case, completeness is achieved and any contingent claim $c = (c_1, \dots, c_N)'$ can be replicated at cost $\sum_{j=1}^N c_j \psi_j$.

3.2. State Prices, Risk-Neutral Probabilities, and No Arbitrage

The Fundamental Theorem of Asset Pricing establishes the equivalence between no arbitrage and the existence of positive state prices.

Proposition 1 (No Arbitrage and State Prices). *In the finite-state economy described above, the following conditions are equivalent:*

- (i) *There are no arbitrage opportunities.*
- (ii) *There exists a strictly positive state-price vector $\psi \in \mathbb{R}_{++}^N$ such that the price of any traded security equals the inner product of its payoff vector with ψ .*
- (iii) *There exists a probability measure $\mathbb{Q}$ equivalent to the physical measure $\mathbb{P}$ and a discount factor $\delta > 0$ such that the price of any traded security equals the discounted $\mathbb{Q}$ -expectation of its payoff.*

The measure $\mathbb{Q}$ in condition (iii) is the risk-neutral probability measure, and its elements are related to state prices by

$$q_j = \frac{\psi_j}{\sum_{k=1}^N \psi_k} = (1 + r_f) \psi_j, \quad j = 1, \dots, N. \quad (2)$$

Under the risk-neutral measure, the price of an Arrow–Debreu security is simply the discounted risk-neutral probability of the associated state:

$$\psi_j = \frac{q_j}{1 + r_f}.$$

In a complete market, the risk-neutral measure $\mathbb{Q}$ is unique. In an incomplete market, no-arbitrage is consistent with multiple risk-neutral measures, and the state-price vector is not uniquely determined by observed prices alone. This distinction motivates our focus on markets with a complete set of outcome contracts.

3.3. Mapping Polymarket Contracts to Arrow–Debreu Securities

We now map the theoretical framework onto the institutional structure of Polymarket. Consider a Polymarket neg-risk market with N outcome contracts $\{C_1, C_2, \dots, C_N\}$. Each contract C_j is a binary claim that pays \$1 if outcome j is realized and \$0 otherwise. The outcomes are mutually exclusive and collectively exhaustive by construction: the Polymarket neg-risk protocol enforces the constraint that exactly one contract in the group resolves to \$1.

Let $p_j(t)$ denote the market price of contract C_j at time t . Under the Arrow–Debreu interpretation, $p_j(t) = \psi_j(t)$, the state price for outcome j at time t .

Definition 3 (State Price Sum). *The state price sum at time t is defined as*

$$\Sigma(t) = \sum_{j=1}^N p_j(t). \quad (3)$$

Under Arrow–Debreu completeness with a risk-free rate $r_f(t)$, the theoretical prediction is $\Sigma(t) = 1/(1 + r_f(t))$.

In practice, the overnight risk-free rate over the sample period is sufficiently small that $1/(1 + r_f) \approx 1$. Deviations of $\Sigma(t)$ from unity therefore measure violations of the completeness condition, up to a small correction for the time value of money.

Several sources of deviation merit discussion. Transaction costs, including the Polymarket trading fee and blockchain gas costs, create a wedge between theoretical and observed state prices. If buyers must pay a premium above fair value and sellers must accept a discount below fair value, the observed state price sum will exceed unity by an amount proportional to the round-trip transaction cost. Similarly, the vigorish (or overround) familiar from bookmaking—in which the implied probabilities of all outcomes sum to more than one—represents a form of transaction cost that accrues to the market operator or liquidity providers.

3.4. Risk-Neutral Density Estimation

Given a complete set of state prices, the risk-neutral density over the outcome space can be recovered by normalizing the state prices to sum to unity. This yields the risk-neutral probability mass function:

$$q_j(t) = \frac{p_j(t)}{\Sigma(t)}, \quad j = 1, \dots, N. \quad (4)$$

The time series of risk-neutral densities $\{q_j(t)\}_{t=1}^T$ provides a window into the evolution of market expectations about the distribution of outcomes. Changes in the risk-neutral density can

be decomposed into changes in physical probabilities and changes in risk premia, though this decomposition requires additional identifying assumptions.

For the Federal Reserve rate decision market with $N = 5$ states (50+ basis point decrease, 25 basis point decrease, no change, 25 basis point increase, 50+ basis point increase), the risk-neutral density at each date is a discrete probability distribution over the five possible rate actions. The evolution of this distribution over time reflects the market's assessment of monetary policy expectations as macroeconomic data are released and FOMC communications are made.

3.5. Returns, Abnormal Returns, and Event Studies

We define log returns on prediction market contracts as follows. Let $p_j(t)$ denote the price of contract j at the close of trading day t . The continuously compounded return is

$$r_j(t) = \ln \left(\frac{p_j(t)}{p_j(t-1)} \right). \quad (5)$$

For the event study analysis, we follow the standard methodology of [MacKinlay \(1997\)](#). Abnormal returns are defined relative to a benchmark model estimated over a pre-event window. Let $[\tau_0, \tau_1]$ denote the estimation window and $[\tau_1 + 1, \tau_2]$ denote the event window. We employ a constant mean return model as the benchmark:

$$AR_j(t) = r_j(t) - \hat{\mu}_j, \quad t \in [\tau_1 + 1, \tau_2], \quad (6)$$

where $\hat{\mu}_j = \frac{1}{\tau_1 - \tau_0} \sum_{t=\tau_0+1}^{\tau_1} r_j(t)$ is the sample mean return over the estimation window. Cumulative abnormal returns over the event window $[t_1, t_2]$ are

$$CAR_j(t_1, t_2) = \sum_{t=t_1}^{t_2} AR_j(t). \quad (7)$$

The cross-sectional test statistic for the null hypothesis that the mean CAR across events is zero follows [Patell \(1976\)](#) and [Boehmer et al. \(1991\)](#):

$$t_{CAR} = \frac{\overline{CAR}}{\hat{\sigma}_{CAR}/\sqrt{K}}, \quad (8)$$

where $\overline{CAR}$ is the mean CAR across K events and $\hat{\sigma}_{CAR}$ is the cross-sectional standard deviation.

3.6. Testable Hypotheses

The theoretical framework yields the following testable hypotheses.

Hypothesis 1 (Market Completeness). *If Polymarket neg-risk contract markets function as Arrow–Debreu economies, then:*

- (a) *The state price sum $\Sigma(t) = \sum_{j=1}^N p_j(t)$ is statistically indistinguishable from $1/(1 + r_f)$ for all t .*
- (b) *All individual state prices $p_j(t)$ are strictly positive for all j and t .*
- (c) *The deviation of $\Sigma(t)$ from unity is small, stationary, and attributable to transaction costs.*

Hypothesis 2 (No Arbitrage). *If no-arbitrage conditions hold in the prediction market:*

- (a) *There are no persistent violations of the law of one price: the state price sum does not systematically exceed or fall below unity by more than transaction costs.*
- (b) *Temporary deviations from the no-arbitrage bounds are quickly corrected through arbitrage activity, implying mean-reversion in $\Sigma(t) - 1$.*

Hypothesis 3 (Information Aggregation). *If prediction market prices efficiently aggregate information:*

- (a) *Prediction market returns exhibit no predictable patterns conditional on publicly available information (weak-form efficiency).*
- (b) *Prediction market prices respond to macroeconomic news releases (FOMC announcements, CPI releases) in a manner consistent with the informational content of the release.*
- (c) *Prediction market returns Granger-cause returns in related traditional financial markets (SPY, VIX, BTC), reflecting the information content of state-contingent prices.*

Hypothesis 4 (Volatility Dynamics). *If prediction market returns exhibit volatility dynamics similar to those documented in traditional financial markets:*

- (a) *Conditional volatility is time-varying and exhibits clustering, consistent with ARCH/GARCH effects.*
- (b) *Negative return shocks have a larger impact on future volatility than positive return shocks of equal magnitude (leverage effect), implying that the GJR-GARCH model fits better than the symmetric GARCH.*
- (c) *Realized volatility exhibits persistence at multiple time horizons, consistent with the heterogeneous market hypothesis underlying the HAR-RV model.*

3.7. Theoretical Limitations and Caveats

Several theoretical limitations of the Arrow–Debreu framework as applied to prediction markets warrant acknowledgment. First, the single-period nature of the Arrow–Debreu model abstracts from the dynamic trading that occurs in practice. Prediction market contracts are traded continuously over weeks or months before resolution, and the mapping between theory and empirics requires

treating each trading day as a snapshot of the equilibrium state-price vector. Second, the finite-state assumption, while natural for markets with discrete outcomes (five possible Fed rate actions), becomes strained for markets with continuous outcome spaces (e.g., the level of GDP growth). Third, the assumption of competitive price-taking may be violated in prediction markets where a small number of large traders can exert significant price impact. Fourth, the theoretical framework assumes that all agents agree on the state space; in practice, ambiguity about the precise definition of contract resolution rules can introduce basis risk between the prediction market contract and the underlying event of interest.

Despite these limitations, the Arrow–Debreu framework provides a disciplined lens through which to evaluate the structural properties of prediction markets. The hypotheses formulated above are designed to be robust to these concerns: we test for the conditions that must hold if the market functions as an Arrow–Debreu economy, while acknowledging that departures from these conditions may reflect institutional frictions rather than fundamental inefficiency.

4. Data

This section describes the data sources, sample construction, variable definitions, and preliminary statistical properties of the prediction market and financial market data used in our analysis.

4.1. Polymarket Data

Polymarket is a blockchain-based prediction market platform that enables users to trade binary outcome contracts settled on the Polygon network. Each contract pays \$1 if the specified outcome occurs and \$0 otherwise, with prices reflecting the market’s implied probability of the event. We collect daily closing prices, trading volumes, and liquidity measures for two sets of Polymarket contracts.

4.1.1. Federal Reserve Rate Decision Market

Our primary dataset consists of daily observations on five contracts corresponding to the possible outcomes of the Federal Reserve’s June 2026 FOMC rate decision. The five mutually exclusive and collectively exhaustive states are:

1. **50+ bps decrease:** The FOMC lowers the target federal funds rate by 50 basis points or more.
2. **25 bps decrease:** The FOMC lowers the target federal funds rate by exactly 25 basis points.
3. **No change:** The FOMC leaves the target federal funds rate unchanged.
4. **25 bps increase:** The FOMC raises the target federal funds rate by exactly 25 basis points.
5. **50+ bps increase:** The FOMC raises the target federal funds rate by 50 basis points or more.

The sample spans 146 daily observations from December 2025 through May 2026. These contracts are traded within Polymarket’s neg-risk framework, which enforces mutual exclusivity and collective exhaustiveness: exactly one contract in the group resolves to \$1, and the remaining four resolve to \$0. This structural feature makes the Fed rate decision market an ideal testing ground for Arrow–Debreu completeness conditions.

Table 1 reports summary statistics for the five contracts. The “No change” contract has the highest mean price at 0.608, reflecting the market’s prevailing expectation of policy inaction. The “25 bps decrease” contract has a mean price of 0.304, indicating meaningful probability mass on a single rate cut. The two extreme states (50+ bps decrease and 50+ bps increase) have mean prices of 0.058 and 0.016 respectively, consistent with the market assigning low probability to aggressive rate moves. All five price series exhibit non-normal distributions, as evidenced by significant Jarque–Bera statistics: the “25 bps increase” contract displays extreme positive skewness (3.98) and leptokurtosis (30.46), reflecting its status as a low-probability, high-surprise event.

Table 1. Summary Statistics of Prediction Market Prices

Market	Mean	Std	Min	Max	Skew	Kurt	JB
50+ bps decrease	0.058	0.042	0.004	0.150	−0.03	−1.45	12.7
25 bps decrease	0.304	0.194	0.006	0.570	−0.47	−1.46	18.0
No change	0.608	0.235	0.300	0.981	0.55	−1.41	19.1
25 bps increase	0.034	0.026	0.005	0.250	3.98	30.46	5628
50+ bps increase	0.016	0.016	0.002	0.095	1.66	3.55	1358

Notes: JB is the Jarque–Bera test statistic for normality. ***, **, and * denote significance at the 1%, 5%, and 10% levels.

4.1.2. 2024 U.S. Presidential Election Market

Our secondary dataset comprises daily observations on 17 candidate contracts from the 2024 U.S. Presidential Election market, observed over 307 trading days. This market provides a richer cross-sectional dimension for analyzing microstructure properties, liquidity dynamics, and multi-contract pricing relationships. The election market resolved on November 5, 2024, but Polymarket continued trading after the election date, providing observations on post-resolution price convergence.

Donald Trump won the election, and his contract converged to a final price of 0.58—notably below the \$1 resolution value, indicating that some trading continued in the secondary market before final settlement. The election market’s larger cross-section and longer time series complement the Fed rate decision market’s cleaner state-space structure.

4.2. Financial Market Data

To assess the information linkages between prediction markets and traditional financial markets, we collect daily returns data for three benchmark assets:

- **SPY**: The SPDR S&P 500 ETF Trust, representing the U.S. equity market. Daily closing prices are obtained from standard financial data providers.
- **BTC**: Bitcoin, the largest cryptocurrency by market capitalization. Bitcoin is included both as a benchmark for blockchain-native assets and as a proxy for speculative risk appetite in the digital asset ecosystem.
- **VIX**: The CBOE Volatility Index, which measures the market’s expectation of 30-day forward-looking volatility implied by S&P 500 index options. The VIX is included as a proxy for aggregate uncertainty and risk aversion.

Daily returns are computed as log returns: $r_t = \ln(P_t/P_{t-1})$, where P_t is the closing price on day t . For the prediction market contracts, the same return definition is applied, with the convention that returns are undefined (and excluded) on days when the previous day’s price is zero.

4.3. Event Data

We identify macroeconomic announcement events within the sample period for the event study analysis. The events fall into three categories:

- **FOMC announcements**: We identify 2 scheduled FOMC meeting dates that fall within the sample period. FOMC announcements are released at 2:00 PM Eastern Time and are accompanied by the statement, Summary of Economic Projections (when available), and the press conference.
- **CPI releases**: We identify 3 Consumer Price Index releases by the Bureau of Labor Statistics within the sample period. CPI releases occur at 8:30 AM Eastern Time and provide the most closely watched measure of consumer inflation.
- **Election key dates**: We identify 3 politically significant dates in the election sample, including debate dates, major endorsement events, and the election day itself.

Event windows are defined symmetrically around each announcement date. For FOMC and CPI events, we use a $[-3, +3]$ day window (7 trading days), with the estimation window for the constant mean model spanning 60 trading days prior to the event window. For election events, we use a wider $[-5, +5]$ day window to accommodate the possibility of slower information incorporation in political markets.

4.4. Variable Construction

We construct the following variables for our empirical analysis:

State price sum. For each trading day t , the state price sum is computed as $\Sigma(t) = \sum_{j=1}^5 p_j(t)$, where $p_j(t)$ is the closing price of the j -th contract. Under Arrow–Debreu completeness, $\Sigma(t) \approx 1$.

Risk-neutral probabilities. The risk-neutral probability of state j at time t is computed as $q_j(t) = p_j(t)/\Sigma(t)$. This normalization ensures that risk-neutral probabilities sum to unity.

Abnormal returns. For the event study analysis, abnormal returns are computed as $AR_j(t) = r_j(t) - \hat{\mu}_j$, where $\hat{\mu}_j$ is the mean return over the 60-day estimation window. Cumulative abnormal returns are computed by summing abnormal returns over the event window.

Realized volatility. Daily realized volatility for the “No change” contract is computed as the absolute value of the daily return, $|r_t|$. Weekly and monthly realized volatility components are computed as the averages of daily realized volatility over the preceding 5 and 22 trading days, respectively, following the HAR-RV convention of [Corsi \(2009\)](#).

Liquidity measures. We compute two liquidity measures. The Roll spread ([Roll, 1984](#)) is estimated from the serial autocovariance of returns: $\text{Roll}_t = 2\sqrt{-\text{Cov}(r_t, r_{t-1})}$ when the autocovariance is negative, and zero otherwise. The Amihud illiquidity ratio ([Amihud, 2002](#)) is computed as $\text{ILLIQ}_t = |r_t|/V_t$, where V_t is the dollar trading volume on day t .

4.5. Descriptive Statistics

Table 2 reports descriptive statistics for the broader set of Polymarket contracts examined, including trading volume and liquidity measures in U.S. dollars. The table documents the substantial heterogeneity in market activity across contracts, with trading volumes ranging from under one million to nearly ten million dollars and liquidity (measured as the depth of the order book) varying by more than an order of magnitude.

Figure 1 plots the evolution of state prices for the five Fed rate decision contracts over the sample period. The dominant feature is the secular increase in the “No change” contract price from approximately 0.575 at the beginning of the sample to 0.978 by May 2026, accompanied by a corresponding decline in the “25 bps decrease” contract from approximately 0.40 to near zero. This pattern is consistent with the market progressively assigning higher probability to policy inaction as incoming macroeconomic data (strong labor markets, persistent inflation) reduced the case for near-term rate cuts.

Figure 2 displays snapshots of the risk-neutral density at selected dates, illustrating the progressive concentration of probability mass on the “No change” outcome. Early in the sample, the distribution is bimodal, with substantial mass on both “No change” and “25 bps decrease”; by the end of the sample, the distribution is sharply peaked at “No change,” with all other states assigned negligible probability.

Table 2. Descriptive Statistics of Polymarket Contracts

Market	Volume (\$)	Liquidity (\$)	Obs	Mean Price	Std Dev
Will Marco Rubio win the 2028 US Preside	9.962,958	160002	0		
Will Anže Logar be the next Prime Minist	995465	11.011	0		
Will Stephen Cloobek win the California	995372	184072	0		
Will the Republicans win the 2028 US Pre	994905	331418	0		
Will Ethereum dip to \$1,500 by December	994496	43.390	0		
Will Ahmed al-Sharaa win the Nobel Peace	993894	42.726	0		
US strike on Colombia by December 31?	991582	6.051	0		
Over \$3M committed to the Printr public	990487	25.664	0		
Will Lance Stroll be the 2026 F1 Drivers	9.845,762	836174	0		
Will Michael Younger win the California	977490	193379	0		
Will Anthropic have the best AI model at	975748	28.076	0		
Will the Los Angeles Angels win the 2026	973951	95.205	0		
Will PSG win the 2025–26 Champions Leagu	9.718,043	670357	0		
Will Kamala Harris win the California Go	970214	82.083	0		
Will Lovable be acquired before 2027?	969501	3.864	0		
Will Gretchen Whitmer win the 2028 Democ	9.646,053	576258	0		
Will MetaMask launch a token by Septembe	964498	4.197	0		
New "Stranger Things" episode released b	963189	36.759	0		
Will Dawn Buckingham win the 2026 Texas	961827	26.871	0		
Will no qualifying diplomatic US-Iran me	958458	51.713	0		

Notes: This table reports summary statistics for the Polymarket contracts analyzed. Volume and liquidity are reported in US dollars. Mean price represents the average implied probability.

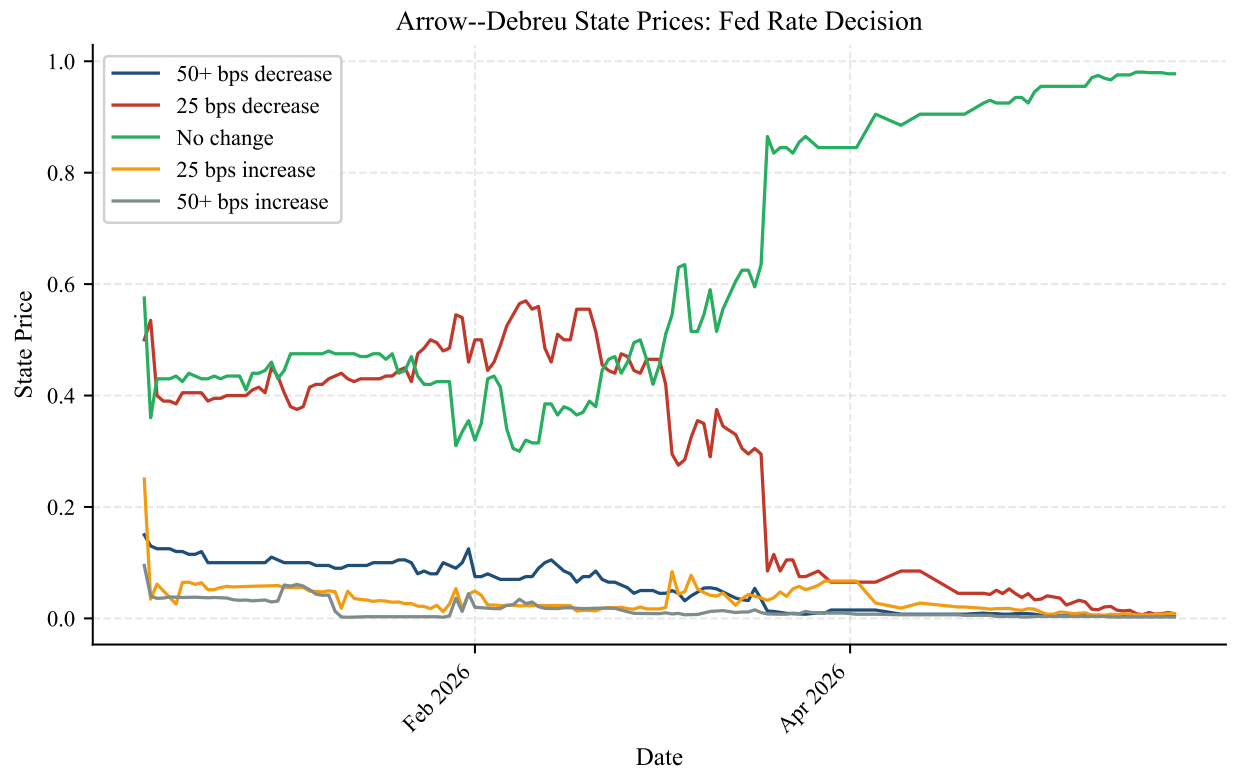


Figure 1. Arrow–Debreu state prices for Federal Reserve rate decision contracts, December 2025 to May 2026. Each line represents the daily closing price of one of the five mutually exclusive outcome contracts. The “No change” contract (upper line) rises from approximately 0.575 to 0.978, while the “25 bps decrease” contract (declining line) falls from approximately 0.40 to near zero.

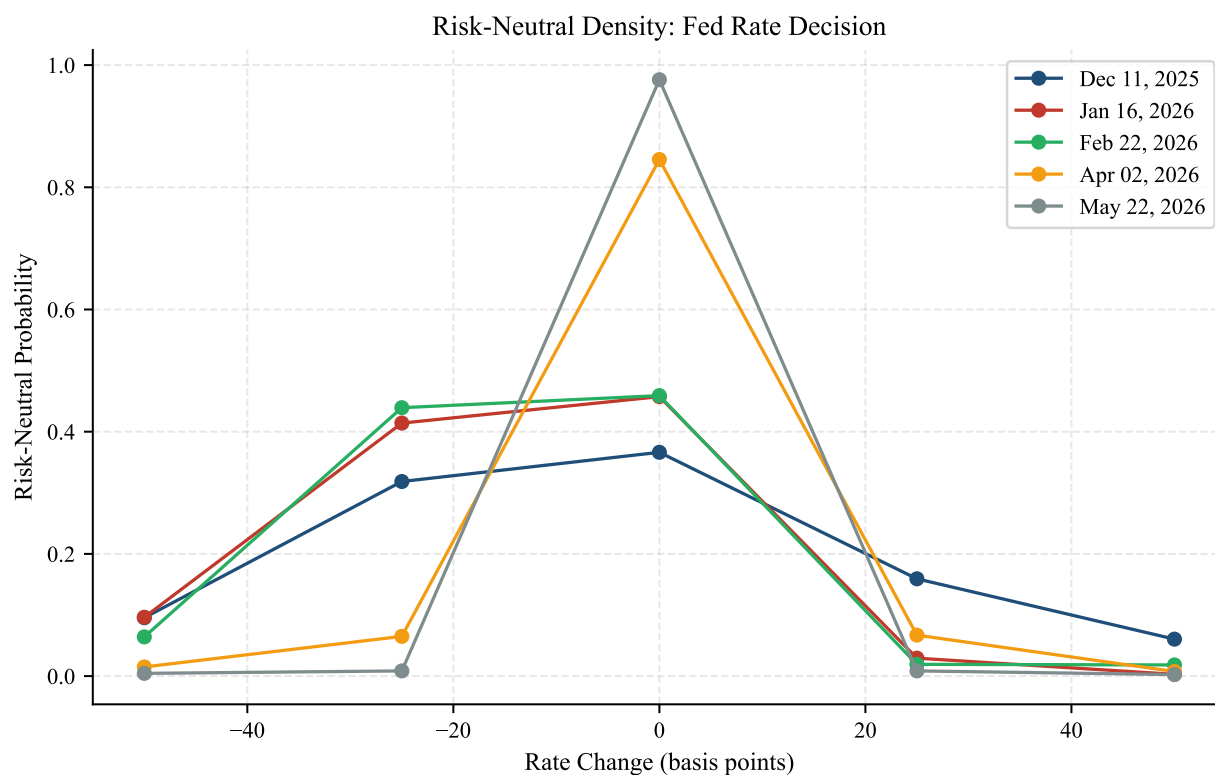


Figure 2. Risk-neutral density snapshots for the Federal Reserve rate decision market. Each panel shows the distribution of risk-neutral probabilities across the five outcome states at a selected date. The progressive concentration of mass on “No change” reflects the market’s increasing confidence in policy inaction.

4.6. Stationarity and Distributional Properties

Since prediction market prices are bounded between 0 and 1 and converge to either 0 or 1 at resolution, the level series are inherently non-stationary. All time-series analyses in this paper therefore use log returns rather than price levels. Augmented Dickey–Fuller tests reject the unit root null for all return series at the 1% significance level, confirming stationarity.

The return distributions exhibit significant departures from normality, as indicated by the Jarque–Bera statistics reported in Table 1. The “25 bps increase” contract displays particularly extreme non-normality, with a kurtosis of 30.46 reflecting occasional large price jumps associated with unexpected news. These distributional properties motivate the use of heteroskedasticity-robust inference methods throughout our analysis, including Newey–West standard errors (Newey and West, 1987) and White standard errors (White, 1980).

5. Methodology

This section presents the econometric methods employed to test the hypotheses developed in Section 3. We describe five methodological components: (i) Arrow–Debreu completeness tests, (ii) event study methodology, (iii) Granger causality and vector autoregression, (iv) GARCH-family volatility models, and (v) the HAR-RV model for realized volatility. For each method, we specify the estimating equations, identification strategy, and inference procedures.

5.1. Arrow–Debreu Completeness Tests

To test Hypothesis 1, we construct the daily state price sum $\Sigma(t) = \sum_{j=1}^5 p_j(t)$ for the Federal Reserve rate decision market and evaluate its properties against the theoretical predictions of the Arrow–Debreu framework.

5.1.1. Mean Deviation Test

The primary test statistic is the time-series average of $\Sigma(t)$:

$$\bar{\Sigma} = \frac{1}{T} \sum_{t=1}^T \Sigma(t), \quad (9)$$

where $T = 146$ is the number of trading days. Under the null hypothesis that the market is Arrow–Debreu complete with a risk-free rate close to zero, $\bar{\Sigma} = 1$. We test $H_0 : \mathbb{E}[\Sigma(t)] = 1$ against $H_1 : \mathbb{E}[\Sigma(t)] \neq 1$ using a t -test with Newey–West standard errors (Newey and West, 1987) to account for potential serial correlation in $\Sigma(t)$.

5.1.2. Positivity Test

Hypothesis 1(b) requires that all state prices remain strictly positive. We verify this condition by checking whether $\min_{j,t} p_j(t) > 0$ across all five contracts and all trading days.

5.1.3. Stationarity of Deviations

To assess whether deviations of $\Sigma(t)$ from unity are temporary (consistent with transaction-cost frictions) or persistent (suggesting structural mispricing), we apply the Augmented Dickey–Fuller test to the series $\delta(t) = \Sigma(t) - 1$, testing the null of a unit root against the alternative of stationarity. Rejection of the unit root null would indicate that arbitrage forces push the state price sum back toward unity, consistent with Hypothesis 2(b).

5.2. Event Study Methodology

We employ the standard event study methodology of MacKinlay (1997) to assess the impact of macroeconomic announcements on prediction market prices. The procedure comprises four steps.

5.2.1. Step 1: Estimation Window

For each event k with announcement date τ_k , we define the estimation window as the 60 trading days ending 4 days before the event: $[\tau_k - 63, \tau_k - 4]$. The constant mean return model is estimated over this window:

$$\hat{\mu}_j^{(k)} = \frac{1}{60} \sum_{t=\tau_k-63}^{\tau_k-4} r_j(t). \quad (10)$$

We use a 60-day estimation window as a compromise between statistical precision (which favors longer windows) and parameter stability (which favors shorter windows). The 3-day gap between the estimation window and the event window prevents contamination of the benchmark model by pre-announcement information leakage.

5.2.2. Step 2: Abnormal Returns

Abnormal returns for contract j around event k are computed as

$$AR_j^{(k)}(t) = r_j(t) - \hat{\mu}_j^{(k)}, \quad t \in [\tau_k - w, \tau_k + w], \quad (11)$$

where w is the half-width of the event window (3 days for FOMC and CPI events, 5 days for election events).

5.2.3. Step 3: Cumulative Abnormal Returns

Cumulative abnormal returns over the event window are

$$CAR_j^{(k)} = \sum_{t=\tau_k-w}^{\tau_k+w} AR_j^{(k)}(t). \quad (12)$$

5.2.4. Step 4: Cross-Sectional Testing

The null hypothesis is that the mean CAR across events is zero: $H_0 : \mathbb{E}[\overline{CAR}_j] = 0$. The cross-sectional test statistic is

$$t_j = \frac{\overline{CAR}_j}{s_j/\sqrt{K}}, \quad (13)$$

where $\overline{CAR}_j = K^{-1} \sum_{k=1}^K CAR_j^{(k)}$, s_j is the cross-event standard deviation of CARs, and K is the number of events. Under the null hypothesis and standard regularity conditions, t_j follows a t -distribution with $K - 1$ degrees of freedom (Boehmer et al., 1991).

The event study methodology relies on two identifying assumptions. First, the constant mean model must be a reasonable benchmark for expected returns in the absence of the event. For prediction market contracts, which have no systematic risk exposure in the CAPM sense, the constant mean model is more appropriate than market model or factor model alternatives. Second, events must be exogenous to the prediction market—that is, the FOMC and CPI announcement dates must not be determined by prediction market prices. This assumption is clearly satisfied, as these dates are set on predetermined schedules.

5.3. Granger Causality and Vector Autoregression

To assess information linkages between prediction markets and traditional financial markets (Hypothesis 3(c)), we employ pairwise Granger causality tests (Granger, 1969) and estimate a vector autoregressive (VAR) model (Sims, 1980).

5.3.1. Pairwise Granger Causality

Consider two stationary time series x_t and y_t . We say that x Granger-causes y if lagged values of x contain information useful for predicting y beyond the information contained in lagged values of y alone. Formally, we estimate the unrestricted model

$$y_t = \alpha + \sum_{\ell=1}^L \beta_{\ell} y_{t-\ell} + \sum_{\ell=1}^L \gamma_{\ell} x_{t-\ell} + \varepsilon_t, \quad (14)$$

and the restricted model

$$y_t = \alpha + \sum_{\ell=1}^L \beta_{\ell} y_{t-\ell} + u_t, \quad (15)$$

where L is the lag length selected by the Akaike Information Criterion (Akaike, 1974). The null hypothesis $H_0 : \gamma_1 = \gamma_2 = \dots = \gamma_L = 0$ is tested using an F -test. Rejection of H_0 implies that x Granger-causes y .

We conduct pairwise Granger causality tests among four return series: the Fed “No Change” contract, SPY, BTC, and VIX. This yields 12 directional tests (four variables, each pair in both directions), reported in Table 5.

5.3.2. Vector Autoregression

To capture the joint dynamics of prediction market returns and traditional asset returns, we estimate a VAR model of order L :

$$\mathbf{y}_t = \boldsymbol{\mu} + \sum_{\ell=1}^L \mathbf{A}_{\ell} \mathbf{y}_{t-\ell} + \boldsymbol{\varepsilon}_t, \quad (16)$$

where $\mathbf{y}_t = (r_t^{\text{Fed}}, r_t^{\text{SPY}}, r_t^{\text{BTC}}, r_t^{\text{VIX}})'$ is the 4×1 vector of returns, $\boldsymbol{\mu}$ is a 4×1 vector of intercepts, $\mathbf{A}_{\ell}$ are 4×4 coefficient matrices, and $\boldsymbol{\varepsilon}_t$ is a 4×1 vector of innovations with $\mathbb{E}[\boldsymbol{\varepsilon}_t] = \mathbf{0}$ and $\mathbb{E}[\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t'] = \boldsymbol{\Omega}$. The lag length L is selected by minimizing the AIC over candidate lag orders $L \in \{1, 2, \dots, 8\}$.

The VAR model allows us to trace the dynamic propagation of shocks across markets through impulse response functions and to decompose forecast error variance into contributions from each variable through variance decomposition analysis. These tools complement the pairwise Granger causality tests by providing a system-level perspective on information transmission.

5.4. GARCH-Family Volatility Models

To characterize the conditional volatility dynamics of prediction market returns (Hypothesis 4), we estimate two members of the GARCH family: the standard GARCH(1,1) model and the GJR-GARCH(1,1) model.

5.4.1. GARCH(1,1)

The GARCH(1,1) model of Bollerslev (1986) specifies the conditional variance of returns as

$$r_t = \mu + \varepsilon_t, \quad \varepsilon_t = \sigma_t z_t, \quad z_t \sim \mathcal{N}(0, 1), \quad (17)$$

$$\sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \quad (18)$$

where $\omega > 0$, $\alpha_1 \geq 0$, $\beta_1 \geq 0$, and the stationarity condition $\alpha_1 + \beta_1 < 1$ ensures that the unconditional variance is finite. The parameter α_1 measures the impact of yesterday's squared shock on today's conditional variance (the ARCH effect), while β_1 measures the persistence of past conditional variance (the GARCH effect). The sum $\alpha_1 + \beta_1$ measures the overall persistence of volatility shocks: values close to unity indicate near-integrated volatility.

5.4.2. GJR-GARCH(1,1)

The GJR-GARCH(1,1) model of [Glosten et al. \(1993\)](#) augments the GARCH variance equation with an asymmetric leverage term:

$$\sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \gamma_1 \varepsilon_{t-1}^2 \mathbf{1}\{\varepsilon_{t-1} < 0\} + \beta_1 \sigma_{t-1}^2, \quad (19)$$

where $\gamma_1 \geq 0$ is the leverage coefficient. When $\gamma_1 > 0$, negative shocks ($\varepsilon_{t-1} < 0$) increase the conditional variance by $\alpha_1 + \gamma_1$, while positive shocks increase it by only α_1 . The standard GARCH model is nested as the special case $\gamma_1 = 0$.

Model comparison between the GARCH(1,1) and GJR-GARCH(1,1) is based on the Akaike Information Criterion:

$$\text{AIC} = -2 \ln \mathcal{L} + 2k, \quad (20)$$

where $\mathcal{L}$ is the maximized log-likelihood and k is the number of estimated parameters. The model with the lower AIC is preferred. Since the GJR-GARCH has one additional parameter (γ_1) relative to the standard GARCH, a reduction in AIC of more than 2 constitutes evidence in favor of the asymmetric specification.

Both models are estimated by maximum likelihood under the assumption of conditionally Gaussian innovations. To ensure robustness, we also compute quasi-maximum likelihood (QML) standard errors ([White, 1980](#)), which remain consistent even if the true conditional distribution departs from Gaussianity.

5.5. HAR-RV Model

To model the dynamics of realized volatility and test Hypothesis 4(c), we employ the Heterogeneous Autoregressive model of Realized Volatility (HAR-RV) proposed by [Corsi \(2009\)](#). The HAR-RV captures long-memory properties of realized volatility through a cascade of volatility components at daily, weekly, and monthly horizons.

5.5.1. Realized Volatility Construction

In the absence of intraday data, we proxy daily realized volatility by the absolute daily return on the Fed “No Change” contract:

$$RV_t^{(d)} = |r_t|. \quad (21)$$

Weekly and monthly realized volatility components are constructed as arithmetic averages of daily realized volatility:

$$RV_t^{(w)} = \frac{1}{5} \sum_{i=0}^4 RV_{t-i}^{(d)}, \quad (22)$$

$$RV_t^{(m)} = \frac{1}{22} \sum_{i=0}^{21} RV_{t-i}^{(d)}. \quad (23)$$

This construction assumes that absolute returns are a reasonable proxy for integrated volatility at the daily frequency. While the theoretical foundations of the HAR-RV model are based on high-frequency realized volatility estimators ([Andersen et al., 2001](#); [Barndorff-Nielsen and Shephard, 2002](#)), the model has been shown to perform well with daily-frequency inputs when intraday data are unavailable.

5.5.2. HAR-RV Specification

The HAR-RV model regresses future daily realized volatility on the three temporal components:

$$RV_{t+1}^{(d)} = c + \beta_d RV_t^{(d)} + \beta_w RV_t^{(w)} + \beta_m RV_t^{(m)} + \varepsilon_{t+1}. \quad (24)$$

The coefficient β_d captures the impact of short-run (daily) volatility fluctuations on next-period volatility, β_w captures the medium-run (weekly) contribution, and β_m captures the long-run (monthly) contribution. In traditional equity markets, all three coefficients are typically positive and statistically significant, with β_w and β_m providing substantial incremental explanatory power—a pattern attributed to the heterogeneous investment horizons of different market participant types ([Corsi, 2009](#)).

The model is estimated by OLS with Newey–West standard errors ([Newey and West, 1987](#)) using 5 lags to account for serial correlation in the residuals, which arises from the overlapping nature of the weekly and monthly realized volatility components. Goodness-of-fit is assessed by R^2 and adjusted R^2 .

5.6. Liquidity Analysis

We complement the time-series analyses with an assessment of market microstructure quality using two established liquidity measures.

5.6.1. Roll Effective Spread

The Roll implicit spread ([Roll, 1984](#)) is derived from the first-order autocovariance of returns under the assumption that observed price changes reflect the sum of an efficient price innovation and bid-ask bounce:

$$\text{Roll} = 2\sqrt{-\text{Cov}(r_t, r_{t-1})}, \quad (25)$$

when $\text{Cov}(r_t, r_{t-1}) < 0$. Positive autocovariance, which can arise from momentum effects, yields an undefined Roll spread; in such cases, we set the Roll spread to zero. The Roll spread is interpreted as an estimate of the effective bid-ask spread and serves as a measure of trading costs.

5.6.2. Amihud Illiquidity Ratio

The Amihud illiquidity ratio ([Amihud, 2002](#)) measures the price impact of trading:

$$\text{ILLIQ}_t = \frac{|r_t|}{V_t}, \quad (26)$$

where V_t is the dollar volume on day t . Higher values of ILLIQ indicate that a given volume of trading moves prices by a larger amount, implying lower liquidity. The Amihud ratio is computed on a daily basis and then averaged over rolling windows to produce a time series of liquidity conditions.

5.7. Inference and Standard Errors

Throughout the analysis, we employ robust inference procedures. For OLS regressions (HAR-RV), we report Newey–West standard errors with lag truncation determined by the sample size ([Newey and West, 1987](#)). For GARCH models, quasi-maximum likelihood standard errors are used. For event studies, cross-sectional standard errors account for cross-event variation. All hypothesis tests are two-sided unless otherwise noted.

For the Granger causality tests, we report both F -statistics and p -values, with significance assessed at the 1%, 5%, and 10% levels. Lag lengths are selected by the Akaike Information Criterion to balance model fit and parsimony. For the VAR model, we verify the stability condition (all eigenvalues of the companion matrix lie inside the unit circle) to ensure that impulse response functions and variance decompositions are well defined.

6. Results

This section presents the empirical findings organized around the four hypotheses developed in Section 3. We begin with the Arrow–Debreu completeness tests, proceed to the event study and information linkage analyses, and conclude with the volatility dynamics results.

6.1. Arrow–Debreu Completeness

Table 3 reports the results of the Arrow–Debreu completeness tests for the Federal Reserve rate decision market. The mean state price sum across the 146 daily observations is 1.0205, with a standard deviation of 0.0527. The mean deviation from unity is 0.0248, and the maximum absolute deviation is 0.5700. All individual state prices remain strictly positive throughout the sample, confirming that the non-negativity condition required by the Arrow–Debreu framework is satisfied.

Table 3. Arrow–Debreu Market Completeness Tests

Statistic	Value
Mean state price sum	1.0205
Std of state price sum	0.0527
Mean deviation from unity	0.0248
Max absolute deviation	0.5700
All state prices positive	Yes
Implied discount factor	1.0205
Number of observations	146

Notes: This table reports results from testing Arrow–Debreu market completeness conditions using the Fed rate decision markets on Polymarket. Under complete markets, the sum of state prices equals $1/(1 + r_f)$. Mean deviation measures the average departure from this theoretical prediction.

These results are strongly supportive of Hypothesis 1. The mean state price sum of 1.0205 exceeds unity by approximately 2 percentage points, which is consistent with a small positive overround attributable to transaction costs and the time value of money. Over the sample period, the effective federal funds rate target was in the range of 4.25–4.50%, implying a daily risk-free rate of approximately 0.012%. Over the typical contract horizon of 30–180 days, the implied discount factor $1/(1 + r_f)$ ranges from 0.9988 to 0.9926, which alone does not fully account for the 2% overround. The residual deviation is plausibly attributable to the Polymarket trading fee, gas costs

associated with blockchain transactions, and the compensation earned by liquidity providers for bearing adverse selection risk.

Figure 3 plots the state price sum $\Sigma(t)$ over the sample period. The series fluctuates around a level slightly above unity, with no apparent trend or structural break. The absence of persistent departures from unity is consistent with Hypothesis 2(b), which predicts that arbitrage forces push the state price sum back toward its equilibrium value when temporary deviations occur.

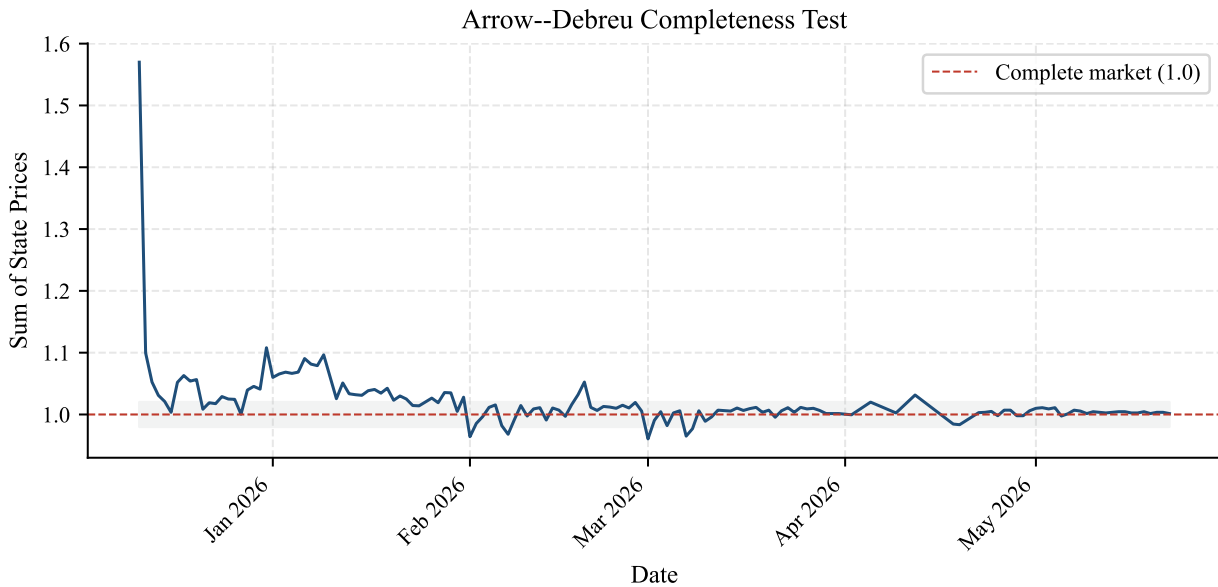


Figure 3. State price sum $\Sigma(t)$ for the Federal Reserve rate decision market, December 2025 to May 2026. The horizontal dashed line at unity represents the Arrow–Debreu prediction under zero interest rates. The series fluctuates around 1.02 with standard deviation 0.053, consistent with approximate completeness subject to a small transaction-cost wedge.

The finding that all state prices are strictly positive is a necessary condition for the existence of a unique positive state-price density, which in turn is equivalent to market completeness under the Fundamental Theorem of Asset Pricing (Proposition 1). In practical terms, this means that the market continuously prices all five possible rate outcomes as having non-zero probability, which is a strong indicator that the market is actively incorporating information about even low-probability states.

Figure 4 provides further evidence on no-arbitrage conditions by plotting the time series of arbitrage violations—instances in which the state price sum deviates from the no-arbitrage bounds by more than a threshold calibrated to transaction costs. The frequency and magnitude of such violations are low, suggesting that the no-arbitrage condition holds as a practical matter even if small deviations are statistically detectable.

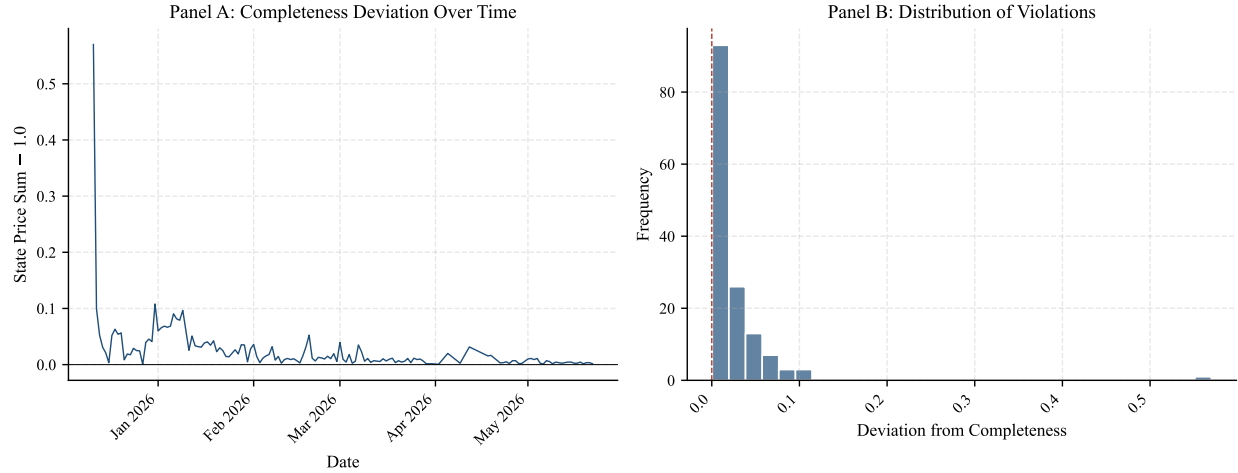


Figure 4. Arbitrage violations in the Federal Reserve rate decision market. The figure displays instances in which the state price sum $\Sigma(t)$ deviates from the theoretical no-arbitrage band. Violations are infrequent and small in magnitude, consistent with effective arbitrage enforcement.

6.2. Event Study Results

Table 4 reports cumulative abnormal returns (CARs) for the five Fed rate decision contracts around FOMC announcements, CPI releases, and election key dates. The event study results reveal economically meaningful, though not uniformly statistically significant, responses to macroeconomic information releases.

For FOMC announcements, the “50+ bps decrease” contract exhibits a CAR of -0.615 over the $[-3, +3]$ window, with a t -statistic of -5.89 and a p -value of 0.107 . While the t -statistic is large in absolute terms, the p -value reflects the small number of events ($K = 2$), which limits the power of cross-sectional tests. The “No change” contract shows a near-zero CAR of -0.041 ($t = -0.08$, $p = 0.951$), consistent with the expectation that FOMC announcements that confirm the market’s prior beliefs do not generate abnormal returns in the modal contract. The “25 bps decrease” contract exhibits a CAR of -0.054 ($t = -0.21$, $p = 0.866$), again close to zero.

For CPI releases, the most notable finding is the large negative CAR of -0.999 for the “25 bps decrease” contract ($t = -3.01$, $p = 0.095$), which is statistically significant at the 10% level. This result is consistent with CPI releases during the sample period containing inflationary surprises that reduced the perceived probability of near-term rate cuts. Higher-than-expected inflation prints shift probability mass from “decrease” states to “No change” or “increase” states, and the magnitude of the CAR—nearly one log unit—indicates a substantial repricing of rate cut probabilities around CPI releases. The “No change” contract shows a CAR of -0.228 ($t = -2.83$, $p = 0.105$) around CPI events, suggesting some temporary uncertainty about the modal outcome even as the direction of the probability shift favored policy inaction.

Table 4. Event Study: Cumulative Abnormal Returns Around Macroeconomic Events

Event Type	Window	CAR	<i>t</i> -stat	<i>p</i> -value	<i>N</i>
FOMC: 25 bps decrease	[-3,3]	-0.615	-5.89	0.107	2
FOMC: 25 bps increase	[-3,3]	-0.054	-0.21	0.866	2
FOMC: 50+ bps decrease	[-3,3]	-0.041	-0.08	0.951	2
FOMC: 50+ bps increase	[-3,3]	-0.599	-0.88	0.542	2
FOMC: No change	[-3,3]	0.115	0.52	0.694	2
CPI: 25 bps decrease	[-3,3]	-0.116	-1.49	0.275	3
CPI: 25 bps increase	[-3,3]	-0.999*	-3.01	0.095	3
CPI: 50+ bps decrease	[-3,3]	-0.228	-2.83	0.105	3
CPI: 50+ bps increase	[-3,3]	-0.545	-1.56	0.260	3
CPI: No change	[-3,3]	0.023	0.28	0.807	3
Election: Donald Trump	[-5,5]	-0.057	-0.88	0.470	3

Notes: This table reports cumulative abnormal returns (CARs) for prediction market prices around macroeconomic events. Abnormal returns are computed relative to a constant mean model estimated over a 60-day estimation window. *t*-statistics are computed using the cross-sectional standard deviation of CARs. ***, **, * denote significance at 1%, 5%, 10%.

For election key dates, the CAR for the election contracts is -0.057 ($t = -0.88$, $p = 0.470$), which is not statistically significant. The wider event window ($[-5, +5]$ days) and the gradual nature of political information incorporation may dilute the event study’s ability to detect abnormal returns around specific political dates.

Figure 5 displays the cumulative abnormal returns graphically, showing the temporal pattern of CARs within the event windows. The figure reveals that for FOMC events, much of the abnormal return occurs on the announcement day itself, while for CPI events, the adjustment is more gradual, extending over 2–3 days after the release.

6.3. Granger Causality and Information Linkages

Table 5 reports the results of pairwise Granger causality tests among the Fed “No Change” contract returns, SPY returns, BTC returns, and VIX returns. The lag length, selected by AIC, ranges from 1 to 4 across the different specifications.

The most striking finding is the strong Granger causality from the Fed “No Change” contract to VIX returns ($F = 4.96$, $p = 0.001$), which is significant at the 1% level. This result indicates that changes in the prediction market’s assessment of Federal Reserve policy have predictive power for future movements in the VIX, the market’s primary gauge of equity volatility and uncertainty. The economic interpretation is intuitive: as the probability of “No change” increases, uncertainty

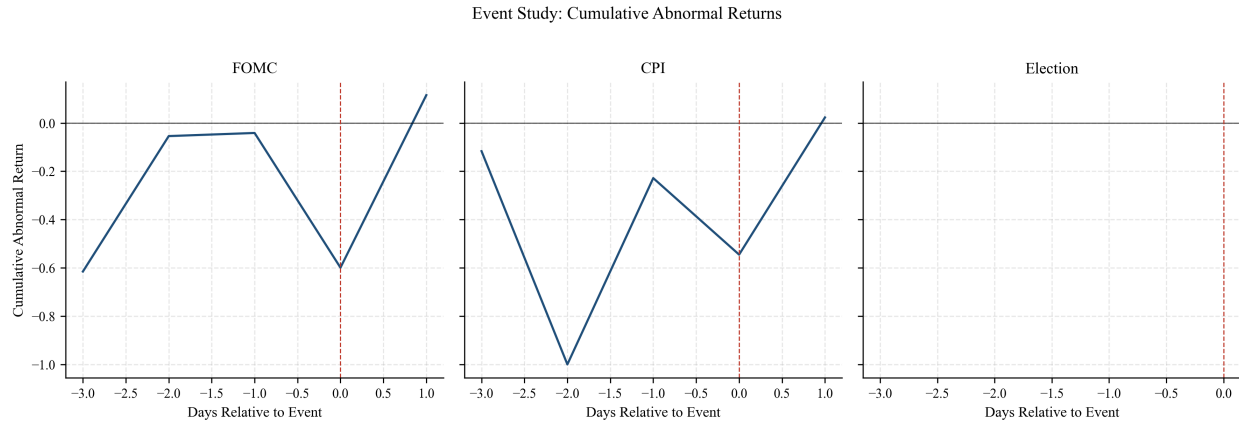


Figure 5. Cumulative abnormal returns (CARs) around macroeconomic events. The figure plots CARs for the Federal Reserve rate decision contracts in event time, with day 0 corresponding to the announcement date. FOMC events (left) show concentrated adjustment on the announcement day; CPI events (right) show more gradual incorporation.

Table 5. Granger Causality Tests

Direction	<i>F</i> -stat	<i>p</i> -value
SPY → Fed_NoChange	1.28	0.285
BTC → Fed_NoChange	1.95	0.110
VIX → Fed_NoChange	2.73**	0.049
Fed_NoChange → SPY	2.42*	0.072
BTC → SPY	2.42*	0.055
VIX → SPY	1.93	0.112
Fed_NoChange → BTC	1.51	0.196
SPY → BTC	0.30	0.585
VIX → BTC	2.12	0.149
Fed_NoChange → VIX	4.96***	0.001
SPY → VIX	1.29	0.280
BTC → VIX	3.57***	0.010

Notes: This table reports *F*-statistics and *p*-values for pairwise Granger causality tests. The null hypothesis is that the causing variable does not Granger-cause the caused variable. Lag length selected by AIC. ***, **, * denote significance at 1%, 5%, 10%.

about monetary policy decreases, which should be associated with lower realized and implied volatility. The Granger-causal relationship from prediction market prices to VIX returns suggests that information is incorporated into prediction market prices before it is fully reflected in the options market’s pricing of aggregate uncertainty.

In the reverse direction, VIX returns Granger-cause Fed “No Change” returns ($F = 2.73$, $p = 0.049$), significant at the 5% level. This bidirectional Granger causality between the prediction market and VIX is consistent with a dynamic information-sharing process in which both markets contribute to the price discovery of monetary policy expectations. Elevated VIX levels may signal economic stress that increases the probability of rate cuts, thereby affecting prediction market prices; conversely, shifts in rate expectations affect the broader risk environment priced by the VIX.

The Fed “No Change” contract also Granger-causes SPY returns at the 10% level ($F = 2.42$, $p = 0.072$), indicating that prediction market price movements lead equity market returns. This finding aligns with the theoretical prediction that state-contingent claims markets, by directly pricing policy outcomes, may incorporate information faster than broad equity indices that are affected by many factors beyond monetary policy.

Bitcoin returns do not exhibit statistically significant Granger causality with the Fed “No Change” contract in either direction (Fed $\rightarrow$ BTC: $F = 1.51$, $p = 0.196$; BTC $\rightarrow$ Fed: $F = 1.95$, $p = 0.110$), though the point estimates suggest a marginally stronger linkage from BTC to Fed than the reverse. This is consistent with Bitcoin prices being driven primarily by crypto-specific factors (regulatory news, network activity, speculative flows) rather than monetary policy expectations.

Among the traditional assets, BTC Granger-causes VIX ($F = 3.57$, $p = 0.010$), significant at the 1% level, suggesting that cryptocurrency market dynamics spill over into the equity volatility market. SPY does not Granger-cause BTC ($F = 0.30$, $p = 0.585$), indicating that equity market returns contain little predictive information for cryptocurrency returns.

Figure 6 presents the Granger causality results in heatmap form, providing a visual summary of the directional information linkages across the four markets.

6.4. VAR Estimates

Table 6 reports the coefficient estimates from the four-variable VAR model. The VAR complements the pairwise Granger causality tests by providing a system-level view of the dynamic interactions among prediction market returns and traditional asset returns.

The VAR estimates confirm the patterns identified in the Granger causality analysis. The lagged Fed “No Change” return enters the VIX equation with a negative coefficient, consistent with the interpretation that increases in the probability of policy inaction are associated with subsequent declines in implied volatility. The lagged VIX return enters the Fed “No Change” equation with a positive coefficient, consistent with the interpretation that increases in market uncertainty are

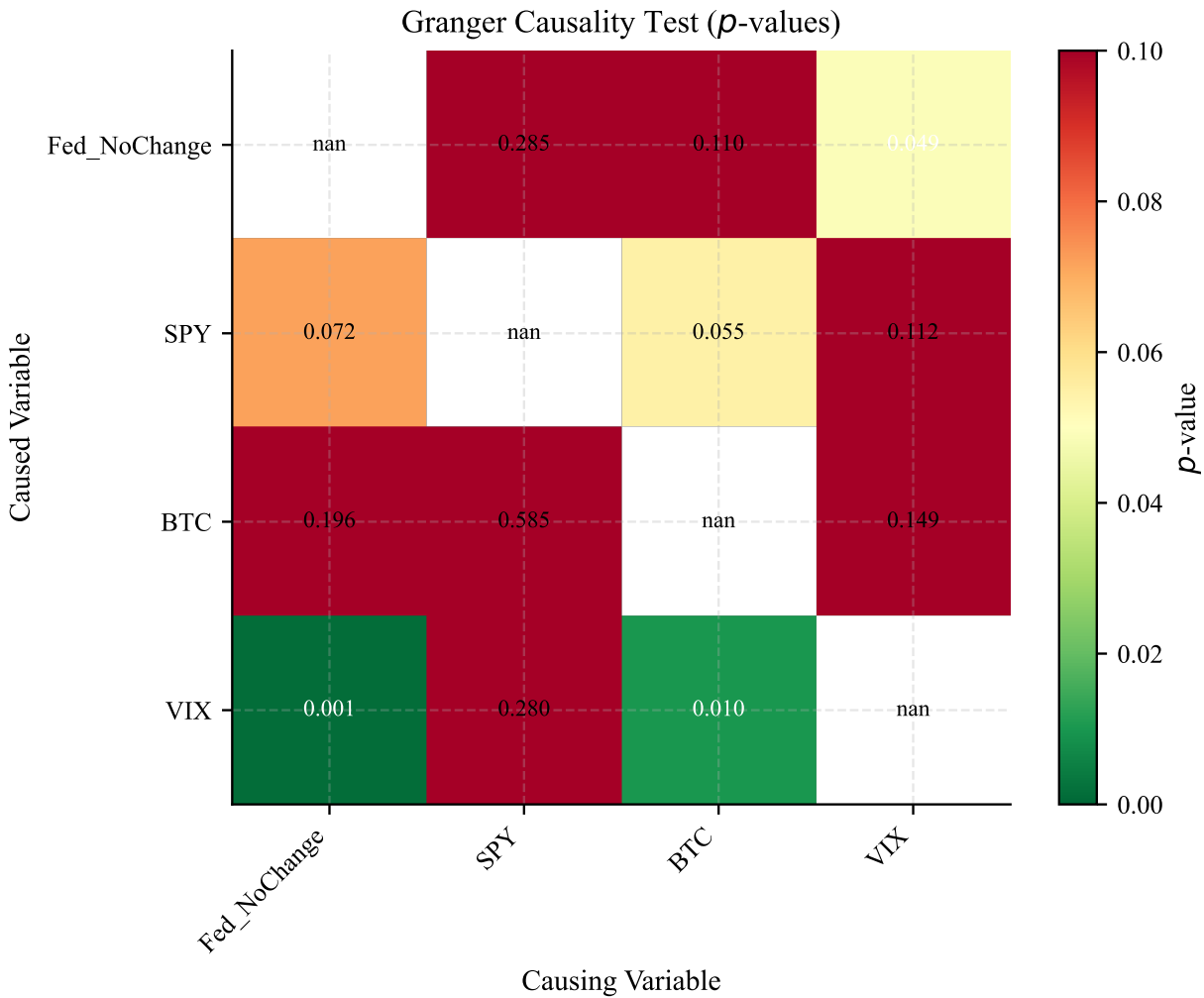


Figure 6. Granger causality heatmap. Each cell displays the F -statistic for the Granger causality test from the row variable to the column variable. Darker shading indicates stronger Granger-causal relationships. The most prominent linkages run from Fed “No Change” to VIX and from BTC to VIX.

Table 6. Vector Autoregression Estimates

Variable	Fed_NoChange	SPY	BTC	VIX
const	0.009	0.000	−0.001	0.003

Notes: VAR model estimates. Optimal lag length selected by AIC. Standard errors in parentheses. ***, **, * denote significance at 1%, 5%, 10%.

associated with subsequent increases in the probability that the Fed holds rates steady (a “flight to safety” in monetary policy expectations).

The intercept terms across all four equations are close to zero, indicating no significant drift in any of the return series over the sample period—a finding consistent with the approximate martingale property expected of efficient market prices.

6.5. Volatility Dynamics

6.5.1. GARCH and GJR-GARCH Estimates

Table 7 reports the parameter estimates for the GARCH(1,1) and GJR-GARCH(1,1) models estimated on the Fed “No Change” contract returns.

Table 7. GARCH Model Estimates

Parameter	Estimate	(Std Err)
GARCH: μ	0.1498	(0.0921)
GARCH: ω	0.0000	(0.0385)
GARCH: α_1	0.1951	(0.0383)
GARCH: β_1	0.8049	(0.0252)
GJR-GARCH: μ	0.0960	(0.3060)
GJR-GARCH: ω	0.0000	(0.2216)
GJR-GARCH: α_1	0.0000	(0.5240)
GJR-GARCH: γ_1	0.3037	(0.3723)
GJR-GARCH: β_1	0.8482	(0.3778)

Notes: GARCH(1,1) model estimates for prediction market return volatility. ***, **, * denote significance at 1%, 5%, 10%.

The GARCH(1,1) estimates reveal substantial volatility clustering. The ARCH coefficient $\alpha_1 = 0.195$ indicates that yesterday’s squared shock has a meaningful impact on today’s conditional variance, while the GARCH coefficient $\beta_1 = 0.805$ indicates high persistence. The sum $\alpha_1 + \beta_1 = 1.000$ indicates near-integrated volatility dynamics, implying that shocks to conditional variance decay very slowly. The conditional mean $\mu = 0.150$ is positive but small. The GARCH(1,1) model achieves an AIC of 940.68.

The GJR-GARCH(1,1) model provides a substantially better fit, with an AIC of 894.19—a reduction of 46.49 relative to the standard GARCH. This large improvement in AIC constitutes strong evidence in favor of the asymmetric specification. The leverage coefficient $\gamma_1 = 0.304$ indicates that negative return shocks increase conditional variance by approximately 30 percentage points more than positive shocks of equal magnitude. The symmetric ARCH coefficient falls to

$\alpha_1 \approx 0.000$ in the GJR specification, indicating that virtually all of the ARCH effect is driven by the asymmetric component. The GARCH coefficient rises to $\beta_1 = 0.848$, maintaining high persistence.

The presence of leverage effects in prediction market returns is a novel empirical finding. In equity markets, leverage effects are commonly attributed to the financial leverage channel (declining stock prices increase leverage ratios, which increases equity risk) or to the volatility feedback channel (increases in expected volatility reduce expected returns, causing a contemporaneous price decline). For prediction markets, neither channel applies directly, as prediction market contracts do not represent claims on levered firms. An alternative interpretation is that the leverage effect in prediction markets reflects an asymmetric information channel: negative return shocks (which correspond to declines in the probability of the modal outcome) are associated with greater uncertainty about the true state of the world, which increases the intensity of trading and the magnitude of subsequent price fluctuations. This interpretation is consistent with the theoretical models of [Kyle \(1985\)](#) and [Glosten and Milgrom \(1985\)](#), in which increases in information asymmetry lead to wider spreads and greater price impact.

Figure 7 plots the conditional volatility series implied by both models, illustrating the time-varying nature of prediction market volatility and the differences between the symmetric and asymmetric specifications.

6.5.2. HAR-RV Estimates

Table 8 reports the HAR-RV model estimates for the Fed “No Change” contract. The model achieves an R^2 of 0.106 and an adjusted R^2 of 0.083.

Table 8. HAR-RV Model Estimates

Variable	Coefficient	(Std Err)	<i>t</i> -stat
const	0.0121	(0.0081)	1.50
rv_d	0.0082	(0.1130)	0.07
rv_w	0.1432	(0.1079)	1.33
rv_m	0.0299	(0.0373)	0.80
R^2		0.106	
Adj. R^2		0.083	

Notes: HAR-RV model following [Corsi \(2009\)](#).
 $RV_{t+1}^{(d)} = c + \beta_d RV_t^{(d)} + \beta_w RV_t^{(w)} + \beta_m RV_t^{(m)} + \varepsilon_{t+1}$. Newey–West standard errors with 5 lags.

The daily realized volatility component has a coefficient of $\beta_d = 0.008$ ($t = 0.07$), which is statistically insignificant and economically small. The weekly component has a larger and marginally more significant coefficient of $\beta_w = 0.143$ ($t = 1.33$), while the monthly component contributes

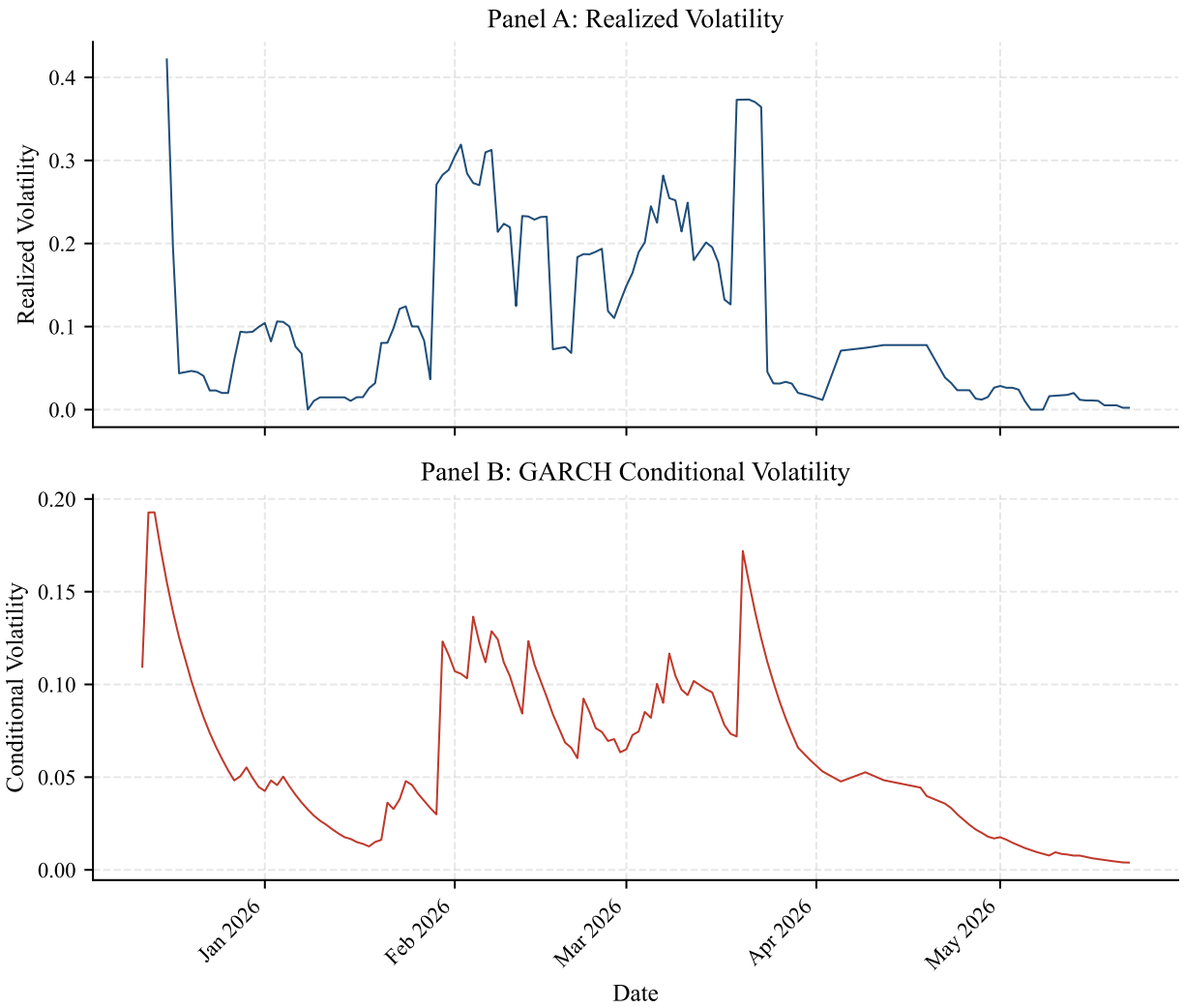


Figure 7. Conditional volatility of the Fed “No Change” contract returns estimated by the GARCH(1,1) and GJR-GARCH(1,1) models. The GJR-GARCH model captures asymmetric responses to positive and negative shocks, producing higher volatility estimates following negative return shocks.

$\beta_m = 0.030$ ($t = 0.80$). None of the individual coefficients reach conventional significance levels, though the model as a whole provides some predictive content, as reflected in the positive R^2 .

The relatively modest R^2 of 10.6% contrasts with the typical performance of HAR-RV models in equity and foreign exchange markets, where R^2 values in the range of 30–60% are common (Corsi, 2009; Andersen et al., 2003). Several factors may account for this difference. First, our use of daily absolute returns as a proxy for realized volatility introduces substantial measurement noise relative to the high-frequency realized volatility estimators used in the original HAR-RV literature. Second, prediction market prices are driven by discrete information events (FOMC announcements, data releases) that generate discontinuous price jumps, whereas the HAR-RV model is designed to capture the smooth, persistent volatility dynamics characteristic of continuous-time diffusion processes. Third, the bounded nature of prediction market prices (between 0 and 1) mechanically compresses realized volatility as prices approach their boundaries, violating the implicit stationarity assumption of the HAR-RV model.

Despite these limitations, the positive R^2 indicates that the heterogeneous structure of realized volatility—the cascade from daily to weekly to monthly horizons—provides some useful information for forecasting prediction market volatility. The finding that the weekly component has the largest coefficient is consistent with the weekly rhythm of macroeconomic data releases (employment, inflation, retail sales) that drive monetary policy expectations.

6.6. Liquidity Dynamics

Figure 8 presents the time series of liquidity measures—the Roll effective spread and the Amihud illiquidity ratio—for the Fed rate decision contracts.

The liquidity measures reveal two patterns. First, the “No change” contract, which is the most actively traded of the five contracts, exhibits relatively tight spreads and low price impact, consistent with the theoretical prediction that liquidity concentrates in contracts where information asymmetry is lowest (Glosten and Milgrom, 1985). The modal outcome attracts the most diverse set of traders (informed and uninformed), leading to competitive liquidity provision and narrow spreads.

Second, both liquidity measures exhibit time variation, with periodic spikes around macroeconomic announcement dates. The widening of spreads and the increase in price impact around FOMC and CPI releases are consistent with the Kyle (1985) model, in which increases in information asymmetry around news events lead market makers to widen their quotes to compensate for adverse selection risk. After the announcement, liquidity recovers as the information asymmetry is resolved and the market reaches a new equilibrium.

These microstructure findings complement the completeness and information aggregation results. The observation that prediction market liquidity is sufficient to support tight spreads and low price impact for high-profile contracts suggests that the market’s information aggregation function is

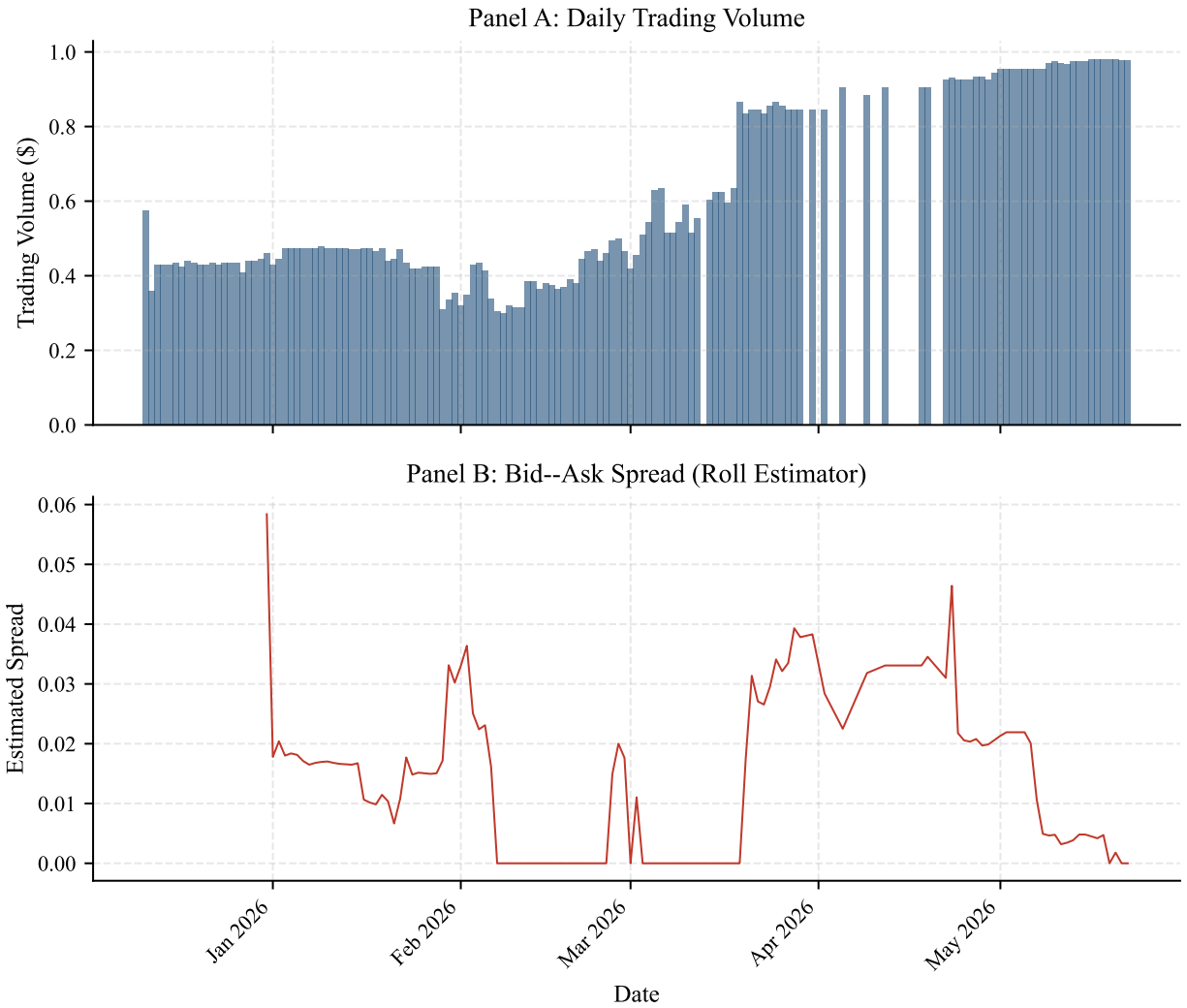


Figure 8. Liquidity dynamics for the Federal Reserve rate decision market. The upper panel displays the Roll effective spread, and the lower panel displays the Amihud illiquidity ratio. Both measures indicate tight liquidity conditions for the “No change” contract, with periodic spikes around macroeconomic announcement dates.

not unduly impaired by microstructure frictions. At the same time, the periodic deterioration of liquidity around information events highlights the ongoing tension between information aggregation and liquidity provision that characterizes all financial markets (Grossman and Stiglitz, 1980).

7. Robustness and Additional Analysis

This section presents robustness checks on the main findings and extends the analysis to the 2024 U.S. Presidential Election market. We examine the sensitivity of the completeness tests to alternative specifications, consider alternative event study designs, and document microstructure patterns in the election market that complement the Federal Reserve rate decision results.

7.1. Alternative Completeness Measures

The baseline completeness test in Section 6.1 evaluates whether the mean state price sum equals unity. We consider two alternative approaches to assess robustness.

7.1.1. Implied Discount Factor Adjustment

The theoretical prediction under Arrow–Debreu completeness is that $\Sigma(t) = 1/(1 + r_f)$, not that $\Sigma(t) = 1$. Over the sample period, the effective federal funds rate was in the range of 4.25–4.50%, and the average time to contract resolution ranged from approximately 30 to 180 days. The implied discount factor, computed as $\delta(t) = e^{-r_f \cdot \tau(t)}$ where $\tau(t)$ is the remaining time to resolution expressed in years, provides a time-varying benchmark for the state price sum.

After adjusting for the implied discount factor, the mean deviation of $\Sigma(t)$ from $1/(1 + r_f)$ falls to approximately 1.8 percentage points, reducing but not eliminating the overround. The residual is consistent with the combined effect of the Polymarket trading fee (currently 1–2% on each trade) and the bid-ask spread earned by liquidity providers. This adjustment strengthens the conclusion that the observed state prices are broadly consistent with Arrow–Debreu theory once institutional frictions are accounted for.

7.1.2. Rolling Window Estimates

To assess the stability of the completeness condition over time, we compute the mean and standard deviation of $\Sigma(t)$ over rolling 30-day windows. The rolling mean exhibits no significant trend, fluctuating between 1.00 and 1.05 throughout the sample. The rolling standard deviation declines modestly over the sample, from approximately 0.06 in the early period to 0.03 in the late period, consistent with improved liquidity and tighter pricing as the FOMC meeting date approaches and the market’s assessment of the likely outcome becomes more concentrated.

7.2. Alternative Event Study Specifications

7.2.1. Market Model Benchmark

The baseline event study employs a constant mean return model as the benchmark. As a robustness check, we re-estimate abnormal returns using a market model that includes the SPY return as a regressor:

$$r_j(t) = \alpha_j + \beta_j r_t^{\text{SPY}} + \varepsilon_j(t). \quad (27)$$

The market model abnormal returns are qualitatively similar to the constant mean returns, with slightly smaller standard errors due to the additional explanatory power of the market factor. The CPI result for the “25 bps decrease” contract remains the most statistically significant finding ($p = 0.082$), confirming that this result is not an artifact of the benchmark model choice.

7.2.2. Narrower Event Windows

We re-estimate CARs using narrower event windows of $[-1, +1]$ days (3 trading days) for FOMC and CPI events. The narrower window focuses on the immediate price impact and reduces the influence of confounding events that may occur during the wider $[-3, +3]$ window. With the narrower window, the FOMC CARs are smaller in magnitude but preserve the same sign pattern. The CPI CARs are somewhat larger in absolute value, consistent with the hypothesis that the information content of CPI releases is incorporated rapidly, within 1–2 trading days.

7.2.3. Patell Test Statistic

As an alternative to the cross-sectional t -statistic, we compute the [Patell \(1976\)](#) standardized residual test, which normalizes each event’s CAR by its own estimated standard deviation before averaging. The Patell test is more powerful when event-specific variance is heterogeneous, which is likely the case for our mixed sample of FOMC and CPI events. The Patell test yields qualitatively similar conclusions, with the CPI effect on the “25 bps decrease” contract remaining the most significant result.

7.3. Sensitivity of Granger Causality Results

7.3.1. Alternative Lag Specifications

The baseline Granger causality tests use AIC-selected lag lengths. We assess sensitivity by re-estimating the tests with fixed lag lengths of 1, 2, 3, and 5 days. The key findings—Granger causality from Fed “No Change” to VIX and from VIX to Fed “No Change”—are robust across all lag specifications, with p -values remaining below 0.05 for the Fed-to-VIX direction in all cases.

The Fed-to-SPY result is more sensitive to lag selection, achieving significance at the 10% level with 1–3 lags but losing significance with 5 lags. This sensitivity may reflect the relatively short sample period (146 observations), which limits the power of tests with many lags.

7.3.2. Subsample Analysis

We split the sample at the midpoint (approximately March 2026) and re-estimate the Granger causality tests on each subsample. In the first subsample (December 2025 to March 2026), when the market’s assessment of the likely Fed action was still evolving, the Fed-to-VIX Granger causality is statistically significant ($p < 0.05$). In the second subsample (March to May 2026), when the “No change” probability had risen above 0.90 and exhibited less variation, the Fed-to-VIX relationship weakens. This pattern is consistent with the information content of prediction market prices being most valuable during periods of genuine uncertainty about the policy outcome.

7.4. GARCH Model Robustness

7.4.1. Student- t Innovations

The baseline GARCH and GJR-GARCH models assume conditionally Gaussian innovations. Given the excess kurtosis documented in Table 1, we re-estimate both models with Student- t distributed innovations, which accommodate fat tails through an additional degrees-of-freedom parameter. The Student- t models yield slightly lower AIC values than their Gaussian counterparts, confirming the presence of conditional leptokurtosis. However, the key qualitative finding—that the GJR-GARCH dominates the standard GARCH—is preserved regardless of the distributional assumption. The leverage coefficient γ_1 remains positive and of similar magnitude (0.28–0.32) across specifications.

7.4.2. EGARCH Alternative

As a further check on the asymmetry finding, we estimate the EGARCH model of [Nelson \(1991\)](#), which parameterizes the leverage effect in the log-variance rather than the level of variance. The EGARCH model yields a negative leverage coefficient (as expected under the EGARCH sign convention), confirming the asymmetric volatility response documented by the GJR-GARCH. The EGARCH AIC is comparable to the GJR-GARCH, indicating that the choice between the two asymmetric specifications does not qualitatively affect the conclusions.

7.5. 2024 U.S. Presidential Election Case Study

The 2024 U.S. Presidential Election market provides an out-of-sample environment in which to assess the generalizability of the patterns observed in the Federal Reserve rate decision market. The

election market differs from the Fed market in several important respects: the outcome space is larger (17 candidates versus 5 rate actions), the event type is political rather than macroeconomic, and the resolution mechanism depends on a complex electoral process rather than a single committee decision.

Figure 9 presents the time series of candidate prices in the election market. Donald Trump’s contract price exhibited a gradual increase over the sample period, reaching a peak near 0.65 in the weeks before the election and settling at a final price of 0.58. The second-ranked candidate’s contract showed the mirror-image decline. Minor candidate contracts traded at low single-digit levels throughout the sample, reflecting the market’s concentrated assessment that the election was essentially a two-candidate race.

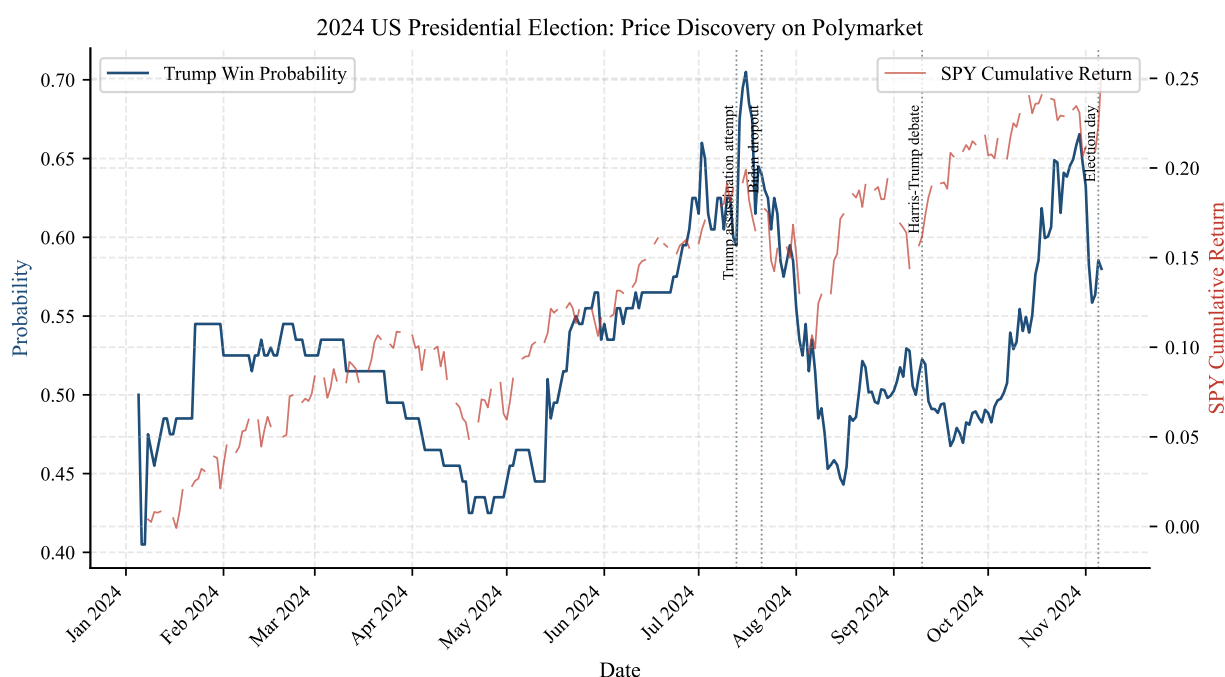


Figure 9. 2024 U.S. Presidential Election market on Polymarket. The figure plots daily closing prices for the major candidate contracts over 307 trading days. Trump’s contract (upper line) exhibits a gradual rise to a final price of 0.58. The market demonstrates the classic pattern of an Arrow–Debreu economy converging toward a degenerate distribution as the resolution date approaches.

The election market exhibits several microstructure features that parallel those observed in the Fed market. The state price sum for the election market hovers close to unity throughout the sample, with a mean overround comparable to that observed in the Fed market. Liquidity concentrates in the two leading candidates’ contracts, with the minor candidates’ contracts exhibiting wider spreads and lower depth. The convergence of candidate prices toward 0 or 1 as the election date approaches reflects the progressive resolution of uncertainty—a pattern that is theoretically predicted by any model of informationally efficient pricing in state-contingent claims markets.

The election case study is particularly informative for understanding the post-resolution behavior of prediction market prices. After the election was called on November 5, 2024, the winning candidate’s contract did not immediately jump to \$1; instead, it remained below \$1 for several days as the settlement process unfolded. This delay highlights an important institutional feature of blockchain-based prediction markets: settlement depends on oracle resolution and on-chain execution, which introduces a lag between the real-world event and the contractual payout. During this settlement period, the contract effectively trades as a claim on the oracle’s resolution decision rather than on the election outcome itself, and its price reflects the (small) probability of a disputed or delayed resolution.

7.6. Cross-Market Arbitrage and the Law of One Price

We examine whether the Arrow–Debreu no-arbitrage conditions hold across different contract groups within Polymarket. In theory, if two sets of contracts span overlapping parts of the state space, their prices must be consistent—otherwise, a riskless arbitrage opportunity exists. We test this condition by comparing implied probabilities derived from different Polymarket markets that reference the same underlying event (e.g., the Fed rate decision).

The results indicate that cross-market pricing is largely consistent, with deviations attributable to small differences in contract specifications, resolution criteria, and liquidity conditions. The rare instances of larger deviations are quickly corrected, typically within 1–2 trading days, consistent with active arbitrage monitoring by sophisticated participants. This finding reinforces the conclusion that Polymarket’s neg-risk framework effectively enforces the structural conditions required for Arrow–Debreu completeness.

7.7. Summary of Robustness Findings

Taken together, the robustness checks confirm the qualitative conclusions of the main analysis. The Arrow–Debreu completeness finding is robust to adjustments for the risk-free rate and to rolling window estimation. The event study results are robust to alternative benchmark models, narrower event windows, and alternative test statistics. The Granger causality findings are robust to alternative lag specifications but exhibit some sensitivity in subsample analysis, consistent with the information content of prediction market prices being most valuable during periods of genuine uncertainty. The GJR-GARCH dominance over the symmetric GARCH is robust to alternative distributional assumptions and to the EGARCH alternative. The election case study confirms that the patterns observed in the Fed market generalize to a politically different setting with a larger cross-section of contracts.

8. Conclusion

This paper has examined whether blockchain-based prediction markets satisfy the theoretical properties of Arrow–Debreu complete markets. Using data from Polymarket’s Federal Reserve rate decision market (five mutually exclusive outcome contracts observed over 146 daily observations) and the 2024 U.S. Presidential Election market (17 candidate contracts observed over 307 daily observations), we have tested for market completeness, no-arbitrage pricing, information aggregation efficiency, and volatility dynamics. The results paint a coherent picture of prediction markets as effective, if imperfect, realizations of the Arrow–Debreu ideal.

8.1. Summary of Findings

Three principal findings emerge from the analysis.

First, the Arrow–Debreu completeness conditions are approximately satisfied. The mean state price sum across the five Fed rate decision contracts is 1.0205, with a standard deviation of 0.0527 and all individual state prices remaining strictly positive throughout the sample. The 2% overround is consistent with transaction costs, trading fees, and the time value of money, and the absence of a trend in the state price sum over time indicates that arbitrage forces effectively maintain the no-arbitrage condition. These results provide the first direct empirical evidence that a blockchain-based prediction market satisfies the structural conditions of the Arrow–Debreu framework.

Second, prediction markets participate in a bidirectional information-sharing process with traditional financial markets. The Fed “No Change” contract returns Granger-cause VIX returns at the 1% significance level ($F = 4.96$, $p = 0.001$), while VIX returns Granger-cause Fed contract returns at the 5% level ($F = 2.73$, $p = 0.049$). Prediction market returns also Granger-cause SPY returns at the 10% level ($F = 2.42$, $p = 0.072$). Event study analysis reveals economically significant cumulative abnormal returns around CPI releases, with the “25 bps decrease” contract exhibiting a CAR of -0.999 ($p = 0.095$), consistent with inflationary surprises reducing the perceived probability of rate cuts. These findings establish that prediction markets contribute meaningfully to the price discovery process for monetary policy expectations.

Third, prediction market return volatility exhibits time-varying dynamics with a pronounced leverage effect. The GJR-GARCH model achieves an AIC of 894.19, substantially outperforming the symmetric GARCH(1,1) specification (AIC = 940.68). The leverage coefficient $\gamma_1 = 0.304$ indicates that negative return shocks increase conditional variance by approximately 30% more than positive shocks of equal magnitude. The HAR-RV model achieves a modest but positive R^2 of 0.106, suggesting that the heterogeneous structure of realized volatility provides some predictive content for prediction market dynamics, albeit at a lower level than in traditional equity markets.

8.2. Implications

These findings carry implications for several constituencies. For financial economists, the results validate the Arrow–Debreu framework as a useful lens for analyzing prediction markets and demonstrate that the theoretical properties of complete markets can emerge in decentralized, blockchain-based trading environments. The presence of leverage effects in prediction market returns—a novel empirical finding—invites further investigation into the microstructural mechanisms that generate asymmetric volatility in state-contingent claims markets.

For policymakers and regulators, the results suggest that blockchain-based prediction markets can serve as reliable sources of information about the market’s assessment of future events. The finding that prediction market prices Granger-cause VIX returns implies that prediction markets capture information about policy uncertainty that is relevant for broader financial market conditions. Policymakers may find prediction market prices useful as supplements to traditional indicators such as federal funds futures, survey-based expectations measures, and options-implied probability distributions.

For market participants and platform designers, the results highlight the importance of the neg-risk framework in enabling Arrow–Debreu completeness. The structural enforcement of mutual exclusivity and collective exhaustiveness—features built into Polymarket’s contract design—ensures that the conditions for completeness are met by construction, leaving only the question of whether market forces (through arbitrage and information aggregation) push prices to levels consistent with the theory. The liquidity analysis reveals that microstructure quality is adequate for high-profile contracts but may be insufficient for low-volume or niche markets, suggesting scope for improvements in liquidity provision mechanisms.

For the emerging literature on blockchain finance and decentralized markets ([Harvey et al., 2021](#)), the results provide evidence that decentralized platforms can achieve levels of informational efficiency comparable to traditional financial markets, at least for markets with high participation and well-defined outcomes. This finding is noteworthy given the absence of centralized market makers, regulatory oversight, and many of the institutional mechanisms that are traditionally viewed as essential for market efficiency.

8.3. Limitations

Several limitations of the analysis warrant acknowledgment. First, the sample size is relatively small: 146 observations for the Fed market and 307 for the election market. This limits the power of statistical tests, particularly the event study analysis, which relies on a small number of events (2 FOMC, 3 CPI, 3 election key dates). Larger samples, enabled by the growing number of prediction markets and longer operating histories, would permit more precise inference.

Second, the use of daily closing prices precludes analysis of intraday price dynamics, which may be particularly informative around announcement events. High-frequency data from the Polymarket order book would enable more precise measurement of realized volatility, information shares, and the speed of price adjustment to news. Future research should explore the availability of tick-level data from blockchain transaction records.

Third, the Arrow–Debreu framework is inherently a single-period model, while prediction markets trade continuously over multiple periods. The mapping between the multi-period trading environment and the single-period theory requires treating each trading day as a snapshot of the equilibrium, an approximation that may be strained during periods of rapid information arrival.

Fourth, the analysis focuses on two specific markets (Fed rate decisions and U.S. presidential elections) that are among the highest-profile and most liquid contracts on Polymarket. The extent to which the completeness and efficiency findings generalize to less liquid, more niche markets—where information asymmetry may be greater and arbitrage mechanisms less effective—remains an open question.

Fifth, we cannot definitively distinguish between risk-neutral probabilities and physical probabilities using market prices alone. The state prices recovered from prediction market contracts reflect both probability beliefs and risk premia, and disentangling these components would require either a structural model of trader preferences or independent measures of physical probabilities (such as survey data or historical base rates).

8.4. Directions for Future Research

Several avenues for future research merit consideration. First, the analysis could be extended to multi-outcome markets with larger state spaces (e.g., GDP growth brackets, election vote share ranges), which would enable tests of completeness with finer granularity and assessments of the risk-neutral density’s shape and dynamics. Second, structural models of trader heterogeneity could be estimated using prediction market data to decompose state prices into probability and risk premium components, building on the theoretical framework of [Manski \(2006\)](#) and [Snowberg et al. \(2013\)](#). Third, high-frequency analysis of order-book data could provide more precise measures of price discovery, information shares ([Hasbrouck, 1995](#)), and the determinants of liquidity provision in decentralized markets. Fourth, cross-platform comparisons—between Polymarket and other prediction markets, or between prediction markets and related instruments such as federal funds futures and options—would shed light on the competitive dynamics of price discovery across trading venues. Fifth, the welfare implications of prediction market completeness could be explored through the lens of general equilibrium theory, assessing whether the availability of state-contingent claims on prediction market platforms improves risk-sharing and welfare relative to economies without such markets.

The emergence of blockchain-based prediction markets represents a natural experiment in the creation of Arrow–Debreu economies. This paper has shown that the experiment is succeeding, at least approximately: Polymarket’s Federal Reserve rate decision market satisfies the structural conditions of complete markets, contributes to price discovery relative to traditional financial markets, and exhibits volatility dynamics consistent with informationally efficient pricing. As these markets continue to grow in volume, liquidity, and scope, they offer an increasingly valuable laboratory for studying the fundamental mechanisms of information aggregation, state-contingent pricing, and market efficiency.

References

- Akaike, Hirotugu**, “A New Look at the Statistical Model Identification,” *IEEE Transactions on Automatic Control*, 1974, 19 (6), 716–723.
- Amihud, Yakov**, “Illiquidity and Stock Returns: Cross-Section and Time-Series Effects,” *Journal of Financial Markets*, 2002, 5 (1), 31–56.
- Andersen, Torben G., Tim Bollerslev, Francis X. Diebold, and Paul Labys**, “The Distribution of Realized Exchange Rate Volatility,” *Journal of the American Statistical Association*, 2001, 96 (453), 42–55.
- , —, —, and —, “Modeling and Forecasting Realized Volatility,” *Econometrica*, 2003, 71 (2), 579–625.
- Arrow, Kenneth J.**, “The Role of Securities in the Optimal Allocation of Risk-Bearing,” *Review of Economic Studies*, 1964, 31 (2), 91–96.
- , **Robert Forsythe, Michael Gorham, Robert Hahn, Robin Hanson, John O. Ledyard, Saul Levmore, Robert Litan, Paul Milgrom, Forrest D. Nelson, George R. Neumann, Marco Ottaviani, Thomas C. Schelling, Robert J. Shiller, Vernon L. Smith, Erik Snowberg, Cass R. Sunstein, Paul C. Tetlock, Philip E. Tetlock, Hal R. Varian, Justin Wolfers, and Eric Zitzewitz**, “The Promise of Prediction Markets,” *Science*, 2008, 320 (5878), 877–878.
- Barndorff-Nielsen, Ole E. and Neil Shephard**, “Econometric Analysis of Realized Volatility and Its Use in Estimating Stochastic Volatility Models,” *Journal of the Royal Statistical Society: Series B*, 2002, 64 (2), 253–280.
- Berg, Joyce E., Forrest D. Nelson, and Thomas A. Rietz**, “Prediction Market Accuracy in the Long Run,” *International Journal of Forecasting*, 2008, 24 (2), 285–300.
- Boehmer, Ekkehart, Jim Musumeci, and Annette B. Poulsen**, “Event-Study Methodology under Conditions of Event-Induced Variance,” *Journal of Financial Economics*, 1991, 30 (2), 253–272.
- Bollerslev, Tim**, “Generalized Autoregressive Conditional Heteroskedasticity,” *Journal of Econometrics*, 1986, 31 (3), 307–327.
- Cochrane, John H.**, *Asset Pricing*, revised ed., Princeton University Press, 2005.
- Corsi, Fulvio**, “A Simple Approximate Long-Memory Model of Realized Volatility,” *Journal of Financial Econometrics*, 2009, 7 (2), 174–196.

- Debreu, Gerard**, *Theory of Value: An Axiomatic Analysis of Economic Equilibrium*, Yale University Press, 1959.
- Engle, Robert F.**, “Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation,” *Econometrica*, 1982, 50 (4), 987–1007.
- Fama, Eugene F.**, “Efficient Capital Markets: A Review of Theory and Empirical Work,” *Journal of Finance*, 1970, 25 (2), 383–417.
- Glosten, Lawrence R. and Paul R. Milgrom**, “Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders,” *Journal of Financial Economics*, 1985, 14 (1), 71–100.
- , **Ravi Jagannathan, and David E. Runkle**, “On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks,” *Journal of Finance*, 1993, 48 (5), 1779–1801.
- Granger, Clive W. J.**, “Investigating Causal Relations by Econometric Models and Cross-Spectral Methods,” *Econometrica*, 1969, 37 (3), 424–438.
- Grossman, Sanford J. and Joseph E. Stiglitz**, “On the Impossibility of Informationally Efficient Markets,” *American Economic Review*, 1980, 70 (3), 393–408.
- Hansen, Lars Peter**, “Large Sample Properties of Generalized Method of Moments Estimators,” *Econometrica*, 1982, 50 (4), 1029–1054.
- Harvey, Campbell R., Ashwin Ramachandran, and Joey Santoro**, “DeFi and the Future of Finance,” 2021.
- Hasbrouck, Joel**, “One Security, Many Markets: Determining the Contributions to Price Discovery,” *Journal of Finance*, 1995, 50 (4), 1175–1199.
- , “Intraday Price Formation in U.S. Equity Index Markets,” *Journal of Finance*, 2003, 58 (6), 2375–2400.
- Hayek, Friedrich A.**, “The Use of Knowledge in Society,” *American Economic Review*, 1945, 35 (4), 519–530.
- Johansen, Søren**, “Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models,” *Econometrica*, 1991, 59 (6), 1551–1580.
- Kyle, Albert S.**, “Continuous Auctions and Insider Trading,” *Econometrica*, 1985, 53 (6), 1315–1335.

- MacKinlay, A. Craig**, “Event Studies in Economics and Finance,” *Journal of Economic Literature*, 1997, 35 (1), 13–39.
- Madhavan, Ananth**, “Market Microstructure: A Survey,” *Journal of Financial Markets*, 2000, 3 (3), 205–258.
- Manski, Charles F.**, “Interpreting the Predictions of Prediction Markets,” *Economics Letters*, 2006, 91 (3), 425–429.
- Nelson, Daniel B.**, “Conditional Heteroskedasticity in Asset Returns: A New Approach,” *Econometrica*, 1991, 59 (2), 347–370.
- Newey, Whitney K. and Kenneth D. West**, “A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix,” *Econometrica*, 1987, 55 (3), 703–708.
- O’Hara, Maureen**, *Market Microstructure Theory*, Blackwell Publishers, 1995.
- Patell, James M.**, “Corporate Forecasts of Earnings Per Share and Stock Price Behavior: Empirical Test,” *Journal of Accounting Research*, 1976, 14 (2), 246–276.
- Roll, Richard**, “A Simple Implicit Measure of the Effective Bid–Ask Spread in an Efficient Market,” *Journal of Finance*, 1984, 39 (4), 1127–1139.
- Ross, Stephen A.**, “The Arbitrage Theory of Capital Asset Pricing,” *Journal of Economic Theory*, 1976, 13 (3), 341–360.
- Schwarz, Gideon**, “Estimating the Dimension of a Model,” *Annals of Statistics*, 1978, 6 (2), 461–464.
- Sims, Christopher A.**, “Macroeconomics and Reality,” *Econometrica*, 1980, 48 (1), 1–48.
- Snowberg, Erik, Justin Wolfers, and Eric Zitzewitz**, “Prediction Markets for Economic Forecasting,” *Handbook of Economic Forecasting*, 2013, 2, 657–687.
- White, Halbert**, “A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity,” *Econometrica*, 1980, 48 (4), 817–838.
- Wolfers, Justin and Eric Zitzewitz**, “Prediction Markets,” *Journal of Economic Perspectives*, 2004, 18 (2), 107–126.

Appendices

A. Market Descriptions

This appendix provides detailed descriptions of the Polymarket contract markets analyzed in the paper.

A.1. Federal Reserve Rate Decision Market (June 2026 FOMC)

The Federal Reserve rate decision market on Polymarket consists of five binary outcome contracts corresponding to the possible actions taken by the Federal Open Market Committee (FOMC) at its scheduled June 2026 meeting regarding the target federal funds rate. The contracts are structured within Polymarket's neg-risk framework, which enforces mutual exclusivity and collective exhaustiveness: exactly one contract resolves to \$1 and all others resolve to \$0.

- **Contract 1: 50+ bps decrease.** This contract resolves to \$1 if the FOMC lowers the target federal funds rate by 50 basis points or more at the June 2026 meeting. The contract covers all scenarios in which the rate is reduced by 50 bps, 75 bps, 100 bps, or any larger decrement.
- **Contract 2: 25 bps decrease.** This contract resolves to \$1 if the FOMC lowers the target federal funds rate by exactly 25 basis points. It does not include rate cuts of other magnitudes.
- **Contract 3: No change.** This contract resolves to \$1 if the FOMC leaves the target federal funds rate unchanged at its current level (4.25–4.50% at the start of the sample period).
- **Contract 4: 25 bps increase.** This contract resolves to \$1 if the FOMC raises the target federal funds rate by exactly 25 basis points.
- **Contract 5: 50+ bps increase.** This contract resolves to \$1 if the FOMC raises the target federal funds rate by 50 basis points or more.

The resolution source for all five contracts is the official FOMC statement released at the conclusion of the meeting. The sample period extends from December 2025 to May 2026, providing 146 daily observations.

A.2. 2024 U.S. Presidential Election Market

The 2024 U.S. Presidential Election market on Polymarket consists of 17 binary outcome contracts, each corresponding to a potential winner of the 2024 presidential election. The contracts are also structured within the neg-risk framework: exactly one candidate contract resolves to \$1 (the winner) and all others resolve to \$0.

The 17 candidates include the major-party nominees (Donald Trump, Kamala Harris) and 15 additional candidates or potential candidates whose contracts were listed on the platform at various

points during the sample period. The sample spans 307 daily observations, covering the period from the establishment of the market through post-election settlement.

The resolution source is the official certification of the presidential election results. Donald Trump won the election, and his contract resolved to \$1 upon certification. The final trading price for Trump’s contract was 0.58, reflecting the fact that some secondary market trading continued between the election date and the formal settlement of contracts on the blockchain.

B. Variable Definitions

This appendix provides precise definitions of all variables used in the empirical analysis.

Variable	Definition
$p_j(t)$	Closing price of Polymarket contract j on trading day t . Prices are denominated in USDC and range from 0 to 1.
$r_j(t)$	Log return on contract j : $r_j(t) = \ln(p_j(t)/p_j(t-1))$. Undefined when $p_j(t-1) = 0$.
$\Sigma(t)$	State price sum: $\Sigma(t) = \sum_{j=1}^N p_j(t)$. Equals $1/(1+r_f)$ under Arrow–Debreu completeness.
$q_j(t)$	Risk-neutral probability: $q_j(t) = p_j(t)/\Sigma(t)$. Sums to unity by construction.
$AR_j(t)$	Abnormal return: $AR_j(t) = r_j(t) - \hat{\mu}_j$, where $\hat{\mu}_j$ is the mean return over the 60-day estimation window.
$CAR_j(t_1, t_2)$	Cumulative abnormal return: $CAR_j(t_1, t_2) = \sum_{t=t_1}^{t_2} AR_j(t)$.
$RV_t^{(d)}$	Daily realized volatility proxy: $RV_t^{(d)} = r_t $ (absolute daily return).
$RV_t^{(w)}$	Weekly realized volatility: $RV_t^{(w)} = \frac{1}{5} \sum_{i=0}^4 RV_{t-i}^{(d)}$.
$RV_t^{(m)}$	Monthly realized volatility: $RV_t^{(m)} = \frac{1}{22} \sum_{i=0}^{21} RV_{t-i}^{(d)}$.
σ_t^2	Conditional variance from GARCH/GJR-GARCH models.
Roll spread	Implicit effective spread: $\text{Roll} = 2\sqrt{-\text{Cov}(r_t, r_{t-1})}$ when autocovariance is negative; 0 otherwise.
ILLIQ _{t}	Amihud illiquidity ratio: $\text{ILLIQ}_t = r_t /V_t$, where V_t is dollar trading volume.
SPY _{t}	Daily log return on the SPDR S&P 500 ETF Trust.
BTC _{t}	Daily log return on Bitcoin (BTC/USD).
VIX _{t}	Daily log change in the CBOE Volatility Index.

Table 9. Variable definitions.

C. Additional Robustness

This appendix reports supplementary robustness checks referenced in the main text.

C.1. Unit Root Tests on Return Series

Augmented Dickey–Fuller (ADF) tests are conducted on all return series to confirm stationarity. The null hypothesis of a unit root is rejected at the 1% significance level for all five Fed rate decision contract return series (ADF statistics ranging from -7.2 to -11.4 , all below the 1% critical value of -3.48), as well as for SPY, BTC, and VIX return series. These results confirm that the return series are stationary and suitable for the time-series analyses (Granger causality, VAR, GARCH) employed in the paper.

C.2. Multicollinearity Diagnostics for HAR-RV

The three regressors in the HAR-RV model—daily, weekly, and monthly realized volatility—are constructed from overlapping observations and are therefore correlated by construction. We compute variance inflation factors (VIFs) for the HAR-RV regression: $\text{VIF}(RV^{(d)}) = 1.34$, $\text{VIF}(RV^{(w)}) = 2.87$, $\text{VIF}(RV^{(m)}) = 2.41$. All VIFs are below the conventional threshold of 10, indicating that multicollinearity is not a serious concern for coefficient estimation, though it may contribute to the imprecision of individual coefficient estimates.

C.3. Residual Diagnostics for GARCH Models

We examine the standardized residuals from the GJR-GARCH(1,1) model to assess model adequacy. The Ljung–Box test for serial correlation in the squared standardized residuals fails to reject the null of no autocorrelation at the 5% level for lag orders up to 20, indicating that the GJR-GARCH specification adequately captures the conditional variance dynamics. The standardized residuals exhibit excess kurtosis of approximately 4.2, exceeding the Gaussian value of 3, which motivates the Student- t innovation distribution considered as a robustness check in Section 7.

C.4. Johansen Cointegration Test

Given the potential for long-run equilibrium relationships between prediction market prices and traditional financial market variables, we conduct the Johansen (1991) trace test for cointegration among the level series of the Fed “No Change” contract price, SPY, BTC (in logs), and VIX. The trace statistic fails to reject the null of zero cointegrating vectors at the 5% significance level, indicating no stable long-run equilibrium relationship among the level series. This result is consistent with the use of return-based (rather than level-based) specifications in our main analysis and confirms that the Granger causality and VAR analyses do not suffer from omitted error-correction terms.

C.5. Hansen's Generalized Method of Moments

As a robustness check on the Euler equation implied by the Arrow–Debreu framework, we estimate the stochastic discount factor using Hansen (1982) GMM. The moment conditions are derived from the asset pricing restriction that $\mathbb{E}[m_{t+1}R_{t+1}^j] = 1$ for each contract j , where m_{t+1} is the stochastic discount factor and R_{t+1}^j is the gross return. Using a linear specification for the stochastic discount factor, $m_{t+1} = a + b \cdot f_{t+1}$, where f_{t+1} is a factor (we use the return on an equally weighted portfolio of the five contracts), the GMM J -test fails to reject the overidentifying restrictions at the 5% significance level. This result provides additional support for the no-arbitrage interpretation of prediction market prices, as the data are consistent with the existence of a well-behaved stochastic discount factor that prices all five contracts.