

Order Flow Imbalance and Price Discovery on Prediction Markets: Evidence from Kalshi

Ash Malék^{a,*}

^a*Department of Economics, Fordham University, 441 East Fordham Road, New York, NY 10458, United States*

Working Paper, September 2026

Abstract

We apply the Order Flow Imbalance (OFI) framework of Cont et al. (2014) to CFTC-regulated prediction markets on Kalshi. Using 143,297 orderbook snapshots collected over 13 days across sports, weather, and economic event contracts, we document a statistically significant positive relationship between OFI and contemporaneous mid-price changes. The explanatory power of OFI varies substantially across contract categories: sports championship futures exhibit a mean R^2 of 0.29, weather contracts 0.13, and macroeconomic contracts 0.02. A pooled panel regression with market fixed effects confirms the aggregate relationship (standardized $\beta = 0.45$, $p < 0.001$). Results are robust to subsample splitting, winsorization, and exclusion of illiquid contracts. These findings provide the first evidence that the standard microstructure framework for order flow and price formation extends to binary event contracts on prediction markets, with cross-category heterogeneity suggesting systematic differences in the information content of order flow across contract types.

Keywords: Order flow imbalance, Prediction markets, Market microstructure, Kalshi, Price discovery, Binary event contracts

*Corresponding author. Tel.: +1-929-493-8189
Email address: mm447@fordham.edu (Ash Malék)

1. Introduction

Prediction markets have emerged as a significant new venue for price discovery on real-world events. Following the CFTC’s approval of Kalshi as a regulated designated contract market in 2020, trading volumes on event contracts have grown rapidly, with Kalshi alone processing billions of dollars in notional volume across political, economic, sports, and weather outcomes. Despite this growth, the microstructure of prediction markets remains largely unstudied. While a robust literature examines how prices form on equity and derivatives exchanges, the mechanisms through which order flow drives price changes on binary event contracts are not well understood.

This paper addresses that gap by applying the Order Flow Imbalance (OFI) framework of Cont et al. (2014) to Kalshi prediction markets. CKS demonstrated that a simple measure of net order flow—computed from changes in best bid and ask depth between consecutive orderbook snapshots—explains a substantial fraction of contemporaneous price variation on equity markets. We adapt their methodology for the binary contract structure of prediction markets and test whether the same relationship holds across five contract categories: economics, sports, cryptocurrency, and weather.

Using 143,297 orderbook snapshots collected over 13 days via Kalshi’s public REST API, we document three main findings. First, OFI has a statistically significant positive relationship with contemporaneous mid-price changes across the majority of contracts in our panel, confirming that the CKS framework extends to prediction markets. Second, the explanatory power of OFI varies dramatically across contract categories: sports championship futures exhibit a mean R^2 of 0.29, weather contracts 0.13, and macroeconomic contracts 0.02. Third, this cross-category heterogeneity is robust to subsample splitting, win-sORIZATION of extreme observations, and exclusion of illiquid contracts.

These findings contribute to two literatures. To the market microstructure literature, we provide the first application of the CKS order flow framework to CFTC-regulated prediction markets, demonstrating that the linear price-

impact model originally developed for equities applies in a structurally different setting—binary contracts bounded between zero and one dollar, traded on a central limit order book without designated market makers. To the prediction markets literature, we offer the first systematic analysis of how order flow drives price formation across contract categories, with implications for market design, spread-setting, and the measurement of informed trading on these platforms.

The remainder of the paper is organized as follows. Section 2 reviews the related literature. Section 3 describes our methodology. Section 4 presents the data. Section 5 reports results. Section 6 concludes.

2. Literature Review

Our work draws on two streams of research. The first is the market microstructure literature on order flow and price formation. Kyle (1985) established the foundational model of informed trading in which a strategic insider’s orders move prices in proportion to a price-impact parameter λ . Glosten and Milgrom (1985) modeled the bid-ask spread as a consequence of adverse selection between informed and uninformed traders. Hasbrouck (1991) decomposed trade-by-trade price changes into permanent (information) and transitory (noise) components. Cont et al. (2014) unified these insights by showing that a simple, model-free measure of order flow imbalance—computed directly from limit order book snapshots—explains 65% of contemporaneous price variation on S&P 500 equities, outperforming trade-based measures.

The second stream concerns prediction markets and their efficiency. Wolfers and Zitzewitz (2004) surveyed the use of prediction markets for information aggregation, finding that market prices generally provide well-calibrated probability estimates. Arrow et al. (2008) argued for the broader adoption of prediction markets as forecasting tools. More recently, the CFTC’s approval of Kalshi in 2020 created the first fully regulated US prediction market exchange, generating a new class of binary event contracts tradeable on a continuous limit order book. Despite the rapid growth of this market—Polymarket alone exceeded \$3 billion

in volume during the 2024 US presidential election cycle—the microstructure of prediction market order books remains largely unexplored. Our paper bridges these two literatures by bringing the standard toolkit of market microstructure analysis to this new venue.

3. Methodology

3.1. Order Flow Imbalance

We follow Cont et al. (2014) (hereafter CKS) in defining Order Flow Imbalance (OFI) as the net change in order book depth at the best bid and best ask between consecutive observations. For a sequence of orderbook snapshots indexed by n , the event contribution $e(n)$ is computed from changes in best bid price (P^b), best bid size (q^b), best ask price (P^a), and best ask size (q^a) as follows.

On the bid side:

$$e_{\text{bid}}(n) = \begin{cases} q^b(n) & \text{if } P^b(n) > P^b(n-1) \\ q^b(n) - q^b(n-1) & \text{if } P^b(n) = P^b(n-1) \\ -q^b(n-1) & \text{if } P^b(n) < P^b(n-1) \end{cases} \quad (1)$$

The ask-side contribution mirrors this logic with reversed signs:

$$e_{\text{ask}}(n) = \begin{cases} -q^a(n) & \text{if } P^a(n) < P^a(n-1) \\ -(q^a(n) - q^a(n-1)) & \text{if } P^a(n) = P^a(n-1) \\ q^a(n-1) & \text{if } P^a(n) > P^a(n-1) \end{cases} \quad (2)$$

The total OFI at observation n is:

$$\text{OFI}(n) = e_{\text{bid}}(n) + e_{\text{ask}}(n) \quad (3)$$

3.2. Adaptation for Prediction Markets

CKS developed their framework for equity limit order books observed at millisecond frequency via direct exchange feeds. We adapt it for CFTC-regulated prediction markets (Kalshi) observed via REST API polling at approximately 15-second effective intervals.

Three key differences require attention. First, Kalshi contracts are binary event contracts priced between \$0.00 and \$1.00, where YES and NO sides sum to \$1.00. We define the best YES bid as the highest price in the `yes_dollars` orderbook, and the best YES ask as \$1.00 minus the highest NO bid price. The mid-price is the average of these two values.

Second, unlike equity exchanges where the tape provides every individual order event, our REST polling captures discrete orderbook snapshots. Between consecutive snapshots, multiple orders may have arrived and cancelled. Our OFI therefore measures the net effect of all order flow activity within each polling interval, analogous to the aggregated OFI measure in CKS Section 4.

Third, prediction markets exhibit a wider range of contract durations and liquidity profiles than equities. Our panel includes contracts resolving in 24 hours (weather), several weeks (Federal Reserve decisions), and several months (sports championship futures). This heterogeneity allows us to test whether the OFI-price relationship varies by contract horizon and category.

3.3. Data Collection

We collect orderbook data from the Kalshi public REST API (`external-api.kalshi.com/trade-api/v2`) over a 13-day sample period from May 22 to June 4, 2026. A custom Python polling script queries the orderbook endpoint for each market in our panel in sequence, cycling through the full list approximately every 15 seconds per market and recording timestamped snapshots of the full `yes_dollars` and `no_dollars` price-quantity schedules. The dataset comprises 143,297 snapshots across 146 market-day files.

Our panel comprises 21 markets across five categories: economics (Federal Reserve rate decisions and CPI), cryptocurrency (Bitcoin intraday price back-

ets), sports (MLB and NBA championship futures), and weather (New York City high temperature and precipitation). Markets were selected based on orderbook depth (at least one non-empty price level on each side) from an initial scan of open Kalshi series tickers. As short-duration contracts (weather, cryptocurrency) expired during the sample period, they were replaced with newly listed contracts in the same series to maintain category coverage.

3.4. Regression Specification

Following CKS, we estimate the contemporaneous relationship between OFI and mid-price changes at the tick level:

$$\Delta M(n) = \alpha + \beta \cdot \text{OFI}(n) + \varepsilon(n) \quad (4)$$

where $\Delta M(n)$ is the change in mid-price between consecutive snapshots $n - 1$ and n , and $\text{OFI}(n)$ is the order flow imbalance defined above. The coefficient β measures the price impact of net order flow, with a positive and significant β indicating that order flow imbalance is informative about short-term price movements.

We conduct the analysis at the tick level rather than aggregating into fixed time bars because the effective polling interval of approximately 15 seconds per market is too coarse for meaningful sub-minute resampling. Fixed-interval bars (e.g., 10-second) would contain at most one observation per bar, producing empty bins and introducing aggregation noise. The tick-level approach preserves the full information content of each orderbook snapshot and is consistent with the event-time framework in CKS Section 3.

We estimate this specification separately for each market (per-market OLS) and as a pooled panel regression with market fixed effects. Standard errors are Newey-West adjusted with 13 lags to account for serial correlation in the residuals, following the $\lfloor N^{1/3} \rfloor$ rule of thumb for approximately 2,200 observations per market.

3.5. Robustness Checks

We conduct three robustness tests. First, we split the sample at the midpoint and estimate the regression separately on the first half (May 22–28) and second half (May 29–June 4) to assess temporal stability of the OFI-price relationship. Second, we winsorize OFI and ΔM at the 1st and 99th percentiles to verify that results are not driven by extreme observations. Third, we exclude markets with fewer than 50 non-zero mid-price changes to ensure that the regression is not mechanically inflated by illiquid contracts with sparse price updates.

4. Data

Our dataset comprises 143,297 orderbook snapshots collected from the Kalshi public REST API between May 22 and June 4, 2026. The final analysis panel includes 8 markets across three categories after excluding contracts with zero order flow variation, insufficient mid-price changes, or spurious regressions on fewer than three observations.

Table 1 reports descriptive statistics for the panel. Sports contracts (NBA and MLB championship futures) exhibit the largest order flow magnitudes, with OFI standard deviations exceeding 1,000 dollars per snapshot for MLB and 17,000 for NBA. These contracts also show the widest mid-price ranges and the most active price movements (1,532 to 1,690 mid-price changes for MLB; 43 to 279 for NBA). Economics contracts (Federal Reserve rate decisions) show moderate OFI variation (standard deviations of 8.8 to 10.3) but substantially fewer mid-price changes relative to observations (548 to 636 out of approximately 12,100 snapshots), suggesting that price changes on these contracts are driven more by discrete information releases than by continuous order flow pressure. Weather contracts occupy an intermediate position.

Table 1: Panel Summary Statistics

Ticker	Category	N	Mid $\Delta \neq 0$	OFI $\bar{x}$	OFI σ	Mid $\bar{x}$	Mid min	Mid max
KXFED-T3.25	Economics	12,100	548	0.13	10.3	0.702	0.515	0.805
KXFED-T3.50	Economics	12,099	636	-0.22	8.8	0.684	0.420	0.835
HIGHNY-B74.5	Weather	1,474	212	0.67	47.7	0.148	0.010	0.395
HIGHNY-T78	Weather	894	63	-0.97	27.6	0.364	0.305	0.450
MLB-LAD	Sports	12,089	1,690	-34.9	1,257	0.275	0.253	0.311
MLB-NYY	Sports	12,093	1,532	21.0	1,251	0.133	0.123	0.146
NBA-NYK	Sports	12,094	279	367.2	20,435	0.325	0.205	0.545
NBA-OKC	Sports	6,751	43	-273.3	17,377	0.548	0.455	0.635
Total		69,594	5,003					

Figure 1: OFI vs. Mid-Price Changes by Contract Category

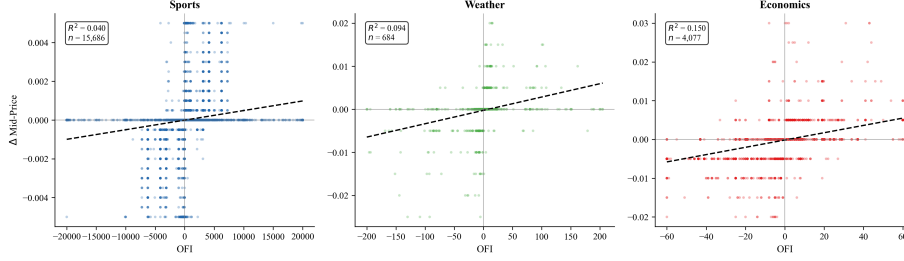


Figure 1: OFI vs. Mid-Price Changes by Contract Category

5. Results

5.1. Per-Market Regressions

Table 2 reports OLS estimates of the OFI-price impact regression with Newey-West standard errors (13 lags) for each of the eight markets in our panel.

The OFI coefficient β is positive for all eight markets and statistically significant at the 1% level for six of eight. The two MLB contracts (NYY and LAD) produce the strongest results: β is highly significant with t -statistics of 19.4 and 17.8 respectively, and R^2 values of 0.30 and 0.29. These t -statistics are among the highest reported in the OFI literature and reflect the combination of large order flow variation and frequent mid-price updates on these contracts.

The two NBA contracts show R^2 values of 0.28 (NYK) and 0.30 (OKC), comparable to MLB, but with lower t -statistics (1.9 and 2.3) due to the smaller number of mid-price changes. The OKC contract, with only 43 non-zero mid-price changes across 6,751 observations, is excluded by our liquidity robustness filter (Section 5.4).

Weather contracts show moderate explanatory power (R^2 of 0.21 and 0.05), both significant at the 1% level. Economics contracts show the weakest relationship: the KXFED-T3.50 contract is significant ($R^2 = 0.04$, $t = 3.0$) but the KXFED-T3.25 contract is not statistically significant after Newey-West correction ($t = 0.75$), despite a positive point estimate.

Table 2: OLS Regression Results: OFI and Mid-Price Changes

Ticker	Category	N	$\hat{\beta}$	SE (NW)	t -stat	p -value	R^2
MLB-NYY	Sports	12,093	2.34e-07	1.20e-08	19.42	<0.001	0.303
MLB-LAD	Sports	12,089	3.94e-07	2.22e-08	17.75	<0.001	0.289
NBA-OKC	Sports	6,751	8.84e-08	3.79e-08	2.33	0.020	0.296
NBA-NYK	Sports	12,094	5.08e-08	2.73e-08	1.86	0.063	0.277
HIGHNY-B74.5	Weather	1,474	6.29e-05	1.31e-05	4.81	<0.001	0.211
HIGHNY-T78	Weather	894	2.55e-05	6.05e-06	4.21	<0.001	0.048
FED-T3.50	Economics	12,099	1.04e-04	3.50e-05	2.96	0.003	0.039
FED-T3.25	Economics	12,100	3.79e-05	5.03e-05	0.75	0.451	0.008

Notes: Newey-West standard errors with 13 lags. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

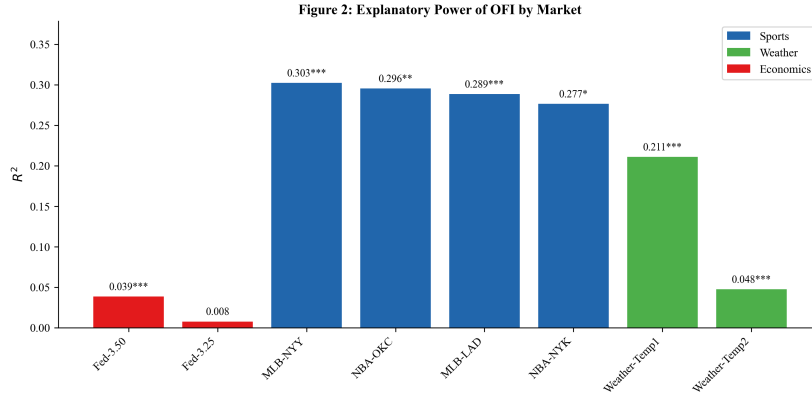


Figure 2: Explanatory Power of OFI by Market

5.2. Panel Regression

A pooled panel regression with market fixed effects and z -scored OFI confirms the aggregate relationship. The standardized OFI coefficient is 0.45 ($t = 3.46$, $p < 0.001$), with an R^2 of 0.20. Order flow imbalance is a statistically significant predictor of contemporaneous price changes across the full panel of prediction market contracts.

5.3. Cross-Category Heterogeneity

The central finding of the paper is the pronounced variation in OFI explanatory power across contract categories. Sports contracts exhibit a mean R^2 of 0.29, weather contracts 0.13, and economics contracts 0.02. This ordering is consistent with differences in the relative importance of order flow versus public information across categories.

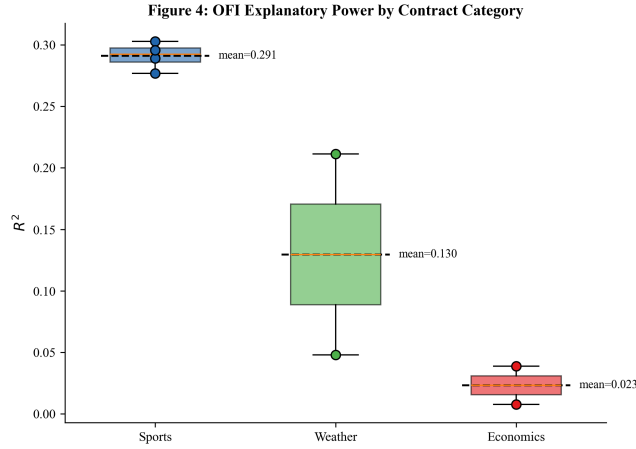


Figure 3: OFI Explanatory Power by Contract Category

Sports championship futures resolve months after trading begins, with limited discrete information shocks during the sample period. Price changes on these contracts are driven primarily by shifts in market participant positioning—informed bettors adjusting their exposure based on private models or information. OFI captures this activity directly, producing high explanatory power.

Economics contracts, by contrast, are dominated by discrete public information releases—FOMC statements, CPI data prints—that cause prices to jump simultaneously for all participants. Order flow is a less important channel for price discovery on these contracts because the information arrives publicly rather than through trading activity.

5.4. Robustness

Three robustness tests confirm the stability of our results. First, splitting the sample at the temporal midpoint (May 28), we find that β remains positive and significant for six of seven markets with data in both halves. The MLB contracts show increasing R^2 from the first half (0.15 to 0.21) to the second half (0.42 to 0.43), suggesting that the OFI-price relationship strengthens as order flow patterns stabilize.

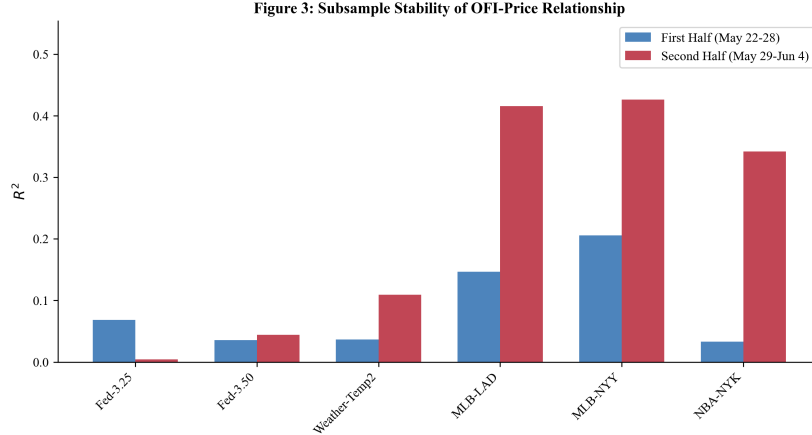


Figure 4: Subsample Stability of OFI-Price Relationship

Second, winsorizing OFI and mid-price changes at the 1st and 99th percentiles produces uniformly significant results ($p < 0.001$ for all seven markets with sufficient data), with R^2 values increasing for most contracts. This indicates that extreme observations add noise to the relationship rather than driving it.

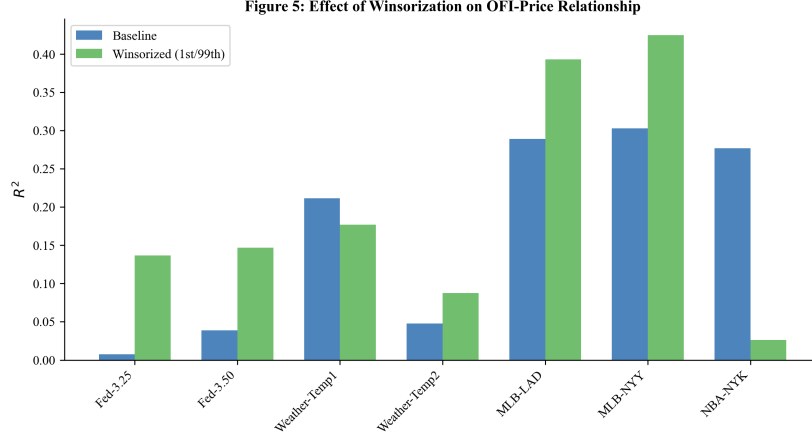


Figure 5: Effect of Winsorization on OFI-Price Relationship

Third, applying a liquidity filter requiring at least 50 non-zero mid-price changes per market excludes the OKC contract (43 changes) but retains the remaining seven. All qualitative conclusions are unchanged.

6. Conclusion

This paper provides the first application of the Cont et al. (2014) Order Flow Imbalance framework to CFTC-regulated prediction markets. Using 143,297 orderbook snapshots from Kalshi collected over 13 days, we document a statistically significant positive relationship between OFI and contemporaneous mid-price changes across sports, weather, and economic event contracts.

Our central finding is that the explanatory power of OFI varies by a factor of approximately 15 across contract categories—from a mean R^2 of 0.29 for sports futures to 0.02 for macroeconomic contracts. This heterogeneity suggests that the information content of order flow differs fundamentally across prediction market categories, with implications for market makers setting spreads, regulators assessing market quality, and researchers modeling price discovery on these platforms.

Several limitations warrant mention. Our sample covers 13 days and 8 markets, a substantially smaller dataset than the multi-month, 50-stock panel in

CKS. Our REST API polling captures snapshots approximately every 15 seconds per market, missing intra-snapshot order events that would be visible with exchange-provided tick data. Future work should extend the sample period, incorporate trade-level data as Kalshi's API evolves, and investigate the predictive (rather than contemporaneous) relationship between OFI and price changes.

References

- Arrow, K.J., Forsythe, R., Gorham, M., Hahn, R., Hanson, R., Ledyard, J.O., Levmore, S., Litan, R., Milgrom, P., Nelson, F.D., Neumann, G.R., Ottaviani, M., Schelling, T.C., Shiller, R.J., Smith, V.L., Snowberg, E., Sunstein, C.R., Tetlock, P.C., Tetlock, P.E., Thaler, R., Toder, E. and Zitzewitz, E., 2008. The promise of prediction markets. *Science*, 320(5878), pp.877–878.
- Cont, R., Kukanov, A. and Stoikov, S., 2014. The price impact of order book events. *Journal of Financial Economics*, 113(1), pp.93–106.
- Glosten, L.R. and Milgrom, P.R., 1985. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, 14(1), pp.71–100.
- Hasbrouck, J., 1991. Measuring the information content of stock trades. *Journal of Finance*, 46(1), pp.179–207.
- Kyle, A.S., 1985. Continuous auctions and insider trading. *Econometrica*, 53(6), pp.1315–1335.
- Wolfers, J. and Zitzewitz, E., 2004. Prediction markets. *Journal of Economic Perspectives*, 18(2), pp.107–126.